United States
Department of
Agriculture

Forest Service

Gen. Tech. Report
WO-79

April 2009

Proceedings of the Eighth Annual Forest Inventory and Analysis Symposium

Monterey, CA • October 16–19, 2006



United States
Department of
Agriculture

Forest Service

Gen. Tech. Report
WO-79

April 2009

Proceedings of the Eighth Annual Forest Inventory and Analysis Symposium

Edited by
Ronald E. McRoberts
Gregory A. Reams
Paul C. Van Deusen
William H. McWilliams

Monterey, CA
October 16–19, 2006

Sponsored by
USDA Forest Service, National Council for Air and Stream Improvement, Society of American Foresters, International Union of Forestry Research Organizations, and Section on Statistics and the Environment of the American Statistical Association

McRoberts, Ronald E.; Reams, Gregory A.; Van Deusen, Paul C.; McWilliams, William H., eds. 2008. 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.

Documents contributions to forest inventory in the areas of sampling, remote sensing, modeling, information management and analysis for the Forest Inventory and Analysis program of the Forest Service.

KEY WORDS: Sampling, estimation, remote sensing, modeling, information science, policy, analysis

Disclaimer

Papers published in these proceedings were submitted by authors in electronic media. Editing was done to ensure a consistent format. Authors are responsible for content and accuracy of their individual papers. The views and opinions expressed in this report are those of the presenters and authors and do not necessarily reflect the policies and opinions of the U.S. Department of Agriculture. 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.

The U.S. Department of Agriculture (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 USDA, Director, Office of Civil Rights, 1400 Independence Avenue, S.W., Washington, D.C. 20250–9410, or call (800) 795–3272 (voice) or (202) 720–6382 (TDD). USDA is an equal opportunity provider and employer.

Preface

The Eighth Annual Forest Inventory and Analysis Symposium was held October 16–19, 2006, in Monterey, CA. The symposium featured participation, including 55 presentations, by scientists from 12 foreign countries. In addition, the trend for participation by scientists from outside the formal Forest Inventory and Analysis program continues to increase. The symposium organizers thank all participants and presenters and convey special thanks to those who submitted their papers for these proceedings.

Ronald E. McRoberts
Gregory A. Reams
Paul C. Van Deusen
William H. McWilliams

Contents

International Forest Inventories

The Italian National Forest Inventory: Geographical and Positioning Aspects in Relation to the Different Phases of the Project
Giacomo Colle, Antonio Floris, Gianfranco Scrinzi, Giovanni Tabacchi, and Lorenzo Cavini.....1

The New Brazilian National Forest Inventory
Joberto V. de Freitas, Yeda M.M. de Oliveira, Doadi A. Brena, Guilherme L.A. Gomide, José Arimatea Silva, José E. Collares, Patrícia P. de Mattos, Maria Augusta D. Rosot, Carlos R. Sanquetta, Marta de F. Vencatto, Paulo L.C. de Barros, João R. dos Santos, Flávio J. Ponzoni, and Yosio E. Shimabukuro 9

Japanese National Forest Inventory and Its Spatial Extension by Remote Sensing
Yasumasa Hirata, Mitsuo Matsumoto, and Toshiro Iehara.....13

The Design of the Second German National Forest Inventory
Gerald Kändler 19

The Brazilian National System of Forest Permanent Plots
Yeda Maria Malheiros de Oliveira, Maria Augusta Doetzer Rosot, Patrícia Povoa de Mattos, Joberto Veloso de Freitas, Guilherme Luis Augusto Gomide, Marta de Fátima Vencato, and Marilice Cordeiro Garrastazú 25

Mapping Martinique's Forests and Other Natural Lands for Land Planning and Development
Rémi Teissier du Cros and Claude Vidal 31

The Finnish National Forest Inventory
Erkki Tomppo 39

Forest Resource Reporting

The Reporting Revolution—The Southern Endeavor
Mark J. Brown 47

Forest Resources of the United States, 2002: Mapping the Renewable Resource Planning Act Data
Cassandra M. Kurtz, Daniel J. Kaisershot, Dale D. Gormanson, and Jeffrey S. Wazenegger.....53

The Influence of International and Domestic Events in the Evolution of Forest Inventory and Reporting Consistency in the United States
W. Brad Smith 59

Analysis

Forest Inventories Generate Scientifically Sound Information on the Forest Resource, But Do Our Data and Information Really Matter?
Christoph Kleinn and Göran Ståhl 71

A Method for Examining the Impacts of Oregon Land Use Laws on Forest Lands and Farmlands
David L. Azuma, Gary Lettman, and Erica Hanson.....79

Research Applications of Ecosystem Patterns
Robert G. Bailey.....83

Developing Survey Grids To Substantiate Freedom From Exotic Pests
John W. Coulston, Frank H. Koch, William D. Smith, and Frank J. Sapio.....91

Predicting the Ability To Produce Emerald Ash Borer: A Comparison of Riparian and Upland Ash Forests in Southern Lower Michigan
Susan J. Crocker, Deborah G. McCullough, and Nathan W. Siegert.....99

The Distribution of Mercury in a Forest Floor Transect Across the Central United States
Charles H. (Hobie) Perry, Michael C. Amacher, William Cannon, Randall K. Kolka, and Laurel Woodruff 103

Dimensionality and the Sample Unit
Francis A. Roesch 109

Assessment of Black Ash *(Fraxinus nigra)* Decline in Minnesota
Kathleen Ward, Michael Ostry, Robert Venette, Brian Palik, Mark Hansen, and Mark Hatfield 115

Biodiversity

Use of Forest Inventory and Analysis Grid-Based Animal Population Data To Develop an Index of Ecological Diversity
Patricia N. Manley, Kristian K. McIntyre, Matthew D. Schlesinger, Lori A. Campbell, Susan Merideth, and Dennis D. Murphy 121

Wildlife Monitoring Across Multiple Spatial Scales Using Grid-Based Sampling
Kevin S. McKelvey, Samuel A. Cushman, Michael K. Schwartz, and Leonard F. Ruggiero 137

Quantifying Forest Fragmentation Using Geographic Information Systems and Forest Inventory and Analysis Plot Data
Dacia M. Meneguzzo and Mark H. Hansen 143

The Relationship Between Diversity and Productivity in Selected Forests of the Lake States Region (USA): Relative Impact of Species Versus Structural Diversity
W. Keith Moser and Mark Hansen 149

Estimating Number and Size of Forest Patches From FIA Plot Data
Mark D. Nelson, Andrew J. Lister, and Mark H. Hansen 159

Using Forest Service Multiple Species Inventory and Monitoring Protocols to Count Birds at Forest Inventory and Analysis Plots on the Caribbean Landscape: Results, Observations, and Challenges From Year 1 of a 2-Year Study
Sonja N. Oswalt, Thomas J. Brandeis, David W. Steadman, and Scott K. Robinson 165

Predicting Bird Habitat Quality From a Geospatial Analysis of FIA Data
John M. Tirpak, D. Todd Jones-Farrand, Frank R. Thompson, III, Daniel J. Twedt, Mark D. Nelson, and William B. Uihlein, III 171

Estimating Dead Wood During National Forest Inventories: A Review of Inventory Methodologies and Suggestions for Harmonization
Christopher W. Woodall, Jacques Rondeux, Pieter J. Verkerk, and Göran Ståhl 179

Estimating the Quadratic Mean Diameter of Fine Woody Debris for Forest Type Groups of the United States
Christopher W. Woodall and Vicente J. Monleon 185

Attributes of Standing Dead Trees in Forests of the United States
Christopher W. Woodall, James E. Smith, and Patrick D. Miles 191

Modeling

Allometric Equations for Predicting Puerto Rican Dry Forest Biomass and Volume
Thomas Brandeis, Matthew Delaney, Larry Royer, and Bernard Parresol 197

Forest Inventory Predictions From Individual Tree Crowns: Regression Modeling Within a Sample Framework
James W. Flewelling 203

External Validation of a Forest Inventory and Analysis Volume Equation and Comparisons With Estimates From Multiple Stem-Profile Models
Christopher M. Oswalt and Adam M. Saunders.....211

Preliminary Results of Spatial Modeling of Selected Forest Health Variables in Georgia
Brock Stewart, Chris J. Cieszewski, and Eric L. Smith.....221

Implementing the Measurement Interval Midpoint Method for Change Estimation
James A. Westfall, Thomas Frieswyk, and Douglas M. Griffith.....231

Remote Sensing and Mapping

Mapping Forest Inventory and Analysis Data Attributes Within the Framework of Double Sampling for Stratification Design
David C. Chojnacky, Randolph H. Wynne, and Christine E. Blinn.....237

Precise FIA Plot Registration Using Field and Dense LIDAR Data
Demetrios Gatziolis.....243

Exploring the Association of the Minnesota Department of Natural Resources' Satellite-Detected Change With the Forest Inventory and Analysis System of Observed Removals and Mortality
Dale D. Gormanson, Timothy J. Aunan, Mark H. Hansen, and Michael Hoppus.....251

The Utility of the Cropland Data Layer for Forest Inventory and Analysis
Greg C. Liknes, Mark D. Nelson, Dale D. Gormanson, and Mark Hansen.....259

Landscape Scale Mapping of Forest Inventory Data by Nearest Neighbor Classification
Andrew Lister.....265

Estimation of a Cover-Type Change Matrix From Error-Prone Data
Steen Magnussen.....273

Bridging the Gap Between Strategic and Management Forest Inventories
Ronald E. McRoberts.....279

Estimating Uncertainty in Map Intersections
Ronald E. McRoberts, Mark A. Hatfield, and Susan J. Crocker.....289

Sampling and Mapping Forest Volume and Biomass Using Airborne LIDARs
Erik Næsset, Terje Gobakken, and Ross Nelson.....297

Forest-Cover–Type Separation Using RADARSAT-1 Synthetic Aperture Radar Imagery
Mark D. Nelson, Kathleen T. Ward, and Marvin E. Bauer.....303

Rapid Mapping of Hurricane Damage to Forests
Eric M. Nielsen.....307

LIDAR Forest Inventory With Single-Tree, Double- and Single-Phase Procedures
Robert C. Parker and David L. Evans.....317

Cartographic Standards To Improve Maps Produced by the Forest Inventory and Analysis Program
Charles H. (Hobie) Perry and Mark D. Nelson.....325

Automatic Crown Cover Mapping To Improve Forest Inventory
Claude Vidal, Jean-Guy Boureau, Nicolas Robert, Nicolas Py, Josiane Zerubia, Xavier Descombes, and Guillaume Perrin.....333

The Finnish Multisource National Forest Inventory: Small-Area Estimation and Map Production
Erkki Tomppo.....341

FIA Operations

Prefield Methods: Streamlining Forest or Nonforest Determinations To Increase Inventory Efficiency
Sara Goeking, Gretchen Moisen, Kevin Megown, and Jason Toombs 351

Is There a Better Metric Than Site Index To Indicate the Productivity of Forested Lands?
Maria E. Blanco Martin, Michael Hoppus, Andrew Lister, and James A. Westfall.....355

Improving Forest Inventory and Analysis Efficiency With Common Land Unit Information
Greg C. Liknes and Mark D. Nelson 363

Phase 2 and Phase 3 Presentation Grids
Joseph M. McCollum and Jamie K. Cochran.....367

Software Development

New Functions and Programs in Hypermap Software Development for Internet-Based Displaying of FIA Data
Chris J. Cieszewski, Roger C. Lowe, Shangbin Liu, Ingvar Elle, and Daniel Markewitz.....375

Next-Generation Simulation and Optimization Platform for Forest Management and Analysis
Antti Mäkinen, Jouni Kalliovirta, and Jussi Rasinmäki 393

SOLE: Enhanced FIA Data Analysis Capabilities
Michael Spinney and Paul Van Deusen 401

The Italian National Forest Inventory: Geographical and Positioning Aspects in Relation to the Different Phases of the Project

Giacomo Colle[1], Antonio Floris[2], Gianfranco Scrinzi[3], Giovanni Tabacchi[4], and Lorenzo Cavini[5]

Abstract.—In this article, we describe in depth the analysis and solutions to manage the multiple coordinates of the sampling objects coming from the three different phases of the second Italian national forest inventory (Inventario Nazionale delle Foreste e dei serbatoi forestali di Carbonio [INFC]). In particular, this article describes the criteria used to determine the sample point coordinates of the first phase, the Global Positioning System (GPS) positioning procedure of the second phase, and the GPS-aided retrieval procedure of sample points during the third phase of the INFC. Because about two-thirds of the Italian territory is hilly or mountainous, with rough orography or with tree coverage, this article also provides a nationwide analysis and survey on the use of GPS technology in forest and natural environments.

Introduction

The Inventario Nazionale delle Foreste e dei serbatoi forestali di Carbonio (INFC) is aimed at updating the data on Italian forests in a way consistent with the current international definitions and commitments detailed in the statements of the Ministerial Conference on the Protection of Forests in Europe (FAO 2000, MCPFE 2003) in relation to the Kyoto Protocol to the United Nations Framework Convention on Climate Change (UNFCCC 1997, 2002). The INFC pays particular attention to the quality of the collected data to obtain a good accuracy and a high precision of results at both the national level and the regional level. This goal was achieved with three-phase sampling for stratification (Fattorini *et al.* 2006), which allowed us to apply a high sampling intensity in the first two phases of the inventory and with the continuous assistance of the surveyors and a careful control of the collected data. In addition, various information sources were used to collect data regarding many aspects of the forest ecosystems (Tabacchi *et al.* 2005, 2007).

This particular sampling design required significant attention to geographical and positioning aspects, due to the necessity to reach twice the ground sample points, during the second and third phases, ensuring that the plot center was in the very same position, to refer the data collected in these two different surveys to the same area.

At the same time, every effort was made to avoid a navigation and positioning procedure too heavy with respect to the time and human resources needed for the whole survey (data collecting and processing).

[1] Research Assistant, Istituto Sperimentale per l'Assestamento Forestale e per l'Alpicoltura, ISAFA-CRA, Piazza Nicolini 6, 38050 Villazzano (Trento), Italy. Phone: +39 0461 381117; fax: +39 0461 381131. E-mail: giacomo.colle@isafa.it.

[2] Forest Technician, Istituto Sperimentale per l'Assestamento Forestale e per l'Alpicoltura, ISAFA-CRA, Piazza Nicolini 6, 38050 Villazzano (Trento), Italy. Phone: +39 0461 381113; fax: +39 0461 381131. E-mail: scrinzi@isafa.it.

[3] Head of Forest Management Department, Istituto Sperimentale per l'Assestamento Forestale e per l'Alpicoltura, ISAFA-CRA, Piazza Nicolini 6, 38050 Villazzano (Trento), Italy. Phone: +39 0461 381117; fax: +39 0461 381131. E-mail: floris@isafa.it.

[4] Head of Forest Biometrics Department/INFC Project Leader, Istituto Sperimentale per l'Assestamento Forestale e per l'Alpicoltura, ISAFA-CRA, Piazza Nicolini 6, 38050 Villazzano (Trento), Italy. Phone: +39 0461 381112; fax: +39 0461 381131. E-mail: giovanni.tabacchi@entecra.it.

[5] Independent Cartographer. E-mail: lorenzocavini@katamail.com.

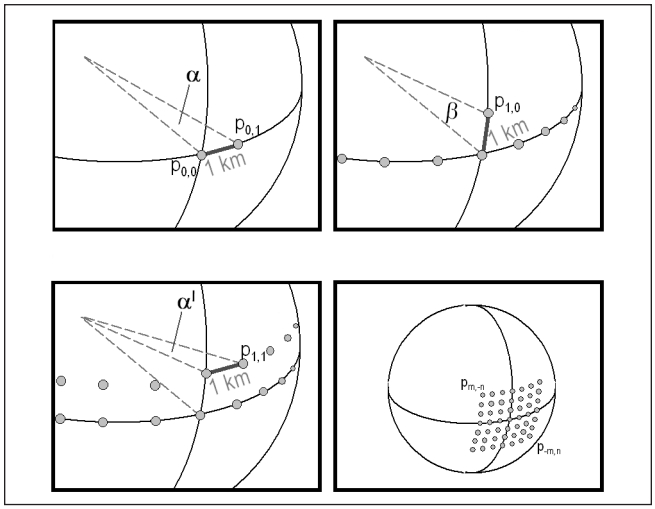
Methods

Production of the National Sampling Point Set

The three-phase INFC sampling inventory is established on a set of sampling points covering the whole Italian territory (30,132,845 ha), each of which represents 1 square km of the land. The proposed method achieves this result using a regular grid covering the whole Italian territory; unlike usual methods, the grid is drawn on the World Geodetic System 1984 (WGS-84) ellipsoid and each sample point is randomly chosen within one of the 1-x-1-km "elliptic" grid squares (Cavini *et al.* 2003).

The first step of the grid production sets the starting point $p_{0,0}$ in Monte Mario in Rome, Italy. The point $p_{0,0}$ is the first point of the sampling grid and determines its central meridian and parallel. The next point of the grid, $p_{0,1}$, is located 1 km east along the parallel passing through $p_{0,0}$. The latitude in degrees of $p_{0,1}$ is the same of $p_{0,0}$; the longitude in degrees of $p_{0,1}$ is the longitude of $p_{0,0}$ plus the central angle α corresponding to the 1-km ellipsoidal arc $p_{0,0}p_{0,1}$, where α is function of the latitude of $p_{0,0}$. In the same way, all the points $p_{0,2},...,p_{0,n}$, and similarly $p_{0,-1},...,p_{0,-n}$, are located along the parallel passing on $p_{0,0}$ (fig. 1). The second step of the procedure generates a new point $p_{1,0}$ located 1 km north from $p_{0,0}$ along the central meridian. The longitude of $p_{1,0}$ is the same of $p_{0,0}$; the latitude of $p_{1,0}$ is the latitude of $p_{0,0}$ plus the central angle β corresponding to the 1-km ellipsoidal arc $p_{0,0}p_{1,0}$ on the central meridian, where β is function of the latitude. The points $p_{1,1},...,p_{1,n}$ and $p_{1,-1},...,p_{1,-n}$ along the parallel passing from $p_{1,0}$ are generated in the same way described previously, but now the angle α^I, corresponding to this step, is determined with the latitude of $p_{1,0}$. The procedure is repeated moving north and south along the central meridian; the resulting set of points $p_{m,-n},...,p_{-m,n}$ determines a nationwide sampling grid formed by same-size elliptic squares.

Figure 1.—*Production of the national sampling grid.*



The procedure to randomly generate the geographic coordinates of the sample point inside each grid square is strongly related to the choice of the national sampling grid plotted on the WGS84 ellipsoid. In proximity of the central meridian, grid squares have a true square shape; moving away from the central meridian, their shape gradually changes to a parallelogram with horizontal sides and height of 1 km still having an area of 1 square km (fig. 2). This change in shape occurs because the central angle α corresponding to a 1-km ellipsoidal arc on a grid parallel is function of the latitude and so varies parallel by parallel. Also, the central angle β corresponding to a 1-km ellipsoidal arc on a grid meridian is function of the latitude, but it is constant for all the squares of a grid row. Thus, the procedure to randomly extract a geographic coordinate (φ,λ) inside the square with low-left grid node $p_{i,j}$ first calculates the latitude φ by randomly generating an arc length between 0 and 1^- km, converting it in the corresponding central angle, and summing it with the latitude φ_i of $p_{i,j}$ (fig. 3). The procedure then extracts the longitude λ by randomly generating a new arc length between 0 and 1^- km, summing it with the arc distance D_j of $p_{i,j}$ from the central meridian (corresponding to j km), converting the resulting arc in the corresponding central angle using the previously calculated latitude φ, and finally summing this result with the longitude of the central meridian (fig. 4).

Figure 2.—*Differences in elliptic squares shapes.*

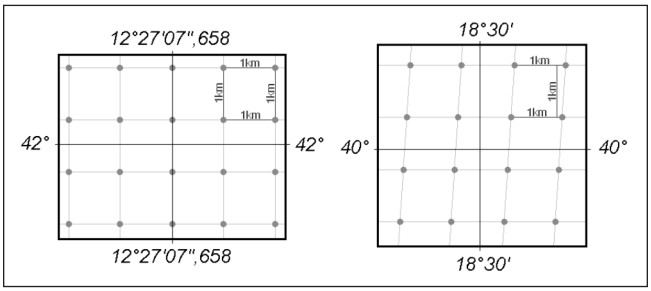


Figure 3.—*Random extraction of sample point latitude (grid squares distortion is exaggerated).*

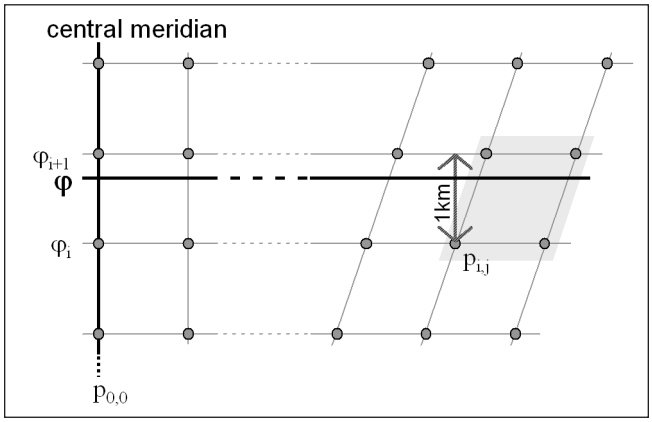
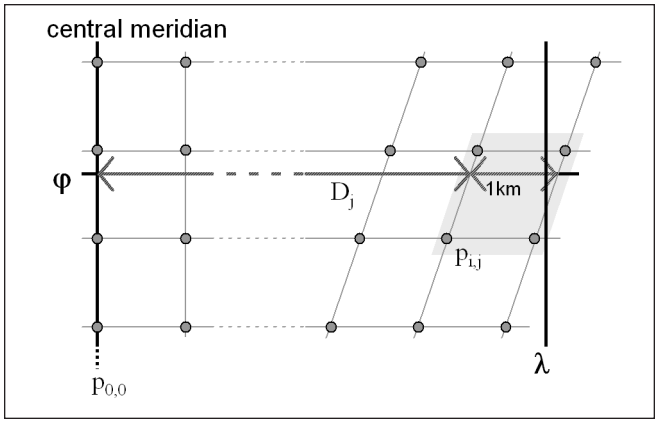


Figure 4.—*Random extraction of sample point longitude (grid squares distortion is exaggerated).*



At the end of the procedure, only sample points inside Italian boundaries were kept, forming the whole sampling set of 301,329 points. Thereafter, the geographic WGS84 coordinates were transformed in the other systems involved in INFC, with the following priority to avoid loss of precision in the various steps: from geodetic system WGS84 to cartographic system UTM-WGS84, with a precision on the order of 10^{-3} m; from geodetic system WGS84 to the geodetic national systems Roma40 and ED50; and from cartographic system UTM-WGS84 to the cartographic national systems Gauss-Boaga and UTM-ED50.[6]

Land Navigation and GPS Positioning in Phase 2 and in Phase 3

Positioning and Data Collection Instruments

Trimble GPS Pathfinder® ProXR and Trimble GeoExplorer® GeoXT receivers were chosen after a comparative test among five Global Positioning System (GPS) receivers of Geographic Information System-mapping class that were eligible to be used in the INFC surveys (Scrinzi *et al.* 2003). The test outcome also provided the expected values of positioning uncertainty (table 1) considering stationary positioning with an average of 170 to 200 single fixings (1 position per second). The uncertainty values under the stand-alone positioning, in particular, were considered sufficient to localize on the ground the phase 3 sample points surveyed in phase 2, by means of a standard GPS navigation procedure (aided by description notes, drafts, and photos taken in phase 2). Therefore, all positioning during the ground surveys of INFC was performed by collecting

Table 1.—*Estimated performance (uncertainty positioning in meters) of the two Global Positioning System receivers chosen.*

Positioning probability (%)	Stand-alone positioning		Postprocessed DGPS	
	Pro XR receiver (meters)	GeoXT receiver (meters)	Pro XR receiver (meters)	GeoXT receiver (meters)
50.0	1.4	3.1	0.9	2.0
66.6	2.0	4.5	1.3	2.8
90.0	3.5	7.7	2.2	4.8

DGPS = Differential Global Positioning System.

[6] All these elaborations were carried out using the software program Cartlab2® (Cima 2002).

and averaging at least 170 to 200 positions and registering a Trimble Standard Storage Format (SSF) file for postprocessed differential correction.

The Pro XR receiver was equipped with an Intermec 700 data logger and its L1 external antenna; the GeoXT receiver was equipped with its external patch antenna. Three possible settings of GPS filter values were chosen: best accuracy, standard, and best performance.[7] The surveyors were suggested to use the highest quality configuration when possible and to perform the GPS calendar planning periodically during the sampling campaign.

Two software applications, INFOR2 and NAV3/RAS3, were expressly developed for phases 2 and 3 of INFC. Each application runs on Microsoft® Windows CE Pocket PC handhelds and provides several ad hoc navigation and positioning functions integrated with the Trimble GPS technology. INFOR2 and NAV3/RAS3 also provide data processing and database storage functionalities for qualitative classifications and quantitative measurements specific for their own inventory phase.

Procedure for Land Navigation and GPS Positioning in Phase 2

During the second phase of INFC, 30,000 sample points were surveyed to discriminate forests from other wooded lands, identify different forest types, and collect information about other qualitative attributes of forest stands (Gasparini and Tosi 2004). More than 100 crews made up of national and local Forest Service personnel were involved in the data collection. The effort needed for this phase required the definition of a detailed procedure to locate on the ground the sample area center, called C, with the best accuracy and collect the necessary positioning data for its retrieval in the following phase of the inventory (Scrinzi and Floris 2004). The procedure was designed to achieve the utmost objectivity in the localization of C and it consists of two subsequent steps depending on the distance of the surveyor to C. This method was inspired by the ground sampling procedures adopted by the British Columbia Resource Inventory Committee for the vegetation resources inventories. (BCRIC 2001, 2002).

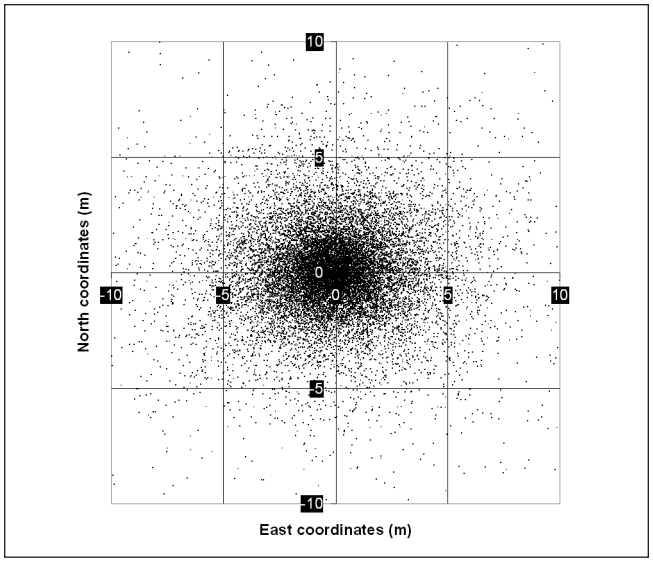
In the first step, when the surveyor is far from C, a standard navigation procedure is used: distance and bearing to C are obtained from an instantaneous GPS position and printed digital orthophotos and cartographic maps help the surveyor to navigate toward C. If the GPS signal is obstructed by the orography of the terrain or by a high canopy density, the procedure allows the surveyor to perform a traditional *open traverse* starting from a known coordinate location: the surveyor works with a compass and a laser rangefinder and clinometer, and a special functionality of the INFOR2 software calculates the coordinates of each traverse station and the distance and bearing to C.

The second step starts when the surveyor is about 15 to 25 m from C; here, the location of C involves a second point, F, chosen subjectively near C where the forest crowns allow the best transparency to the GPS signal and characterized by a nearby existing object (like a tree or a stone). Point F is marked with a temporary stake hidden under the ground and its coordinates are obtained with an averaged GPS positioning. With the surveyor standing at point F, C is located on the ground by calculating the distance d and the azimuth α from the GPS (average) coordinates of F to the theoretical coordinates of C (fig. 5).

Afterwards, a temporary stake is placed at C and its coordinates are collected with GPS positioning as well. These coordinates differ from the theoretical coordinates of C due to the instrumental positioning error of GPS and the error in measuring distance d and azimuth α from F to C; this new point is called C_{GPS} and becomes the true sample plot center for phase 2 surveys. Additional markers like metal plaques are placed near C and F to assist in the phase 3 retrieval of the stakes.

[7] Best accuracy—maximum Position Dilution of Precision (PDOP) of 4, minimum elevation of 15 degrees, and minimum Signal-to-Noise Ratio (SNR) of 6 Amplitude Measurement Units (AMUs). Standard—maximum PDOP of 8, minimum elevation of 12 degrees, and minimum SNR of 4 AMUs. Best performance—maximum PDOP of 12, minimum elevation of 12 degrees, and minimum SNR of 3 AMUs.

Figure 5.—*Scatterplot of C_{GPS} around C.*



Procedure for Sample Point Retrieval in Phase 3

The third phase of INFC, carried out during 2006, implied a further national ground survey to collect dendrometric measurements. The phase 3 points were chosen inside the second phase sample set and stratified by administrative region and forest type (Tabacchi *et al.* 2007). The result of the selection was a sample subset of 6,865 points, each of which still had the sample plot center in C_{GPS}. A retrieval procedure was set up to relocate on the ground the stake placed in C_{GPS} at the end of the previous phase. To perform the retrieval procedure, the surveyor first navigates toward the object of point F using the GPS receiver and locates the F stake; the surveyor then may locate C_{GPS} stake using the distance d and the azimuth α from point F to C. Other quick procedures using the marks placed on the ground were set up to find C_{GPS} stake in different ways. Moreover, if the surveyor misses point F, he or she could perform near C a new average GPS positioning to get accurate values of distance and bearing to C and then retrieve the C_{GPS} stake. All retrieval procedures are supported by a metal detector.

If none of the procedures for C_{GPS} stake retrieval succeeds, the procedure used in phase 2 to locate C from point F must be executed again in phase 3. The new located point C_{GPS} will need new average GPS positioning and becomes the new center of the sample plot. This positioning procedure was also repeated if the Trimble SSF file was missing or corrupted or if the average coordinates of C_{GPS} did not satisfy some given quality requirements (at least 170 positions and a maximum distance between C_{GPS} and C of 20 m). At the end of the procedure, a permanent stake is placed in C_{GPS} to enable the retrieval of the sample point in the future.

GPS Positioning Quality Results

No real-time differential correction procedures were performed during the phase 2 and phase 3 surveys. Real-time Differential GPS procedures, in fact, could not be performed in many Italian forest scenarios due to the weak European Geostationary Navigation Overlay Service[8] signal under tree coverage and in mountainous environments and to the difficulties to receive (for the same reasons) other broadcast signals via radio or cellular phone connections (European Space Agency 2006). On the other hand, positioning accuracy obtained by means of standard mode had already been estimated as sufficient for our main operational purposes (Scrinzi *et al.* 2000), which consisted of finding again in phase 3 the sample points positioned in phase 2. Moreover, this assumption has been confirmed during the third phase field campaign, which is currently in progress: of a total of 6,865 sample points, 6,724 have been surveyed so far and about 98.8 percent of them have been exactly relocated without any particular problem. Therefore, all the following analyses are referred to stand-alone positioning data. By the way, postprocessed differential correction using the European Reference Frame Permanent Network will be performed on phase 3 points, which, at the end of the phase 3 surveys, is supposed to set up a permanent network.

At the end of phase 2, the postprocessing of Trimble SSF files permitted a deep analysis of GPS positioning quality. Approximately 74 percent of phase 2 sample points had SSF C_{GPS} coordinates valid for the analysis. Most of those sample points excluded (21 percent) did not satisfy the 170 positions

[8] European Geostationary Navigation Overlay Service correction data were not fully available until July 2006.

constraint; some of them (5 percent) were provided with a corrupted or missed SSF file, and only few (less than 0.25 percent) were discarded because they had a theoretical outside distance between C_{GPS} and C of more than 20 m.

The first analysis on C_{GPS} coordinates was the measurement of their distance from the C theoretical coordinates; although this analysis is not a real performance test (C is not a tangible object with known and precise coordinates and the surveyors located C using nondifferential corrected GPS positioning in point F), the regular scatterplot of C_{GPS} around C (fig. 6) proves the absence of systematical errors in the phase 2 positioning procedure. In addition, almost 90 percent of C_{GPS} points have a distance from C of less than 5 m (fig. 7), indicating a high proximity of sample area center of phase 2 to the photo interpretation site of phase 1. The analysis on the distance of C_{GPS} from C coordinates showed also that the Pro XR receiver performed better than the GeoXT receiver (fig. 7), albeit the Pro XR was generally assigned to surveys characterized by severe terrain and forest conditions.

The SSF files provided also information about the maximum Horizontal Dilution of Precision (HDOP) of positioning in C_{GPS}: approximately half of the positioning was performed with high horizontal accuracy values (HDOP of less than or equal to 3) and more than 90 percent with fair horizontal accuracy (HDOP of less than or equal to 6) (fig. 8). Moreover, the comparison of the maximum HDOP in C_{GPS} and in F positioning showed that HDOP values registered in F were generally lower than those registered in C_{GPS}, proving a correct procedure for the point F location choice.

Figure 6.—*Distance of C_{GPS} from C.*

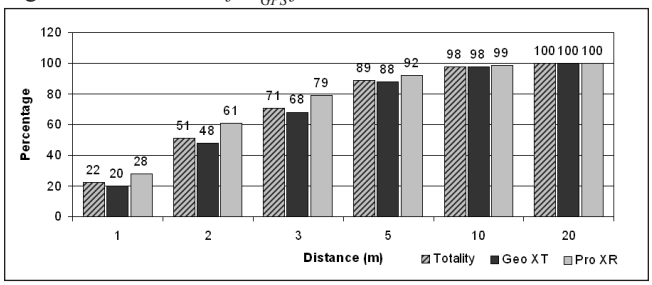


Figure 7.—*Maximum Horizontal Dilution of Precision in C_{GPS}.*

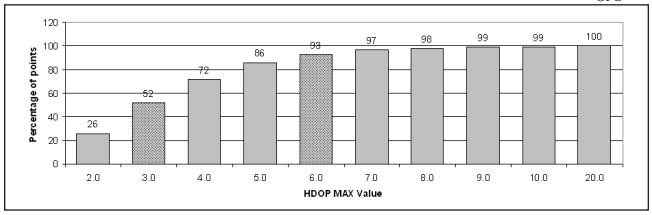
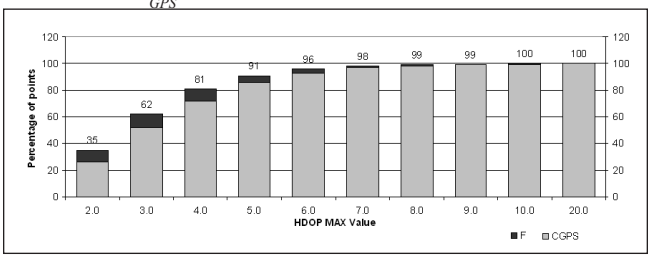


Figure 8.—*Comparison of maximum Horizontal Dilution of Precision in C_{GPS} and in F.*



Discussion

Methods and results previously described suggest some considerations. The production of the national sampling point set was quite hard, but the adopted solution provides two major advantages with respect to standard solutions.[9] First, the appreciable reduction of the area distortion in the eastern regions of Italy (up to 8,000 m^2 per square kilometer) compared to the one obtainable using a unique projected Gauss-Boaga/Roma40 square grid. The "elliptic" grid, in fact, is formed of squares with a constant area of 1 square km on the ellipsoid, each one having the same measurement error. The second advantage is that generating sample point coordinates in the WGS84 geographic system ensures their direct usage in GPS surveys.

Regarding the navigation and positioning procedures adopted in phase 2 and phase 3, the surveys highlighted the almost trouble-free use of GPS receivers in forest environments. Only in very few situations, in fact, the best performance GPS setting had to be used, and the use of the open traverse method in substitution of the GPS-driven one was exceptional (24 cases out of 22,289). The use of point F in the adopted procedures

[9] The two mostly used cartographic systems in Italy are Gauss-Boaga/Roma40 (digital orthophotos and regional maps) and UTM (Universal Transverse Mercator)/ED50 (European Datum of 1950) (national topographic map).

allowed an objective location of point C; moreover, the several markers placed around them guaranteed the invisibility of C_{GPS} position and minimized the time wasted on average GPS positioning during phase 3 surveys.

Regarding the performances of the different GPS receivers used, the Pro XR receiver showed better positioning precision than the GeoXT receiver did. In addition, the Pro XR receiver operated more efficiently under tree coverage, probably because of its specific hardware features and better antenna, and experienced fewer system failures in the field than the GeoXT receiver did. According to our experience, it would be better to use an instrumental configuration with separate devices for the GPS antenna and for the personal digital assistant (PDA) component, although an integrated device like the GeoXT receiver could offer advantages in terms of weight and portability. Moreover, our experience put in evidence that PDAs still show difficulties in working in real multitasking mode. This pitfall was one of the reasons for having two separate software applications in phase 3: one for navigation and positioning and another for collecting attributes.

INFC has been the first Italian experience of using GPS technology in natural environment assessment on a nationwide scale. Data collected during this survey are available for further analyses and can be useful for improving positioning specifications and protocols in future field campaigns. Although a prudential criterion of collecting at least 170 to 200 positions for each point feature was used in INFC, in fact, it is reasonable to assume that this value could be significantly reduced after data distribution analyses.

GPS can be nowadays considered a mature technology in forest assessment applications at a certain scale, such as navigating to a target plot center or positioning stand boundaries. So far, it still is not possible, however, to determine the position of any single tree inside a stand with sufficient accuracy, using only GPS technology.

On this matter, significant improvements are expected in coming years as a result of GALILEO full operational deployment (European Commission 2006) and its integration with other Global Navigation Satellite System services, especially GPS.

Acknowledgments

The authors thank Giuseppe Cappalonga and Laura Marzullo for their help in reviewing this article.

Literature Cited

British Columbia Resource Inventory Committee (BCRIC). 2001. Ministry of Forests procedures and guidelines for operational forest resource survey and mapping using global positioning system technology. British Columbia: British Columbia Resource Inventory Committee. http://ilmbwww.gov.bc.ca/risc/. (30 January 2009).

BCRIC. 2002. Vegetation resources inventory ground sampling procedures. British Columbia: British Columbia Resource Inventory Committee. http://ilmbwww.gov.bc.ca/risc/. (30 January 2009).

Cavini, L.; Scrinzi, G.; Floris, A. 2003. Procedure topografiche di definizione delle coordinate dei punti di campionamento. Trento, Italy: Inventario Nazionale delle Foreste e dei Serbatoi Forestali di Carbonio. Ministero per le Politiche Agricole e Forestali, Direzione Generale Risorse Forestali Montane Idriche Corpo Forestale dello Stato, Consiglio per la Ricerca e la Sperimentazione in Agricoltura, Istituto Sperimentale per l'Assestamento Forestale e per l'Alpicoltura. http://mpf.entecra.it/?q=node/615. (30 January 2009).

Cima, V. 2002. Cartlab1, uno strumento pratico per la conversione di coordinate. Rome, Italy. MondoGIS, Year VII, 30: 10.

European Commission. 2006. GALILEO European satellite navigation system Web site. http://ec.europa.eu/dgs/energy_transport/galileo/index_en.htm. (30 January 2009).

European Space Agency. 2006. European Geostationary Navigation Overlay Service Web site. http://www.esa.int/esaNA/egnos.html. (30 January 2009).

Fattorini, L.; Marcheselli, M.; Pisani, C. 2006. A three-phase sampling strategy for large-scale multiresource forest inventories. Journal of Agricultural, Biological and Environmental Statistics. 11: 296-316.

Food and Agriculture Organization of the United Nations (FAO). 2000. Global forest resources assessment. FAO forestry paper 140. Rome, Italy: Food and Agriculture Organization of the United Nations. ftp://ftp.fao.org/docrep/fao/003/Y1997E/FRA%202000%20Main%20report.pdf.

Gasparini, P.; Tosi, V. 2004. Istruzioni per il rilievo degli attributi di seconda fase. Trento, Italy: Inventario Nazionale delle Foreste e dei Serbatoi Forestali di Carbonio. Ministero per le Politiche Agricole e Forestali, Direzione Generale Risorse Forestali Montane Idriche Corpo Forestale dello Stato, Consiglio per la Ricerca e la Sperimentazione in Agricoltura, Istituto Sperimentale per l'Assestamento Forestale e per l'Alpicoltura. http://www.isafa.it/scientifica/pubblicazioni/pu_infc/ALGT/Ri_Attr/RiAttrb.pdf.

Ministerial Conference on the Protection of Forests in Europe (MCPFE). 2003. Relevant definitions used for the improved Pan-European indicators for sustainable forest management. Vienna, Austria: Ministerial Conference on the Protection of Forests in Europe, Liaison Unit Vienna. 21 p.

Scrinzi, G.; Floris, A. 2004. Procedure di individuazione, materializzazione e ritrovamento dei punti di campionamento. Trento, Italy: Inventario Nazionale delle Foreste e dei Serbatoi Forestali di Carbonio. Ministero per le Politiche Agricole e Forestali, Direzione Generale Risorse Forestali Montane Idriche Corpo Forestale dello Stato, Consiglio per la Ricerca e la Sperimentazione in Agricoltura, Istituto Sperimentale per l'Assestamento Forestale e per l'Alpicoltura. http://www.isafa.it/scientifica/pubblicazioni/pu_infc/ALGT/Navig.pdf.

Scrinzi, G.; Floris, A.; Galvagni, D.; Marzullo, L. 2003. Un metodo di valutazione della precisione di ricevitori GPS, o di metodi di posizionamento, nei rilievi forestali. Linea Ecologica. 2: 11-18.

Scrinzi, G.; Floris, A.; Picci, M. 2000. GPS and forest inventories: accurate positioning and finding of sample points. Comunicazioni di Ricerca dell'Istituto Sperimentale per l'Assestamento Forestale e per l'Alpicoltura. Trento, Italy: Istituto Sperimentale per l'Assestamento Forestale e per l'Alpicoltura. 11 p.

Tabacchi, G.; De Natale, F.; Floris, A.; Gagliano, C.; Gasparini, P.; Scrinzi, G.; Tosi, V. 2007. Italian National Forest Inventory: methods, state of the project, and future developments. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C.; McWilliams, W.H., eds. Proceedings, seventh annual forest inventory and analysis symposium. Gen. Tech. Rep. WO-77. Washington, DC: U.S. Department of Agriculture, Forest Service: 55-66.

Tabacchi, G.; De Natale, F.; Floris, A.; Gasparini, P.; Scrinzi, G.; Tosi, V. 2005. The new Italian National Forest Inventory: a multi-source and multi-resource survey. In: Proceedings, International Union of Forest Research Organizations World Congress. Vienna, Austria: International Union of Forest Research Organizations.

United Nations Framework Convention on Climate Change (UNFCCC). 1997. Kyoto protocol to the United Nations framework convention on climate change. FCCC/CP/1997/7/Add.1. Bonn, Germany.

UNFCCC. 2002. Report of the Conference of the Parties on its seventh session. Addendum. Part 2: Action taken by the conference of the parties. FCCC/CP/2001/13/Add.1. Bonn, Germany.

The New Brazilian National Forest Inventory

Joberto V. de Freitas[1], Yeda M.M. de Oliveira[2], Doadi A. Brena[3], Guilherme L.A. Gomide[4], José Arimatea Silva[5], José E. Collares[6], Patrícia P. de Mattos[7], Maria Augusta D. Rosot[7], Carlos R. Sanquetta[8], Marta de F. Vencatto[7], Paulo L.C. de Barros[9], João R. dos Santos[10], Flávio J. Ponzoni[10], and Yosio E. Shimabukuro[10]

Abstract.—The new Brazilian national forest inventory (NFI) is being planned to be carried out through five components: (1) general coordination, led by the Brazilian Forest Service; (2) vegetation mapping, which will serve as the basis for sample plot location; (3) field data collection; (4) landscape data collection of 10 x 10-km sample plots, based on high-resolution satellite imagery interpretation; and (5) associated programs. The mapping will be based on 1:250,000-scale topographic maps and China-Brazil Earth Resource Satellite images. The standard sample plot distribution will be based on a systematic grid of 20 x 20 km. The NFI will be held based on a 5-year measurement cycle.

Introduction

In the 1980s, Brazil carried out its first and unique national forest inventory (NFI), which aimed at producing information about timber stocks of planted and natural forests (Brena 1995, Machado 1984) as most of the earliest national inventories did (Holmgren and Persson 2002). Since then, only regional forest inventories have been carried out to attend to particular demands on information such as government planning strategies. More recently, some States have taken the initiative in setting up their State forest inventories aiming at monitoring forest resources. These initiatives, however, are completely independent in terms of methodologies and timing. Despite the fact that States' initiatives are positive, and potentially more detailed, it is argued that, ideally, an NFI is the most appropriate alternative to produce information on forest resources at the national level.

Among the motivations that led the Ministry of Environment to propose a new NFI certainly is the forest resource strategic importance, both for the country and at the global level, as well as the lack of reliable information at the national level. Brazil is the largest country in Latin America, occupying 8.8 million square km, of which approximately 4.8 million square km are covered by forests (FAO 2005). Despite the importance of forest resources, the country does not have a regular national forest assessment to support public forest policy formulation aimed at forest conservation and sustainable use.

The process of designing a project for the new NFI started in 2005, when the Ministry of Environment carried out a national workshop to identify the main components and methodological approaches to be considered in the project. A technical committee was then designated to coordinate a participatory approach to set up a nationwide project. Therefore, the conceptual basis for the project took into consideration the contributions of experts and groups of interest from different institutions and regions through workshops as well as international collaboration with more experienced countries. A second national workshop, held in December 2006, presented the first version of the project.

[1] Project Leader, Brazilian Forest Service, SCEN, Trecho 2, Lote 4, Bloco H, 70068-090, Brasilia, Distrito Federal, Brazil. E-mail: joberto.freitas@florestal.gov.br.
[2] Researcher, Embrapa Florestas, Curitiba, Paraná, Brazil.
[3] Professor, Universidade Federal de Santa Maria, Santa Maria, Rio Grande do Sul, Brazil.
[4] Technician, Brazilian Forest Service, Brasilia, Distrito Federal, Brazil.
[5] Professor, Universidade Federal Rural do Rio de Janeiro, Seropédica, Rio de Janeiro, Brazil.
[6] Researcher, Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro, Brazil.
[7] Researcher, Embrapa Florestas, Curitiba, Paraná, Brazil.
[8] Professor, Universidade Federal do Paraná, Curitiba, Paraná, Brazil.
[9] Professor, Universidade Federal Rural da Amazônia, Belém, Pará, Brazil.
[10] Researcher, Instituto Nacional de Pesquisas Espaciais, São José dos Campos, São Paulo, Brazil.

This article aims at to inform on the Brazilian NFI process and to present the main characteristics of the project. The article first describes the institutional framework and then presents the main methodological approaches.

Institutional Framework

The main purpose of the NFI is to generate information on forest resources, both natural and plantations, based on a 5-year measurement cycle, to support the formulation of public policies aiming at forest resources use and conservation. Considering the country extension and the forest resource diversity, a national project requires contributions from different national institutions in an adequate institution framework.

The Brazilian NFI will be coordinated by the Brazilian Forest Service, a recently (2006) created institution that is also responsible for maintaining the Brazilian National Forest Information System. The Brazilian Forest Service will be responsible for providing financial and technical resources to the NFI execution and managing the institutional agreements and technical cooperation with the participating national institutions.

Because the NFI will involve several other institutions in a nationwide survey, there will be technical consultative committees at the national and State levels, aiming at supporting the Brazilian Forest Service for designing guidelines and planning according to regional particularities. Furthermore, the NFI planning process will include an ad hoc committee that may join experts in discussing specific themes such as sampling, biodiversity, and socioeconomics whenever a high level of knowledge support is required to assure the success of the NFI.

Considering the complexity and diversity of the activities and subjects related to the NFI, the project will be based on partnerships with other institutions, which will coordinate specific components of the project to supply the Brazilian Forest Service with the required data to produce the NFI results. Among the partner institutions are the Brazilian Institute of Geography and Statistics, which is responsible for vegetation mapping; the Brazilian National Institute of Spatial Research, which will coordinate satellite image interpretation at the landscape level; and Embrapa Forestry, the forestry branch of the Brazilian Agricultural Research Corporation, which will coordinate the research program to support the NFI. Universities will take part through the quality control program, and private companies or organizations will be involved in field data collection through business contracts.

Methodological Framework

The NFI will be a nationwide and multisource project. It is organized into the following five components:

1. General coordination.
2. Vegetation mapping.
3. Field data collection.
4. Landscape data collection.
5. Associated programs.

General Coordination of the NFI

The general coordination component will be led by the Brazilian Forest Service, acting as the NFI headquarters in Brasilia, the capital of Brazil. The main activities of the component include administrative and technical support, managing the information system, and establishing and refining technical procedures aimed at adopting national standards. In practice, the general coordination component will integrate data gathered from other components, to process them and make available the NFI outcomes for different interest groups. Furthermore, the general coordination component will maintain a permanent strategy of communication, which includes also contact with national and international groups dealing with forest assessment.

Vegetation Mapping

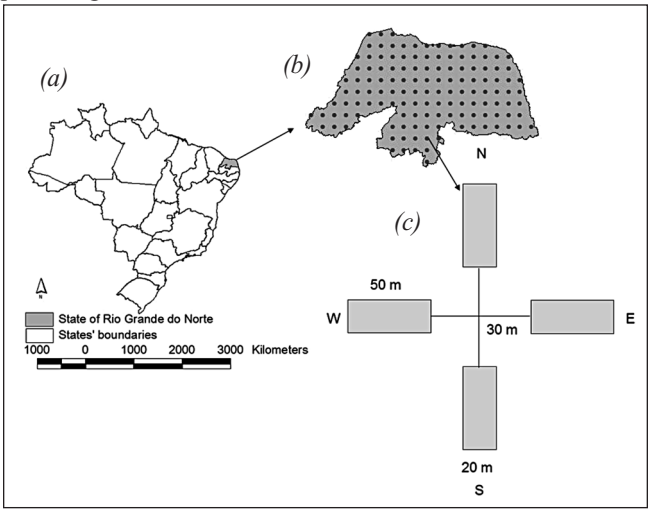
This component encompasses the preparation of topographic maps and the vegetation mapping. It is proposed as a vegetation mapping scheme, to be updated every 5 years, based on topographic maps at a scale of 1:250,000 and China-Brazil Earth Resource Satellite images. The vegetation mapping component will serve to guide the field sample plot selection and support forest area estimates for different poststratification criteria such as biome, vegetation class, State, species group, and so on. A

recent vegetation mapping project of natural vegetation remnants, carried out in each biome and based on Landsat satellite images from 2003, will serve as the basis for the first NFI edition.

Field Data Collection

The sampling design for field data collection will be based on the cluster plot distribution over a systematic grid of 20 x 20 km potential sample points (fig. 1) established from the framework offered by the 1:250,000-scale topographic maps. Plots to be measured in the field will be selected according their status of forest or nonforest and according to the vegetation mapping data. Fixed-area sampling units will be grouped in clusters of four rectangular sample plots, with size and shape defined according to biome characteristics as well as the status of forest plantations. Data collection on sample clusters will report quantitative and qualitative forest attributes or variables by measurement of the classical dendrometric variables, species identification, and recording qualitative variables that are useful for forest ecosystem characterization. Simultaneously to each sample plot measurement, it is proposed that an expedited socioeconomic survey be conducted nearby consisting of two to four interviews aimed at gathering data to describe how local communities view and are using their available forest resources and also informing them about some national programs of forestry incentives (Kleinn *et al.* 2005).

Figure 1.—*Brazil map (a) showing the State of Rio Grande do Norte and details of the sampling design: a 20-x-20-km base grid (b) laid out over the State and (c) the basic cluster sample plot design.*



Data Collection for Landscape Analysis

Using the same sampling framework offered by the 1:250,000-scale topographic maps, an additional sample design will be developed based on a grid of 40 x 40 km sampling points from which data will be collected at the landscape level. Instead of field measurement, this sample unit of 10-x-10-km plots will be based on high-resolution satellite image interpretation. Some of the landscape attributes to be analyzed include forest fragmentation, changes on forest cover and land use, and the condition of permanent protected areas required by law.

Associated Programs

The associated programs component aims at supporting the NFI with improved methods and procedures and producing complementary information that, because of its nature, will not be collected in the sampling frameworks described previously.

The main associated program is the Research and Development Program, which will be of particular importance for the initial edition of the NFI because several demands exist for research and methodological procedures to be incorporated by the different components of the NFI. For example, the NFI field manual will be elaborated and will include content on the state of the art of vegetation measurement protocols in each biome. A second associated program is the Training Program, which will aim at providing human resources able to accomplish the NFI needs with the required standard quality. A third associated program is the Quality Control Program, which will aim at establishing procedures for data quality control and checking a fraction of the measured sample plots. A fourth associated program will be set up by the Brazilian Forest Service to produce annual forest indicators, based on secondary data gathered from different sources. The program is proposed to monitor annually at least three forestry indicators at the national level: (1) area of natural forest under sustainable management, (2) area of plantations, and (3) forest growth and yield data based on permanent sample plot networks already established in every biome (Oliveira *et al.* 2005). These indicators will be recorded annually but will be analyzed for the NFI 5-year measurement cycle. Additional associated programs may be designed according to the needs and priorities identified in the context of NFI purposes.

Future Developments

So far, the new Brazilian NFI has been built through a project preparation phase of which at least two aspects are important to be mentioned. The first important aspect is that the project proposes a methodological approach and an institutional framework for implementing the NFI at the national level. The project preparation phase is an important step to obtain financial resources as well as the required institutional and political support for the project's implementation. The second important aspect is that the project preparation phase has been a participatory process in which the main institutions dealing with natural resources monitoring were involved as well as many experts on forest inventory. This participatory process is important to make the NFI a national project instead of just one more project at the regional level. Given these initial efforts, project implementation is a process yet in development. Some of the next important steps include testing the methodology of field data collection in every biome, which will contribute methodological refinements and more accurate cost estimates; establishing the general coordination component and the NFI information system; establishing the required agreements with the institutions responsible for every component; and initiating the Research and Development Program based on immediate priorities already identified. In the meantime, it will be also important to disseminate the project and its potential value for the society as a means of gaining political support, which is important to guarantee one of the main characteristics for national forest inventories: to be a national permanent program.

Acknowledgments

The authors thank Tasso Azevedo, Tatiana Sá, Peter Holmgren, Mohamed Saket, Alberto Sandoval Uribe, Ron McRoberts, Brad Smith, Liz Mayhew, Michelle Zweede, and Pedro Arraes for their cooperation in this project.

Literature Cited

Brena, D.A. 1995. Inventário florestal nacional: proposta de um sistema para o Brasil. Tese de Doutorado. Curitiba, Paraná, Brazil: Universidade Federal do Paraná. 226 p.

Food and Agriculture Organization of the United Nations (FAO). 2005. Global forest resources assessment 2005: progress towards sustainable forest management. FAO Forestry Paper 147. Rome: Food and Agriculture Organization of the United Nations. 352 p.

Holmgren, P.; Persson, R. 2002. Evolution and prospects of global forest assessments. Unasylva. 210: 3-9.

Kleinn, C.; Ramírez, C.; Holmgren, P.; Valverde, S.L.; Chavez, G. 2005. A national forest resources assessment for Costa Rica based on low intensity sampling. Forest Ecology and Management. 210: 9-23.

Machado, S.A. 1984. Inventário florestal nacional das florestas plantadas nos Estados Paraná e Santa Catarina. Brasília, Distrito Federal, Brazil: Instituto Brasileiro de Desenvolvimento Florestal. 283 p.

Oliveira, Y.M.; Rosot, M.A.D.; Luz, N.B.; Mattos, P.P.; Guimarães, D.P.; Gomide, G.L.A.; Sá, I.B.; Freitas, J.V.; Silva, J.N.M.; Garrastazu, M.C.; Higuchi, N.; Costa, T.C.C. 2005. Sistema nacional de parcelas permanentes: proposta de modelo metodológico. Colombo, Paraná, Brazil: Embrapa Documentos 106. 67 p.

Japanese National Forest Inventory and Its Spatial Extension by Remote Sensing

Yasumasa Hirata[1], Mitsuo Matsumoto[2], and Toshiro Iehara[3]

Abstract.—Japan has two independent forest inventory systems. One forest inventory is required by the forest planning system based on the Forest Law, in which forest registers and forest planning maps are prepared. The other system is a forest resource monitoring survey, in which systematic sampling is done at 4-km grid intervals. Here, we present these national forest inventory systems. High-resolution satellite data were examined to classify forest types with an object-oriented method for spatial extension of monitoring data. These results as well as other information, such as forest registers and forest planning maps, are accumulated in the National Forest Resources Database.

Introduction

Japan consists of four main and many small islands that extend about 3,000 km from southwest to northeast, with mountain chains running throughout. The country spans four climatic zones: subtropical, warm-temperate, cool-temperate, and boreal. Steep mountains, which are mainly covered with forests, approach the coastline and occupy about three-quarters of the land. The diversity of Japanese forests, which range from mangrove to boreal forests, comes from these geographic and climatic characteristics (Ohba 1987, Yamanaka 1990): more than 1,000 species of trees have been recorded (Satake *et al.* 1989).

In Japan, forests cover a total area of about 25 million ha. The figure has remained stable for 40 years, but the ratio of manmade forest to natural forest has varied greatly (MAFF 2003). The increase in demand for timber for reconstruction after World War II caused natural forests to be converted to sugi (*Cryptomeria japonica*) and hinoki cypress (*Chamaecyparis obtusa*) plantations. Natural forests, which had been continuous, became fragmented, particularly from the 1950s to 1970s (Japan FAO Association 1997). At present, 44 percent of forests are manmade and the rest are natural or seminatural. The total stock is 4.04 billion m^3 and the average is 161 m^3/ha, with an annual increment of 110 million m^3/ha (2.7 percent).

Thirty-one percent of manmade forests are national and 69 percent are public or private. The peak of the stand age distribution is between 40 and 50 years. Most individual owners of private forests are small-scale forest holders with less than 5 ha of land. As a result, management practices such as planting, thinning, and harvesting are often done intermittently (Sakai 1999). Although the rotation age in manmade forests was planned to be 40 years, it is tending to become longer. Abandoned areas after final cutting in manmade forests can be found here and there because the price for domestic timber has fallen compared with imported timber.

It is essential for sustainable forest management to monitor forest resources (Shiraishi 1998). We conduct forest inventories by the forest planning system based on the Forest Law in Japan, although a new forest resource monitoring survey was started in 1999 to measure indicators of the Montreal process, which aims to achieve sustainable forest management and estimate carbon stocks (Iehara 1999). This article describes the national forest inventory (NFI) system in Japan and the challenge of increasing the scale of data monitoring using remote-sensing techniques.

[1] Forest Conservation and Management Group Leader, Shikoku Research Center, Forestry and Forest Products Research Institute, 2-915, Asakura-Nishimachi, Kochi 780-8077, Japan. E-mail: hirat09@affrc.go.jp.

[2] Chief of Climate Change Office, Bureau of Climate Change, Forestry and Forest Products Research Institute, 1 Matsunosato, Tsukuba 305-8687, Japan. E-mail: machan@affrc.go.jp.

[3] Chief of Forest Resource Analysis Laboratory, Department of Forest Management, Forestry and Forest Products Research Institute, 1 Matsunosato, Tsukuba 305-8687, Japan. E-mail: iehara@affrc.go.jp.

National Forest Inventory Systems

Japan has two independent forest inventory systems. One forest inventory is required by the forest planning system based on the Forest Law, in which forest registers and forest planning maps are prepared. The other system is a forest resource monitoring survey, in which systematic sampling is done at 4-km grid intervals. Here, we present these NFI systems.

Inventory Under the Forest Planning System

The forest inventory consists of forest registers and forest planning maps. Forest registers are required for all subcompartments of all private and national forests and record information on area, species, diameter at breast height (d.b.h.), volume, and so on. Stand volume is estimated from empirical yield tables by region. The total number of compartments is 370,000 and the total number of subcompartments is 31,000,000. The registers are updated every 5 years. All information of subcompartments is linked with corresponding polygons in forest planning maps, which are prepared on a scale of 1:5,000 for all forest areas. The boundaries of all compartments and subcompartments are delineated on the maps. At present, these geodata are being converted to Geographic Information Service (GIS) data, and more than 70 percent of their boundaries in private forests and all of those in national forests have now been digitized for GIS.

Forest Resource Monitoring Survey

The new Japanese forest resource monitoring inventory is carried out by the Forest Agency of the Ministry of Agriculture, Forestry and Fisheries of Japan and was started in 1999 in both private and national forests to understand the conditions and changes of Japanese forests. The objective of the forest resource monitoring survey is to provide objective data that enables regional forest offices to form regional forest plans for private forests and national forest plans. The survey is also necessary to understand the biodiversity and productivity of forest ecosystems and the carbon cycle. The results are reflected in the national forest plan (Forest Agency 2004).

More than 15,700 permanent plots, which consist of three circular subplots, are set up at 4-km grid intervals systematically (i.e., every 16 km^2). These plots are selected after interpreting aerial photographs to judge whether the plots were forest or nonforest. The inventory staff navigate to a target plot by Global Positioning System (GPS) and identify the center of the plot. A survey team consists of three or more staff members who are specialists in forestry and forest vegetation. In the largest circular subplots, all trees of more than 18 cm d.b.h. are measured, all species of vegetation are recorded, and stumps are checked. In the medium circular subplots, the same items are investigated but the smallest size of trees measured is 5 cm d.b.h. In the smallest subplots, all trees of more than 1 cm d.b.h. are measures, species of vegetation are recorded, stumps are checked, and fallen trees are investigated. The same plots are measured every 5 years; therefore, more than 3,000 permanent plots are measured in a year.

Decision of Survey Points

Points proposed for the forest resource monitoring survey are selected at 4-km grid intervals on the Japan-19 Plane Orthogonal coordinate system. The Forest/nonforest status is interpreted from aerial photographs for all points proposed and other available information is also used. Survey points are marked on the forest planning maps as well as aerial photographs.

Establishment of Permanent Plots

A survey team decides on a survey point by referring to aerial photographs and GPS. A plastic pole is laid on the ground at the survey point. Three circular plots, which have the same center at the survey point, are established. The radii of the small, medium, and large circles are 5.64 m, 11.28 m, and 17.84 m, respectively (fig. 1). Each plot is defined as follows:

- Small circular plot = small circular area.
- Medium circular plot = medium circular area minus small circular area.
- Large circular plot = large circular area minus medium circular area.

Investigation of General Condition of the Plot

Altitude. Altitude is interpreted from the 10-m contours on the map on which survey points are shown.

Aspect of Slope. Slope aspect is classified into N, NE, E, SE, S, SW, W, NW, and flat.

Slope Angle. Slope angle is measured to the nearest degree at the center of the plot using a clinometer and other equipment.

Surface Geology. Surface geology is identified from the surface geology map produced from land cover classification surveys, which are conducted by local governments.

Soil Type. Typical soil type within the plot is decided by visual observation during the survey by referring to existing soil maps.

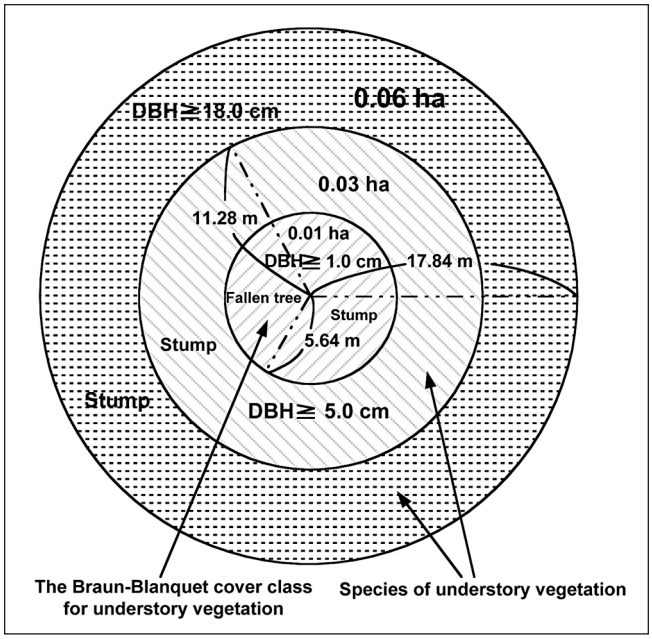
Local Topography. Local topography is selected from flat ridge, lean ridge, convex slope side, concave slope side, equilibrium slope, spur with erosion, spur with deposition, talus cone, alluvial fan, flood flat, alluvial depositional landform, diluvial terrace, upland, and wetland.

Distance From Roadway. Distance from roadway to the plot center is measured to the nearest 100 m in a direct line.

Distance From Village. Distance from the outer edge of a village to the plot center is measured to the nearest 100 m on the map on which survey points are shown.

Soil Erosion Grade. Soil erosion is graded according to the five ranks shown in figure 1.

Figure 1.—*Establishment of permanent plot for forest resource monitoring survey.*



Special Mention for Stand. Damages by disease, insects, animals, and climate are judged and recorded.

Related Data. Land use type is recorded from the forest register and forest planning map. Regulations, laws, stand type, and species on the forest register are recorded.

Stand Structure. The dominant species is judged by visual observation. Stand age is judged from the forest register or factfinding on the spot. Stand type is judged as single-story forest, two-story forest, multistory forest, and unstocked land by visual observation. Category of regeneration is selected from planting, natural seedling, and sprouting by visual observation referring to the forest register.

Operation Records During 5 Years. Operations, if any, such as clear-cutting, regeneration, felling in multistory forest, selective cutting, and thinning, are recorded based on factfinding on the spot.

Complete Tree Tally

The target trees for investigation are those with d.b.h. of 1.0 cm or more in small circular plots, 5.0 cm or more in medium circular plots, and 18.0 cm or more in large circular plots. The following items are checked and measured for all target trees:

Numbering. Trees with d.b.h. of 18.0 cm or more are numbered to identify them in the next investigation.

Species. The species of each tree is identified and recorded.

d.b.h. d.b.h. is measured to the nearest 0.1 cm with a tape measure for diameter.

Tree Height. Around 20 standard trees are selected and measured to the nearest 0.1 m.

Dead Tree. Dead trees are recorded if observed.

Animal Noise. Animal noises are recorded if heard.

Cavity in Stem. Cavity in the stem is recorded if observed.

Investigation of Stumps

Stumps of 5.0 cm in diameter at 20 cm above the ground are measured within small and medium circular subplots and stumps of 18.0 cm in diameter at 20 cm above the ground are measured within large circular subplots. All stumps are measured and recorded in the specific plots and stumps resulting from cutting with the last 5 years are measured and recorded in other general plots. The diameters of five standing trees at 20 cm above the ground are measured when stumps are discovered in the plot.

Investigation of Fallen Trees

Fallen trees whose bases are in small circular plots with a diameter of 5 cm or more at the center of the stem are investigated. Decay grade is judged as shown in table 1. All fallen trees are numbered to identify them in the next investigation.

Table 1.—*Decay class in investigation of fallen trees.*

Class	Status
0	The tree has just fallen; leaves remain on branches.
1	Only cambium is decaying; large branches remain.
2	Alburnum is decaying; large branches remain.
3	Heartwood is also decaying; only trunk remains.
4	Cambium has disappeared; only heartwood remains.
5	Original form of wood no longer exists.

Investigation of Understory Vegetation

Cover rates of tree layer, lower-tree layer, shrub layer, herb layer, and bare land within the small circular plots are estimated by visual observation. The Braun-Blanquet cover class (Braun-Blanquet 1964) is estimated for all species of shrub layer and herb layer within the small circular plots. Species that appear within the medium and large plots are recorded.

Scaling Up With Remote Sensing

Although statistics on forest conditions can be obtained by analyzing the inventory data, the extent and spatial change of forest cover cannot be grasped. Therefore, the Forest Agency started a project in 2002 to develop methodologies to apply high-resolution, remotely sensed data to spatially extend the results of forest resource monitoring surveys. Data concerning forest resource monitoring surveys with remotely sensed data are required for classifying the forest type of natural and man-made forests for forest management as well as for ecological reasons. High-resolution satellite data is expected to be useful for estimating stand factors such as species, height, and density.

The project aims to integrate data from the forest resource monitoring surveys and other information with high-resolution satellite data. The satellite data must have sufficient resolution to understand forest resources, be a sustainable source, and cover all forest areas in Japan. IKONOS and SPOT-5 satellite data have been used as high-resolution data for the project. An object-oriented classification method was used to segment forest patches in terms of forest management and ecological knowledge using eCognition software. Suitable parameters have been investigated in some regions of Japan where vegetation components were different.

The results will be used to evaluate indices of the Montreal process as well as for national and regional forest planning. The methods for evaluating forest functions and for monitoring forest changes from the results should be considered. As a result, the accuracy of forest resource information in forest registers is expected to be increased.

Conclusions

Forest inventories are essential for sustainable forest management at regional and national levels, although some difficulties exist in continuing the inventories both financially and technically. From fiscal year 2004, the national forest resources monitoring survey has entered its second cycle. It is necessary to examine quality assurance and quality control of monitoring data. The results of monitoring surveys as well as other information, such as forest registers, forest planning maps, etc., are accumulated in the National Forest Resources Database, which is also used for reporting under the Kyoto Protocol (Matsumoto 2005).

Literature Cited

Braun-Branquet, J. 1964. Pflanzensoziologie. Berlin, Germany: Springer-Verlag. 865 p. In German.

Forest Agency. 2004. A handbook for forest planning duties. Tokyo: Japan Foundation Information Center. 1,400 p. In Japanese.

Iehara, T. 1999. New Japanese forest resource monitoring survey. Sanrin. 1384: 54-61. In Japanese.

Japan FAO Association. 1997. Forest and forestry in Japan. Tokyo: Japan FAO Association. 65 p.

Matsumoto, M. 2005. Efforts at developing a method to account for greenhouse gas absorbed by forest for the Kyoto Protocol reporting. Shinrin Gijutsu. 754: 3-9. In Japanese.

Ministry of Agriculture, Forestry and Fisheries (MAFF). 2003. Annual statistics book on forestry census. Tokyo: Association of Agriculture and Forestry Statistics. 321 p. In Japanese.

Ohba, H. 1987. Forest of Japan. In: Saito, M., ed. Forest and culture. Tokyo: Tokyo University Press: 29-56. In Japanese.

Sakai, M. 1999. Minds of forest owners for management and forest resource policy. Journal of Forest Economics. 45(1): 3-8. In Japanese.

Satake, Y.; Hara, H.; Watari, S.; Tominari, T., eds. 1989. Wild flowers of Japan—woody plants. Tokyo: Heibonsha Ltd. 626 p. In Japanese.

Shiraishi, N. 1998. How should forest inventory be planned towards 21 century? Sanrin. 1373: 2-10. In Japanese.

Yamanaka, T. 1990. Forest vegetation of Japan. Tokyo: Tsukijishobo Ltd. 223 p. In Japanese.

The Design of the Second German National Forest Inventory

Gerald Kändler[1]

Abstract.—In Germany, a sample-based national forest inventory (NFI) took place for the first time from 1986 to 1990 (in West Germany only); the second one took place from 2001 to 2002. The inventory design is based on a systematic distribution of tracts on regular grids of regionally differing width. The primary sampling unit is a quadrangular tract with sides of 150 m. The tract corners located in forests are the centers of permanently marked subplots in which different sampling procedures are used for the selection of sample trees and the survey of other characteristics. The core sampling technique is horizontal point sampling by means of the angle-count method with a basal-area factor of 4 m^2ha^{-1}. The German NFI is a periodic survey conducted in the whole territory at time intervals that are not predefined but have to be determined anew for each repetition. The results obtained by the second survey document an increase of the growing stock and a high level of periodic annual volume increment in West Germany. Generally, the information provided by the NFI is of great importance for forest policy, especially in connection with international commitments to report on forest resources; as the information is also important for the wood processing industry. Due to the societal, ecological, and economic significance of forests and forestry, the third NFI has now been scheduled to take place from 2011 to 2012.

Introduction

In Germany, the concept of a statistically designed national forest inventory (NFI) emerged in the 1970s. It took some time to convince the upper management of the German Forest Service that a sample-based survey is an efficient tool to collect representative and valid data on forest condition. Finally, the NFI, referred to as the Federal Forest Inventory (Bundeswaldinventur), was included into the Federal Forest Act in 1984. The main objective is to provide an overview on large-scale forest condition and forest productivity by using consistent procedures in a permanent design that will enable remeasurement of the same plots to obtain data on increment and drain. According the Federal Forest Act, the NFI is not a mandatory survey to be repeated at predefined time intervals; instead, its implementation requires an executive order law that has to be decided upon anew each time by the Federal Government and the Federal States (the Länder). This procedure is due to the distribution of competencies concerning forestry between the Federal Government and the Federal States. The Federal Government has only limited executive authority in forestry affairs because the Forest Service is in the sphere of competence of the Federal States. Thus, compared with other countries, (e.g., Finland, Norway, Sweden, or the United States), this monitoring system was introduced quite late in West Germany, with the first Federal Forest Inventory carried out from 1986 to 1990. After the reunification in 1990, the second inventory was conducted from 2001 to 2002. Today, the usefulness of a sample-based inventory is generally accepted and the results obtained are widely demanded.

Design

The German NFI is based on a systematic rectangular grid with clusters (tracts) as primary sampling units (fig. 1). The General Administrative Regulation prescribes a grid width of 4 x 4 km as a so-called basic grid covering the complete surface of Germany with a defined starting point. The sample grid is intensified in some Federal States or parts of them to a 2.83-x-2.83-km or 2-x-2-km quadrangular grid. With the second NFI, 45,098 tracts cover forest and nonforest land throughout

[1] Head of Department, Forest Research Institute of Baden-Württemberg, Department of Biometrics, Wonnhaldestr. 4, 79100 Freiburg i. Br., Germany. E-mail: gerald.kaendler@forst.bwl.de.

Germany. Of these tracts, 18,822 are located with at least one corner in a forest (table 1). The total number of subplots (tract corners) on forest land is 54,026, of which 51,768 are on accessible wooded ground, including openings (wooded ground temporarily without a forest cover).

The tract is a quadrangle with sides of 150 m. The sides of the tract are oriented north-south and east-west, respectively. Samples are taken on the tract lines as line-intersect samples for the forest road inventory (only in the new States in East Germany) and on the tract corners if they hit a forest. Tracts with at least one corner in a forest are forest tracts and are to be surveyed. The sample selection on a tract corner is made according to different methods (fig. 2): trees with a minimum diameter at breast height (d.b.h.) (diameter at 1.3 m above ground) of 7 cm over bark are selected by the angle-count method (horizontal point sampling) with a basal-area factor of 4 [m^2ha^{-1}] if they are alive or recently dead (fine branchwood maintained in full) and if they belong to the same stand type as that in which the subplot center is located. These trees are subject to various measurements and assessments. According to the permanent design, their locations are recorded by polar coordinates. The attributes recorded are sample tree code (e.g., new, remeasured, removed), species, azimuth, horizontal distance from center, canopy class, d.b.h., tree class (according to Kraft), tree age, tree height (only subsample), upper diameter (only subsample), height code, trunk code, trunk damage, and lopping. An additional angle count with a basal-area factor of 1 or 2 [m^2ha^{-1}] is carried out as a basis for describing the forest structure by tree species and layer. Smaller trees are sampled in circular plots: trees higher than 0.5 m and d.b.h. of less than 7 cm over bark are surveyed in a plot with a radius of 1.75 m centered on the tract corner; trees 20 to 50 cm in height are recorded in a circular plot with a radius of 1 m located 5 m away from the tract corner, generally to the north. Attributes of the small trees surveyed in the circular plots are species, tree size class, damage by game or other animals, individual protection, fencing protection, and canopy class. Deadwood (lying or standing—either whole or broken, stumps, or leftovers from hauling) is surveyed in a plot with a radius of 5 m (centered in the corner). In a circle with a 10-m radius, trees up to 4 m in height, shrub layers, and ground vegetation are surveyed. In a circle of 25 m around plot corners, site characteristics and forest edges are

Figure 1.—*Regular sample-grid of the German national forest inventory with tracts. This example is from the State of Baden-Württemberg with a 2-x-2-km grid.*

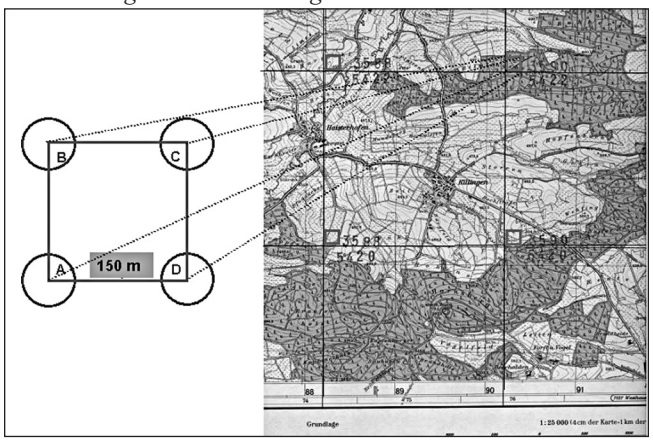


Figure 2.—*Structure of a tract and a subplot on a tract corner.*

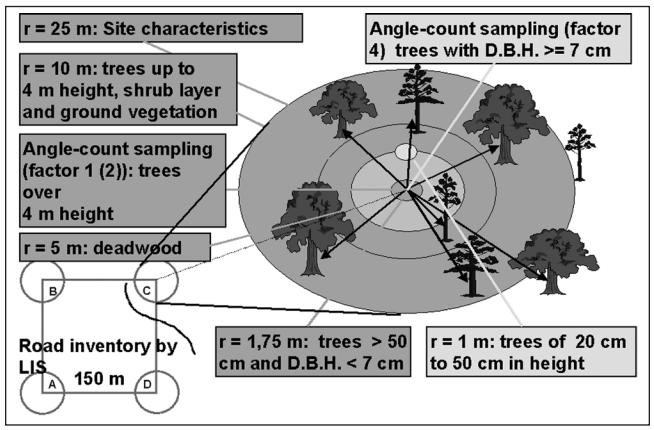


Table 1.—*Number of tracts in regions with different grid widths.*

Grid width	Tracts outside forests	Tracts in forests	Total number of tracts
2 x 2 km	13.737	10.055	23.792
2.83 x 2.83 km	6.174	3.350	9.524
4 x 4 km	6.365	5.417	11.782
Total	26.276	18.822	45.090

recorded. Forest edges are relevant for two reasons: (1) as a characteristic with regard to forest fragmentation and (2) for theoretical statistical reasons. Due to the fact that sample plots are located only in forests, the selection probability of sample trees near a forest border is biased (edge effect bias). Based on the geometry of forest edges captured by polar coordinates related to the plot center, the individual selection probability of sample trees whose limit circle overlaps the forest border can be corrected. To this end, the area of the part of the limit circle lying in the forest is calculated. The ratio of the reduced area to the full limit circle corresponds to the reduction of selection probability.

Altogether, about 150 characteristics for each inventory tract are recorded. As opposed to Germany's first NFI (which occurred from 1986 to 1990 in West Germany), the spectrum of data collected in the second NFI has been expanded to include ecological parameters such as deadwood, closeness to nature, ground vegetation, and forest borders.

The German NFI is conducted as a periodic inventory at time intervals that are not predefined; instead, the repetition of the survey has to be determined anew, if required.

Methodological Note Concerning Growth Estimation

Using horizontal point sampling on permanent plots has the effect that sample composition changes between consecutive surveys due to altered selection probabilities of the sample trees representing different components of forest volume growth. Hence, different estimators of volume growth were developed. A set of literature exists that deals with this issue (for example, Roesch *et al.* 1989, Van Deusen *et al.* 1986). The estimator of volume growth used for the analysis of the German NFI is analogous to one estimator described by Roesch *et al.* (1989) and requires the assessment of the initial volume of the so-called nongrowth trees that were not measured during the preceding inventory because those trees had not yet qualified for the sample. The initial (time 1) volume is assessed by means of growth functions for diameter and height with time 2 tree age and diameter and height, respectively, as predictors.

Organization, Competencies, and Tasks

The German NFI is conducted as a joint mission of the Federal Government and the Federal States that requires close collaboration during preparation and implementation. The Federal Ministry of Agriculture is responsible for the central coordination, scheduling, data management, data processing, and reporting; the States are responsible for the data collection and have to provide the required means (field crews, other personnel, instruments, data logistics, etc.). In each State, a temporary staff unit is in charge of the implementation of the survey. One important task of the State inventory unit is quality assurance of the collected data. Quality assurance is done by a so-called inventory inspection that has to cover a minimum of 5 percent of the plots.

In Germany, only one Federal institution is permanently in charge of inventory and monitoring tasks: the Institute for Forest Ecology and Forest Assessment of the Federal Research Centre for Forestry and Forest Products. At the Federal State level, the Baden-Württemberg Forest Research Institute is the only institution with a research unit permanently dealing with forest inventory issues and tasks (the author's department). In connection with the German NFI, this unit has devised basic methods such as volume functions and taper models (Kublin 2003) as well as the forest development and timber supply model applied for timber supply forecasts based on the NFI data.

Results of the Second National Forest Inventory

Some basic figures on Germany's forests in comparison with the most populous U.S. State, California, and the United States are compiled in table 2. Forest percentage is about the same in the three areas, but growing stock per area is much higher in Germany.

Comparison With Other European Countries

The comparison of forest characteristics between different countries in Europe (with the exemption of Russia) shows that forests are an important natural factor and the basis of multipurpose forestry in a heavily settled country like Germany. Figure 3 displays the total growing stock (volume over bark) of larger

Table 2.—*General information on Germany's forests in comparison with those of the State of California and the United States.*

		Germany	California	United States
Total land area	*1,000 ha*	34,895	41,000	915,896
Forest area	*1,000 ha*	11,076	13,451	303,089
Percent of total land area		*32*	33	33
Total growing stock	*1,000.000 m^3*	3,381	1,907	35,118
Growing stock per hectare	*m^3 ha^{-1}*	315	142	116

Sources: FAO 2005, USDA Forest Service 2006

Figure 3.—*Total growing stock in selected European countries (from FAO 2005).*

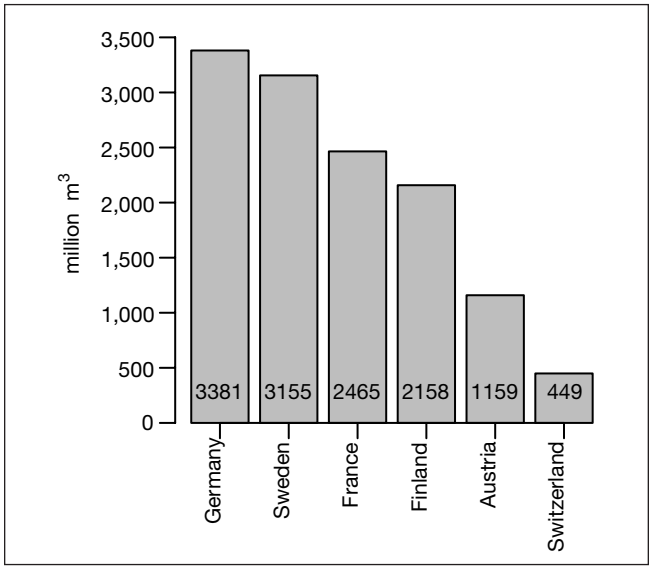
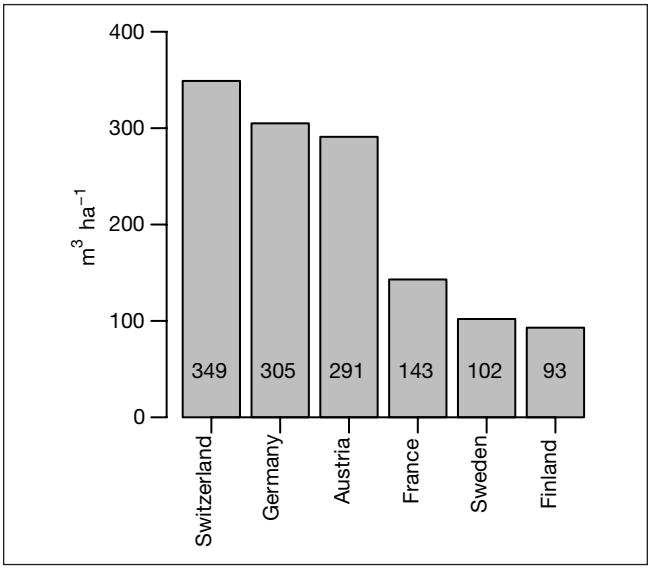


Figure 4.—*Growing stock per hectare in selected European countries (from FAO 2005).*



European countries as well as smaller ones in Central Europe. Although forest area in Germany is not as large as it is in other countries of Northern Europe or in France, Germany's forests are those with the highest growing stock in Europe (excluding Russia). This observation becomes evident when examining the average growing stock per hectare (fig. 4). The highest volume per hectare is attained in Central Europe (Germany, Switzerland, and Austria).

Growth and Drain

With the second NFI, for the first time, growth and drain were assessed on a large, regional scale, but only in the West German Federal States because remeasurement has been done in this area only. The overall number of revisited permanent subplots is 30,538, with a mean period length of 14 (vegetation) years. Growth is expressed as periodic annual increment (PAI) and drain is periodic annual cut and mortality. The average PAI in West Germany is 12.7 $m^3ha^{-1}year^{-1}$ (corresponding to 181.5 ft^3 $ac^{-1}year^{-1}$) The figures obtained substantiate the observation that currently periodic annual growth in West German forests is at a high level, which is in agreement with observations in other European regions (Spiecker *et al.* 1996). Figure 5 shows the comparison of growth and drain from the forests of the major Federal States of Baden-Württemberg, Bavaria, and Lower Saxony for the period from 1987 to 2002 (see also table 3). Generally, drain is below growth expressed as PAI, which leads to an increase of the growing stock in this period, but differences between the regions are evident. In the State of Baden-Württemberg, drain reaches the highest level, almost exhausting increment, whereas in Bavaria and Lower Saxony, only about 67 and 53 percent, respectively, of the increment has been removed.

Assessment of Future Timber Supply

The NFI data is an important base for the assessment of future timber supply in the following 4 decades (from 2003 to 2042). The forecasts of potential timber yield for the period 2003 to 2017 support that an increase of wood harvest is possible. The mean annual cut in West Germany attained about 6.7 $m^3ha^{-1}year^{-1}$ (merchantable volume under bark) in the period 1987 to 2002; in the following period (2003 to 2017), this amount could be raised up to 8.3 $m^3ha^{-1}year^{-1}$ without violating sustainability. The wood processing industry has already responded to these perspectives and has reinvested in additional sawmill capacities.

Figure 5.—*Periodic annual growth and drain in West German Federal States (in cubic meters of volume over bark per hectare and year).*

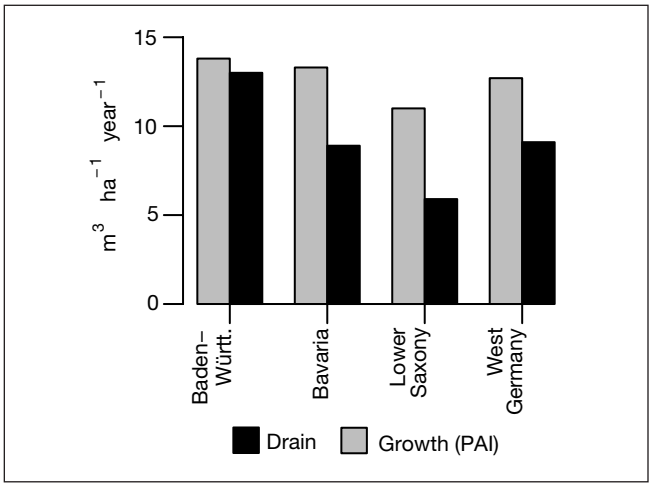


Table 3.—*Growth (periodic annual volume increment) and drain in selected West German Federal States.*

		Baden-Württemberg	Bavaria	Lower Saxony	West Germany
Growth	$m^3\ ha^{-1}\ year^{-1}$	13.8	13.3	11.0	12.7
Drain	$m^3\ ha^{-1}\ year^{-1}$	13.0	8.9	5.9	9.1

Conclusion

In Germany, the NFI is not a permanent institution, unlike, for example, the Forest Inventory and Analysis program in the United States or the NFI agencies in other countries. Instead, the NFI is a temporary mission implemented in close collaboration with Federal and State authorities. Due to this fact, the NFI will remain a periodic survey with a longer time interval. Today, politicians, the Forest Service, the timber industry, and other stakeholders have recognized the usefulness of large-scale, sample-based inventories. The appreciation of the information and data made available by the NFI has increased a great deal. The need for objective data on forest condition and its development is generally accepted. Forest policy has a great demand for objective figures on forests, especially in the scope of international conventions and resulting commitments to report. For example, for the Kyoto Protocol to the United Nations Framework Convention on Climate Change (Kyoto Protocol) signatory States, the part of the national inventory report concerning the carbon budget of forests can be improved significantly if current data from NFIs are available. Another important impact of the second NFI in Germany triggered by the results of the timber supply assessment for the next decades is an increased demand for timber as well as for energy wood. In addition, the timber industry demands reliable data on wood supply for decisions on investments. Therefore, the third NFI is already under way and is scheduled to take place from 2011 to 2012. One reason for the timing is the fact that 2012 is the final year of the first Kyoto Protocol commitment period with the necessity for the signatory States to report on carbon-stock changes in the forests based on reliable data.

Literature Cited

Food and Agriculture Organization of the United Nations (FAO). 2005. Global forest resources assessment (FRA) 2005 country tables. http://www.fao.org/forestry/site/32179/en/. (16 January 2007).

Kublin, E. 2003. Einheitliche Beschreibung der Schaftform—methoden und programme—BDATpro. Forstwissenschaftliches Centralblatt. 122: 183-200.

Roesch, F.A.; Green, E.J.; Scott, C.T. 1989. New compatible estimators for survivor growth and ingrowth from remeasured horizontal point samples. Forest Science. 35(2): 281-293.

Spiecker, H.; Mielikäinen, K.; Köhl, M.; Skovsgaard, J.P. 1996. Growth trends in European forests. Heidelberg, Germany: Springer Germany. 372 p.

U.S. Department of Agriculture (USDA) Forest Service. 2006. FIA database. California: 2001–2005 (interim) tables. http://www.fs.fed.us/pnw/fia/local-resources/pdf/ca_07262006_tables.pdf. (16 January 2007).

Van Deusen, P.C.; Dell, T.R.; Thomas, C.E. 1986. Volume growth estimation from permanent horizontal points. Forest Science. 32(2): 415-422.

Additional Reading

Bundesministerium für Verbraucherschutz, Ernährung und Landwirtschaft. 2001. Survey instructions for Federal Forest Inventory II (2001–2002). English version. http://www.bundeswaldinventur.de/. (22 October 2005).

The Brazilian National System of Forest Permanent Plots

Yeda Maria Malheiros de Oliveira[1], Maria Augusta Doetzer Rosot[2], Patrícia Povoa de Mattos[3], Joberto Veloso de Freitas[4], Guilherme Luis Augusto Gomide[5], Marta de Fátima Vencato[6], and Marilice Cordeiro Garrastazú[7]

Abstract.—The Brazilian National System of Forest Permanent Plots (SisPP) is a governmental initiative designed and being implemented in partnership by the Ministry of Environment (MMA), represented by the National Forest Programme (PNF) and the Brazilian Forest Service (SFB) and the Embrapa Forestry (a research center of the Brazilian Agricultural Research Corporation - Embrapa). The proposal was presented, discussed and approved by the Brazilian forest inventory community in a workshop carried out in September 2004. The methodological model converged to a national network organized by biomes, interconnecting existent initiatives (in the biomes Amazonia and Caatinga) and promoting new networks (in the biomes Cerrado, Pantanal and Atlantic Rain Forest). The purpose of SisPP is to congregate the regional networks as its branches (branch, in this case, means the regional and stratified by biomes permanent plots networks and the community of researchers and other individuals who share the same interests and are involved in the subject). This article intends to describe the main objectives and elements of SisPP as well as the strategies that are being used for its implementation. Another important aspect of SisPP is its integration to the new Brazilian national forest inventory, which would use data from the permanent plots as an input to estimate growth and to analyze forest dynamics in the different Brazilian biomes, thus complementing the stock information obtained from temporary sample plots.

Introduction

In forest science and forest research, some distinct existing characteristics require the establishment of specific studies in large areas, through long-term periods, in complex natural and socioeconomic systems, and in systems where sustainable management is fundamental also for the preservation of the systems' functions for future generations. The combination of these aspects turns forest management into a challenge and defines, at the same time, an important role for scientific research (Kleinn and Köhl 1999). Additionally, sustainable forest management requires information on forest production and patterns of development in the present and future and also under different management regimes. That information should be obtained from silvicultural trials and by the observation of forest growth along the time. Estimates on current stocks of volume and on average growth taxes come from forest inventory data (temporary plots, for example); however, present average growth rates do not provide with the necessary accuracy, the tendencies of future behavior of the forest because they are originated from only one measurement in different stands, which tends to represent several development phases (Curtis and Marshall 2005).

[1] Researcher and Project Leader, Embrapa Florestas, P.O. Box 319, CEP 83411-000, Colombo, Paraná, Brazil. E-mail: yeda@cnpf.embrapa.br.

[2] Researcher, Embrapa Florestas, P.O. Box 319, CEP 83411-000, Colombo, Paraná, Brazil. E-mail: augusta@cnpf.embrapa.br.

[3] Researcher, Embrapa Florestas, P.O. Box 319, CEP 83411-000, Colombo, Paraná, Brazil. E-mail: povoa@cnpf.embrapa.br.

[4] Project Manager, Brazilian Forest Service, Ministry of the Environment, SCEN, Trecho 2, Bl. H 70068-900, Brasília DF, Brazil. E-mail: joberto.veloso@florestal.gov.br.

[5] Technician, Brazilian Forest Service, Ministry of the Environment, SCEN, Trecho 2, Bl. H 70068-900, Brasília, DF, Brazil. E-mail: guilherme.gomide@.florestal.gov.br.

[6] Assistant, Embrapa Florestas, P.O. Box 319, CEP 83411-000, Colombo, Paraná, Brazil. E-mail: marta@cnpf.embrapa.br.

[7] Researcher, Embrapa Florestas, P.O. Box 319, CEP 83411-000, Colombo, Paraná, Brazil. E-mail: marilice@cnpf.embrapa.br.

In the extent of forest management, permanent plots are permanently demarcated areas of forest that are periodically remeasured (Alder and Synnott 1992) with the objective of obtaining information on the growth and the dynamics of the forest. In other words, the use of permanent plots aims to obtain information about the variations related to the number of a certain species of tree, the composition of the forest, and the trees' individual dimensions in a certain period of time. Such plots usually present relatively high costs for their implantation and maintenance and require a commitment of resources for a long period, which many corporations' managers consider a disadvantage. However, frequently, permanent plots also can be also used for analyses and purposes other than the study for which they were installed.

In 2004, the Brazilian Ministry of Environment (MMA), in partnership with the Food and Agriculture Organization (FAO) of the United Nations, launched an announcement for the elaboration of a proposal for a methodological model of the Brazilian National System of Forest Permanent Plots (SisPP). Many reasons supported that Federal Government decision, including the following:

- Brazil is the world's largest producer and consumer of forest products. Strategic sectors of the Brazilian economy, such as metallurgy, paper and cardboard manufacturing, and building construction, are highly dependent on the forest sector.
- Since the end of the 1980s, despite the issue's economical and social importance, Brazil has experienced deficiencies regarding the availability of official information of the forest sector.
- There is a lack of systematic information to subsidize decisionmaking regarding public policies and projects from the civil society and private-sector companies. At the same time, in the sense of filling out such a gap, several actions by different institutions were adopted. Until now, however, no such an integrated system existed that could supply periodic information related to forest administration.
- Cooperative efforts among research institutions, private-sector companies, nongovernmental organizations, and universities are vital for establishing a consensual process regarding data collection and data availability to the different participants to compose a national database.
- It is necessary, also, to ensure the acquisition of new data to complement already existing statistics and to extend the monitoring process to species and biomes that have been seldom studied until now. Also, in this case, the magnitude and costs of the establishment of field plots, their measurement, data processing, and data analyses lead to the need for cooperative efforts involving different organizations.

Thus, since 2000, when the National Forest Programme (PNF), was created, and after that, in 2006, when the Brazilian Forest Service (SFB) was established, the formulation of a long-term forest policy has been desired in the country.

In this scenario, the methodological model for SisPP was developed by researchers of Embrapa Forestry (a research center linked to the Brazilian Agricultural Research Corporation - Embrapa) in partnership with other Embrapa research centers and other institutions (Oliveira *et al.* 2005). The model was presented, discussed, and approved by MMA and invited institutions in a workshop held in Curitiba in the State of Paraná, in September 2004.

The idea of "networks" is a trend in the contemporary administration. In this context, the change of paradigm involves cooperative effort instead of isolated actions. The networks should count with advanced communication means to promote the participants' interaction with supplemental qualifications. In this way, the methodological model for SisPP, developed by Embrapa Forestry, converged to the network scheme and is stratified by the six Brazilian biomes: (1) Amazon Rain Forest, (2) Caatinga (scrubland or steppe ecosystem), (2) Cerrado (Brazilian tropical savannah), (4) Pantanal (wetlands of the central west part of the country), (5) Atlantic Rain Forest (coastal rain forest and other vegetation associations), and (6) Pampa (formerly covered by grasslands and some forests). The model was designed for the interconnection among existing initiatives, such as the Dynamic Monitoring of the Brazilian Amazon Forest Network, and the Caatinga Forest Management Network. In a second phase, which was financially supported by MMA, the Cerrado and Pantanal Permanent Plots Network (which

includes two biomes in only one network) and the Atlantic Forest and Pampa Permanent Plots Network (which also includes two biomes in one network) were created. The Planted Forests Permanent Plots Network is still being organized. SisPP will congregate the regional networks as subsystems or network branches and will serve as the database feeder of SisPP. The system's functions will include compilation, systematization, compatibility, data processing, and results publication of each regional network and of the joint activities.

Objectives

This article intends to describe the main objectives and elements of SisPP as well as the strategies that are being used for its implementation. It also presents some guidelines related to the structure, implementation, and coordination of SisPP, whose main purpose is the permanent monitoring of the natural and manmade forests located in the six Brazilian biomes. SisPP will embrace information about forest growth and yield as well as about the forest reaction to direct or indirect disturbances, including the effects of management regimes, silvicultural responses, and climatic changes.

SisPP Design

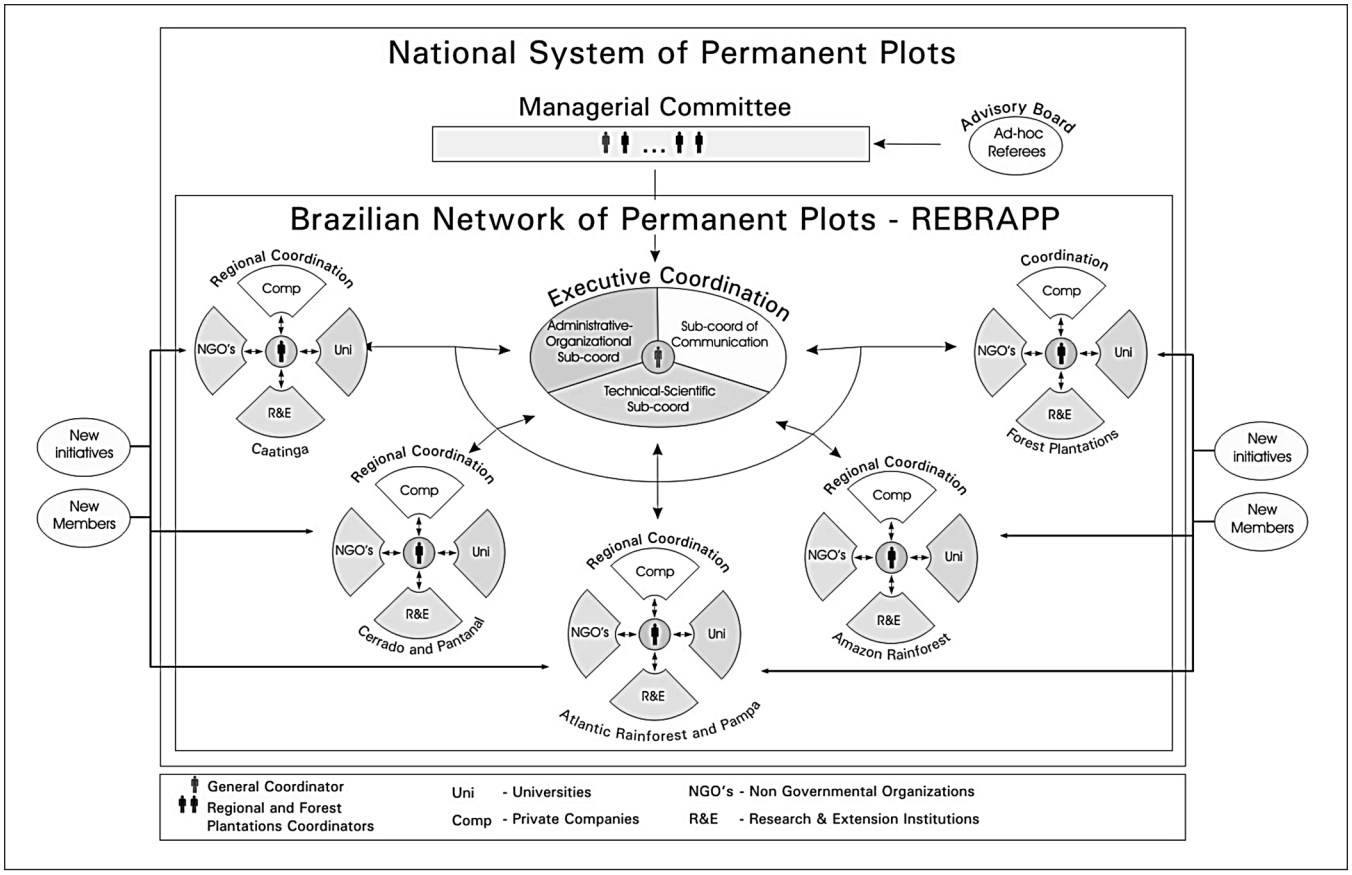
Several network functioning models were studied and none of them fulfilled entirely the basic idea conceived for SisPP; thus, a hybrid model was adopted. The hybrid model includes the following objectives: the independence of the participant institutions, sharing the information generated in each network branch, the definition of mechanisms for information availability, the establishment of a repository of data, and the creation of interconnected sub networks (network branches) agglutinated around a regional coordination. The basic elements idealized for SisPP include the following: system engineering, institutional arrangement, and economical-financial arrangement. The system engineering consists basically of the following initiatives: the establishment of relationships among the network branches; the hierarchization among and inside the Managerial Committee, Advisory Board, and coordination sets (general, regional, and executive); and the system formalization, its organization, and its relationship with the other elements of the system. The institutional arrangement is defined by the integration of public and private institutions potentially involved in the system and the establishment of an organizational chart involving responsibilities that take into account the institutional differences. The analysis of the needs of resources to quantify and qualify equipments, software, and personnel involved in the system results in the economical-financial arrangement.

SisPP has different and interdependent components targeted to enhancing the effectiveness of the system as a whole. The first hierarchical level is the Managerial Committee (figure 1). The committee is composed by a General Coordinator, five Regional Coordinators (one for each biome or biome groups), and an additional coordinator who will represent the planted forests. The General Coordinator manages the network articulation, what means maintaining a continuous data and information flow in the system. The Managerial Committee decides on general matters of SisPP and contributes to constant evaluation and improvement of the system. An *ad hoc* Advisory Board is linked to the Managerial Committee. The board is formed by external elements to SisPP and composed of five researchers of "notorious knowledge" in the subject. The board acts as a referee chamber of the activities developed by the system in meetings to be held at least once a year. The second hierarchical level is the Executive Coordination, which is responsible for the administration of three sub-coordinate sections: (1) Administrative-Organizational, (2) Technical-Scientific, and (3) Communication.

The Executive Coordination will be responsible for the SisPP office, which will also help administer the three sub-coordinate sections. Required human resources include the following: a General Coordinator, a secretary and an administrator linked to the Administrative-Organizational sub-coordinate section, a journalist and a Web designer linked to the Communication sub-coordinate section, and two foresters and a geoprocessing technician linked to the Technical-Scientific sub-coordinate section.

The database of permanent plots is being structured according to the existing protocols of establishment and measurement of permanent plots. All regional networks published specific

Figure 1.—*The Brazilian National Forest Permanent Plots Network organizational chart.*



rules regarding the monitoring process of their permanent plots. SisPP will constitute the forum for technical discussions, aiming for the improvement of those guidelines. For permanent plots already being monitored outside the system, adaptation processes and data conversion formats are being considered. The individuals and participant institutions will have their copyrights and intellectual property guaranteed.

Implementation Strategy

Embrapa Forestry is a research institution that is the forest branch of the Brazilian Agricultural Research Corporation - Embrapa. It is located in Colombo (a metropolitan area of Curitiba) in the State of Paraná, and was designated by MMA to be the institution responsible for SisPP implementation.

The SisPP Executive Coordination is located in a facility (SisPP general building) that permits the collection, storage, manipulation, and updating of the information regarding the existing permanent plots and the creation of a structure for data analyses.

The Executive Coordination is responsible for the following activities:

- Organizing meetings with the permanent plots regional networks.
- Writing and publishing the SisPP statutory agreement.
- Collaborating on the text and publication of the statutory agreement of the regional networks.
- Encouraging the creation of the Planted Forests Permanent Plots Network.

- Building and periodically updating a communication space in the Internet for collaborators and coordinators of the regional networks.
- Publishing a monthly digital informative (newsletter) for the divulgation of the SisPP themes and the regional networks.
- Collaborating on the text and publication of the measurement and maintenance guidelines of the regional networks of permanent plots.

The activities of SisPP also include the creation and updating of a homepage to publish information regarding the concept of the system; its importance, objective, and main expected results in the short, medium and long run; the main team members and their addresses; and recent results. In technical terms, the Executive Coordination is expected to coordinate the trainees' contributions; collaborate on the writing of technical-scientific papers; and organize, coordinate, and promote the First National Congress of Forest Monitoring using Permanent Plots.

Additional activities include: reception, organization, and manipulation of data originated from the regional networks of permanent plots and the documentation and updating of metadata. In addition, they will manage spatialized data using GIS tools.

Results

SisPP has the purpose of providing data for research involving forest growth and yield, surveys and forest inventories, and the monitoring of forest planning. Therefore, is appropriate to define standardized procedures that should be followed as a routine in all permanent plots still to be installed. These procedures should consider the differences in the objectives of the plots that compose SisPP. At the same time, they should represent a compatible group of measurements to be taken on certain basic variables. It is expected that, with the establishment of SisPP, it will be possible to watch aspects regarding such group of variables as well as aspects relative to the plots' size and shape, their physical delimitation in the field, the registration of the initial conditions, the numbering of the trees, the determination of the diameter measurement points and determination of the stands' age, and the mapping of the trees and remeasurement layout.

The computation design to be developed should enable several users to access the system at the same time, control and monitor the updating of data, and permit the obtention of registered information in real time. The information access will take place through the Internet, allowing the registration of any user from any part of the world. The stored data in the centralized base should enable registered queries on the database and queries using raw, processed, or historical data and data crossing from different sites and plots.

SisPP and the Brazilian National Forest Inventory

SisPP has a strong interface with the new Brazilian NFI. The permanent plots installed by the regional networks can contribute with detailed information on forests dynamics and growth as well as give information regarding silvicultural treatments and technical alternatives for multiple use management. On the other hand, it is expected that the Brazilian NFI works in partnership with the regional networks, articulating and sponsoring the installation of some plots in association. The strategy involves the Brazilian NFI protocol that is being followed by the regional networks and also involves the plots' remeasurement and maintenance as part of the structure of the network. This kind of arrangement would be satisfactory for both systems, considering that the network coordinators have been participating actively in the Brazilian NFI project meetings as consultants in the methodology definition and specific measurement protocols for each biome.

Conclusion

The implementation of SisPP is a directive of the Brazilian government and private-sector corporations linked to the forest sector in Brazil. This project is important, especially for those institutions responsible for the elaboration of public policies, forest research, decisionmaking at the strategic and managerial levels, and the medium- and long-run environmental planning. The political circumstances are is quite favorable, considering

the worldwide concerns regarding climate change and the recent creation of the Brazilian Forest Service, which is the head of the National System of Forest Information. On the other hand, the project of the new Brazilian NFI can also strengthen SisPP by articulating an interface between the two systems.

Literature Cited

Alder, D.; Synnott, T.J. 1992. Permanent sample plot techniques for mixed tropical forest. Tropical Forestry Pap. No. 25. Oxford, England: University of Oxford, Oxford Forestry Institute. 124 p.

Curtis, R.O.; Marshall, D.D. 2005. Permanent-plot procedures for silvicultural and yield research. Gen. Tech. Rep. PNW-GTR-634. Portland, OR; U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 86 p.

Kleinn, C.; Köhl, M. eds. 1999. Long-term observations and research in forestry. In: Proceedings, International Union of Forest Research Organizations S4.11 International Symposium. Vienna, Austria: International Union of Forest Research Organizations. 291 p.

Oliveira, Y.M.M. *et al.* 2005. Sistema nacional de parcelas permanentes: modelo metodológico. Série Documentos 106. Colombo, Paraná, Brazil: Embrapa Florestas.
[Also on CD-ROM].

Mapping Martinique's Forests and Other Natural Lands for Land Planning and Development

Rémi Teissier du Cros[1] and Claude Vidal[2]

Abstract.—The Regional Council of Martinique has chosen the French national forest inventory to realize Martinique's forest and other natural lands map. The project is divided into the three following steps: (1) nomenclature proposal and study area delineation; (2) mapping of the vegetation based on 2005 airborne orthophotographs, Geographic Information System-based slope and accessibility analysis maps, and a land cover change map based on the comparison of 2004 and 1951 airborne covers; crossed maps will be produced to show areas that can potentially be developed for agriculture, forestry, or land protection; (3) and presentation of the results.

Context

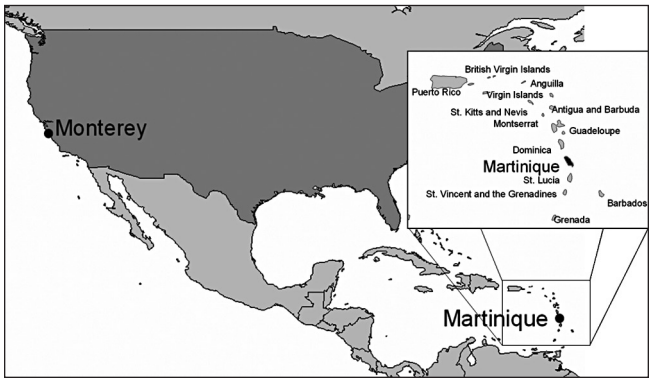
The Island

Martinique is a French overseas department in the Caribbean West Indies with the Atlantic Ocean on its east and the Caribbean Sea on its west, as shown in figure 1. It has a surface area of 1,128 km^2 (278,735 acres) and measures 70 km (43 mi) from north to south and 30 km (18 mi) from east to west. It is the smallest of France's four overseas departments. Its geographic relief is diverse, with a group of mountains in the north topped by the Pitons du Carbet (1,196 m, or 3,924 ft) and the still-active volcano of Montagne Pelée (1,397 m, or 4,583 ft). The rest of the island is hilly except along the mideastern coast, where the Lamentin Plain is located (Ministère de l'outre-mer 2005).

Climate

The island's climate is tropical, with an average annual temperature of 26 °C (78 °F), and is extremely humid from 80- to 87-percent humidity, depending on the season. Its heat, due to sunshine, is cooled by trade winds coming from the Atlantic Ocean.

Figure 1.—*Martinique's localization and a closeup of the Caribbean West Indies.*



Martinique has two seasons: a dry warm season from December to May with extensive drought from February to April and maximum sunshine, and a humid warm season from June to November with high hurricane risk.

Mountainous areas in the north have a cooler and more humid climate than shore areas. Trade winds from the Atlantic Ocean are blocked by the relief, producing important orographic precipitation. For example, average rainfall produces more than 5 m of water a year at the top of the Montagne Pelée.

Population

The most recent population census conducted in 1999 listed 381,427 inhabitants. The population density is high, with 338 inhabitants per square kilometer (875 inhabitants per square mile). Detailed data show that inhabitants are unequally distributed: three nearby cities (Fort de France, Le Lamentin, and Schoelcher) contain 40 percent of the population and the remaining

[1] Vice Director, French National Forest Inventory, Bordeaux Unit, 62 rue de Laseppe, 33000 Bordeaux, France. E-mail: teissierducros@ifn.fr.
[2] Director, French National Forest Inventory, Château des Barres, 45290 Nogent-sur-Vernisson, France. E-mail: cvidal@ifn.fr.

60 percent of the population is generally distributed throughout the southern part of the island. The mountains in the north have a low population density.

Vegetation

Martinique is covered by approximately 47,500 ha (117,375 acres) of forest; no precise data exist for other natural lands. The public-owned forest, representing around 12,000 ha (29,650 acres) is well known and managed by the French national Forest Service, whereas private forests and other natural lands are not well known. Forest fragmentation is highly variable: densely populated areas show a high level of forest fragmentation and areas with fewer population or steeper slopes have much bigger stands. A wide variety of forest types exists, ranging from dry forest to rain forest. The tree diversity is high (Palli 2007, Portecop and Petit le Brun 2003). Also present are rare forest types, including mangrove forest. Protecting the sparsest types of forest is a major issue: Whereas this land used to be naturally shaped by relief and rainfall, it is currently under a great deal of pressure stemming from urbanization and agricultural policy. All districts have used up their available lots and, in general, they look to forest areas for new urban projects. Although comprehensive development area maps exist, communities manage to obtain authorization, and, for the most part, cities are growing over the island's forests. In addition to urbanization, agriculture also represents a source of pressure for forests. More and more young farmers are unable to find land to start their businesses, while some existing farms are looking to increase their surface area. The simplest solution often involves forest clearing. On the other hand, the amount of fallow land is increasing.

A Small-Scale Wood Industry

Martinique's forest also supplies a small-scale wood industry. The wood is mainly used for local-style furniture. About 2,000 ha (4,942 acres) of mahogany forest produce approximately 4,000 m^3 (141,260 cubic ft) of wood each year. Almost all harvested areas are in public-owned forests, although politicians wish to expand harvesting in private-owned forests. Some forest-accessible stands could be cleared and replanted with mahogany; abandoned farmlands could be used in the same way.

A Lack of Forest Data

Overseas departments, unlike mainland France, do not have forest inventories. Vegetation maps (Poupon 1982, Rosette 1985) are old and inaccurate and restricted to climax vegetation (Fiard 1994, Portecop 1978).

As a result, a monitoring study of Martinique's territory is needed to improve data and maps of the island's forests and other natural lands. Politicians and institutions need a tool to manage their territory.

Purpose of the Study and Procedure

Purpose of the Study

In light of all the previous observations, Martinique's politicians decided to conduct an inventory of their natural territories, including forests, heath, and fallow lands. This study, based mostly on photo interpretation and Geographic Information System (GIS) processing, is the first step of a large-scale program conducted over many years, starting with this general study and followed by more specific studies. The main objective of this first step is to produce maps of vegetation, accessibility, slopes, and land use history on forest and other natural areas. These maps will be used to define areas that could potentially be developed toward agriculture or forestry.

The second step of this large-scale program will use the results of the previous study considering properties, environmental constraints, water resources, agricultural or forestry potentials, landscape, and tourism. Other precise thematic studies will be conducted.

The Regional Council of Martinique launched a call for tenders in May 2005 for the first phase of the program. The French national forest inventory (NFI) was chosen to conduct the study.

Procedure

The project itself is divided into three parts, as follows.

Part 1, probably the most strategic, consists of proposing different typologies to be used on the different maps requested for the project. It is associated with a field trip to train some NFI

team members who are not used to Martinique's vegetation. At the same time, a first map is produced to delimit the study area that includes forest, heath, and fallow lands.

Part 2 is the main operational part, in which all maps are compiled and the GIS map crossings are done. The maps are statistically controlled in the field to guarantee a high level of accuracy.

Part 3 involves presenting the results to the various parties concerned by the study.

After each part is completed, the work is validated by an expert group.

Results To Be Delivered

The results produced during this study are mainly GIS maps, including a 2005 vegetation map, a land use map showing changes that have occurred since 1951, and accessibility and slope maps. These maps are intersected to produce different maps showing areas with forestry or agriculture potential.

Data Available for the Study

The study is based on photo-interpretation work; most of the data were GIS and/or photographic material. Since the late 1990s, every 5 years, the National Geographic Institute has produced an accurate orthophotographic layer for the entire French territory, including the overseas land. We were able to use the two most recent layers, from 2000 and 2005. Having two layers is interesting because the quality of each layer varied greatly and it is extremely useful to use both layers on difficult cases even if the map's reference date was 2005. The only drawback is that the emulsion is not infrared but is a natural color that makes it much more difficult to see differences between vegetation types. In addition, we have paper prints of the airborne photographs to use for stereoscopic vision; the use of these prints is necessary to decide on some kinds of vegetation types related to canopy height (Boureau 2002). Another set of airborne photographs is also available for 1951 cover. It is used for a comparison with the 2005 cover to map land use changes.

All GIS processing about map accessibility and slopes use a vectored map with different themes, such as roads and rivers and a digital terrain model updated in 2006.

A complementary agricultural map is also used, showing the different types of agriculture existing in each block. The divisions of this map are based on statements by landowners, however, which may have impacted the accuracy of the map's content. Furthermore, the outlines of the map's various divisions are somewhat vaguely defined. Despite these obstacles, the map is nevertheless a useful tool that helps identify fallow lands.

All these data are supplemented with bibliographic research and discussions with local specialists.

Teams Working on the Study

Four people from the French NFI in Bordeaux are working on this project, including one project manager who used to work in Martinique's forests and three cartographers. Working with three cartographers is an important choice. The quality of work is greatly improved because the cartographers are able to compare their analyses when they have doubts, which is quite common when performing photo interpretation. We are assisted with GIS processing by another cartographer living in Martinique.

The Regional Council of Martinique is supervising the project with the assistance of the French national Forest Service. Each step of the work is validated by a group of experts that is composed of people involved in the two main problematic issues of the agriculture and forestry project.

Results and Outcomes

Study Area

The study area map sets up the perimeter of the study that will be used for the compilation of the other maps. It includes forest, heath, and fallow lands, the three main land types that are targeted by the study. The detail of what should be included in these three types is detailed using the international Food and Agriculture Organization of the United Nations (FAO) definitions (FAO 2000) for forest and heath. The last type is much more difficult to determine. Focusing on the study's objectives, fallow lands are considered to be any field where no distinct agricultural activity is present. It is important that fallow lands

be included because interpretation doubts usually exist about whether they are truly abandoned or neglected; ignoring them would be a loss of interesting areas.

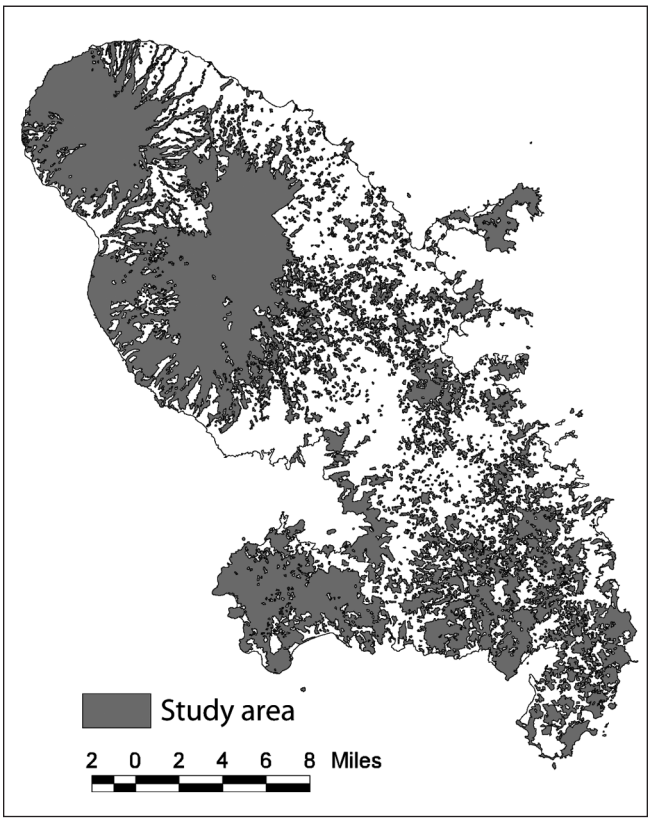
The reference for this work is the 2005 orthophotographic layer. Delineation is performed directly in ArcView. In general, the work scale is around 1:3,000, giving a good accuracy of the layout.

The next step is to determine the thresholds of this layer. The contract requests a minimum width of 75 m (246 ft) and a minimum surface area of 2.25 ha (5.56 acres) because it is common practice in mainland France (IFN 2003). Discussions reveal that these specifications were too weak for the expected outcome of the study for two reasons: (1) mapping with a 2.25-ha limit would omit a large number of fragmented areas and (2) producing a synthesis for a larger surface area would be much more time consuming. The minimum surface area is reduced to 0.5 ha, and we keep the minimum width at 75 m. As a result, the minimum surface area is almost in line with FAO recommendations, and it will be easy to present the results according to international definitions.

Practically speaking, delineation of the study area began before the field trip and was finished afterward. This practice enabled us to select areas that were difficult to identify on orthophotos and to visit them in the field. After the field trip, the study area was updated according to what we saw there.

Once the map is validated by the expert group, the next stage will continue within the study area, as shown in figure 2.

Figure 2.—*General view of Martinique with the study area in gray.*



Vegetation Types Map

The vegetation map is quite significant for Martinique: it is probably the longest part of the project and politicians and other stakeholders have high expectations about it.

The method used for mapping is an adaptation of the one traditionally used in mainland France. It is mostly based on direct delineation on a computer screen and was previously controlled with stereoscopic couple photographs (IFN 2005). The thresholds used are lower, but this does not significantly change the usual practices (see "Study Area"). Two other major differences exist between the method used for Martinique and the one used for mainland France. As is the case for the study area, 2005 orthophotos with a natural color emulsion are used. This practice is a major constraint because, in general, it is highly preferable to use an infrared emulsion to differentiate vegetation types. The second major difference is, of course, the highly diversified tropical vegetation.

Previous vegetation maps had been produced (Poupon 1982, Rosette 1985) but never with a very high level of precision in terms of layout. These maps are used to correlate our observations.

Even if the available material is an important constraint for the vegetation typology definition, 18 types are proposed (see table 1). Our analysis of the project led us to build the typology with a correspondence to the United Nations Educational, Scientific and Cultural Organization's forest classification and FAO land cover definition whenever possible for each type that was useful for the study. Once the map is established, experts and field observation will validate it.

Table 1.— *Summarized presentation of the nomenclature of plant formations that will be mapped as part of the study's geographic scope.*

Family	Name	Definition	Martinique's correspondence with adapted international UNESCO classification[a]	Correspondence with FAO definition
Forest	Mangrove	Mangrove	Mangrove	Forest
Forest	Beach forest	Beach forest	Beach forest	Forest
Forest	Young dry forest	Young forest with homogeneous low cover and high density of shrubs and small trees	Evergreen seasonal tropical forest in lower horizon	Forest
Forest	Medium dry forest	Mid-evolved forest with heterogeneous cover with shrubs and trees	Evergreen seasonal tropical forest in lower horizon	Forest
Forest	Older dry forest	More evolved stands with a homogeneous tree cover	Evergreen seasonal tropical forest in lower horizon	Forest
Forest	Semimoist forest	Forest ranging between 0 and 300 m in elevation with a pluviometry range between 1,800 and 2,500 mm/year	Evergreen seasonal tropical forest in mid/higher horizon	Forest
Forest	Rain forest	Forest ranging between 200 and 1,000 m in elevation with a pluviometry range between 2,500 and 5,000 mm/year	Ombro-evergreen seasonal tropical forest and submontane ombrophilous tropical forest	Forest
Forest	Altimontane shrubs and trees	Altimontane shrubs and trees	Shrubs and trees on higher volcanic crests	Forest
Forest	Bamboo	Bamboo cover of more than 25%	NA	Forest
Forest	Mahogany high forest	Forest with mahogany cover of more than 75%	NA	Forest
Forest	Mosaic forest/housing	Forest fragmented with housing; forest cover of more than 10 and less than 75%	NA	Other lands with tree cover
Fallow land	Fallow land with a high cover of shrubs	Tree cover of less than 10% and shrub cover of more than 40%	NA	Other wooded lands
Fallow land	Fallow land after banana trees	Tree cover of less than 10% and banana tree cover of less than 40%	NA	Other wooded lands and other lands
Fallow land	Fallow land after other croplands	Tree cover of less than 10% and other croplands cover of less than 40%	NA	Other wooded lands and other lands
Fallow land	Other fallow lands	Tree cover of less than 10% and shrub cover of less than 40%	NA	Other wooded lands and other lands
Agriculture with tree cover	Croplands with sparse trees	Tree cover over of more than 10% with an agricultural main use	NA	Other wooded lands
Heath	Savanna	Permanent heath with forest cover of less than 10%	NA	Other wooded lands and other lands
Heath heath lands	Altimontane Low shrubs; tree cover of less than 10%	Altimontane shrubs	Other lands	

[a] *Joseph (1998).*

FAO = Food and Agriculture Organization of the United Nations. NA = Not applicable. UNESCO = United Nations Educational, Scientific and Cultural Organization.

Vegetation types are separated into the following four families:

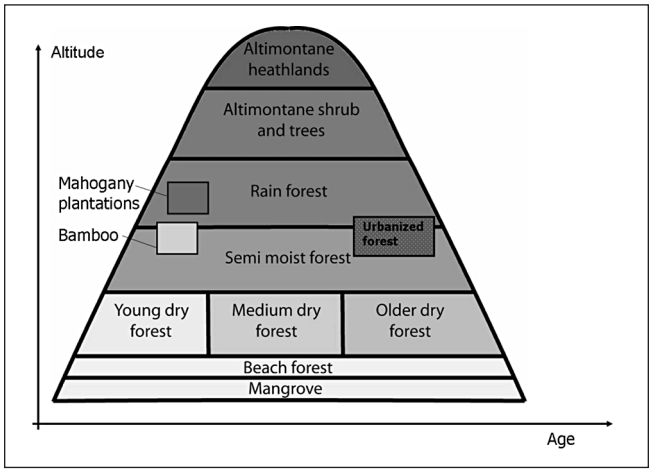
1. Forest types (figure 3**)**. Because it is impossible to distinguish the specific constitution of each type, our typology is based on what is seen from the airborne photographs using stereography. Three main factors are used to determine the type: (1) transparency of the cover, (2) average tree height, and (3) crown aspect/density. The combination of these factors with field knowledge and

external complementary information such as altitude and rainfall enable the identification of the appropriate type.

Other types, such as mangrove, mahogany plantations, and bamboo, have more specific aspects. These types are usually easily recognizable on photos. Urbanized forest areas have a dedicated type. They are highly common and have a peculiar aspect, in that they are composed mostly of species that produce food supplements. Beach forest has a specific species constitution and is subject to considerable human pressure stemming from urbanization and tourism. Only a few beach forests will be in the study area (less than 75 m in width).

2. The second family includes the different kinds of fallow lands depending on the type of agriculture that was previously practiced. Fallow lands vary greatly depending on what was present before and generally do not have the same development potential. Fallow lands that used to contain banana trees have a fast dynamism and can become forests much faster than fallow land that used to be a meadow.
3. The third family contains two kinds of heath (or high grasslands) that are not subject to high height pedoclimatic conditions.
4. The last family has just one type: areas with tree cover that are used for agriculture.

Figure 3.—*Forest typological diagram.*



Accessibility and Slopes

The purpose of the study is to underline areas with development potential. Areas accessible in terms of road servicing and areas that are not too steep to work on are extracted in two layers using the methodology described in the following text. (See figure 4.)

The first layer concerns theoretical accessibility. It is made from the vectorized road layer, with buffers created around roads that are suitable for trucks. The type of road is also attached to the buffer. Buffers have the diameter of the maximum skidding length; we used 75 m, considering that Martinique's conditions are commonly difficult (Anon. 2001). Moreover, these buffers are intersected by rivers, the most common obstacle for wood harvesting. The opposite side of the river is removed, giving us a reachable area for skidding.

The second layer is a simplification of a digital terrain model in different slope classes taken from the value of the maximum working capacity for mechanized agriculture and forestry, considering that 20 percent of the slope is the limit for mechanized agriculture and 30 percent of the slope is the limit for mechanized forestry skidding, as shown in table 2.

These two layers are then intersected to extract potentially developed area by vegetation type.

Figure 4.—*Road buffer map, slopes classes map, and vegetation map are intersected to show areas with development potential.*

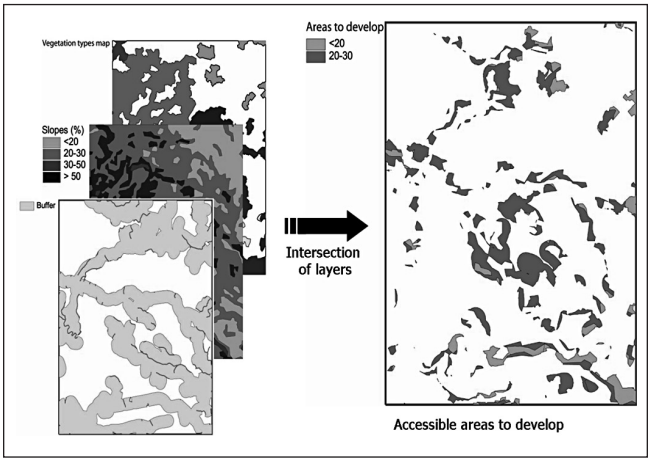


Table 2.—*Slope classes*

Slope	Definition
< 10%	Optimal for all types of use
10 to 20%	Mechanized agriculture still possible and optimal for mechanized skidding
20 to 30%	Optimal for mechanized skidding but mechanized agriculture no longer possible
30 to 50%	Mechanized skidding more complicated but still possible
> 50%	May not be used

Land Use Change

Politicians were curious to see changes in land use that have occurred over the past 50 years. A full airborne photograph cover was available. This cover had been realized in 1951 and the 2005 cover was compared with it.

We have only paper prints of the aerial cover, which we have to scan and then orthorectify to superimpose them on the 2005 cover. The 1951 cover had to be orthorectified, which presented difficulties because so many changes have occurred.

The comparison was done only within the study area (see table 3) to see the range of possibilities. This is not a comprehensive land use change map. Changes are determined on every plot that was mapped as forest or heath in 2005. Polygons from 2005 are modified according to previous land use, as observed on the 1951 photographs.

Maps Crossing

Synthesis maps are produced by intersecting the different layers. These maps are tools for finding the potential areas to develop. The maps will be used by the stakeholders to enable them to focus on the interesting areas.

Current Status of the Project

At the moment when this article was written, the study area had been delimited and the forest types had been listed. They have been submitted for validation by local administrations. The 1951 airborne photographs have been orthorectified. Tests have been carried out for every layer and all of our methodological problems have been resolved.

Conclusion

This project is a local tool for improving land planning and the management of the island. It will have numerous effects on various topics. This is the first time that Martinique will have highly accurate maps for assessing its policy concerning land planning and management for both forestry and agriculture. Many people are waiting for the results of the project and it will be the basis for many other studies. These are some of several reasons why our work has to be perfect.

If we consider the impact of this study in France, forest statistics for one of the country's overseas territories are going to be relevant for the first time. If this study is carried out on the other territories, France will have good data for its entire surface area. It will be possible to use the results to answer to international questionnaires about forests, such as the Forest Resources Assessment for FAO or the Kyoto Protocol. At this time, only two French departments, Martinique and French Guyana, are being inventoried. In these two locations, the inventory is more focused on remote sensing. Other departments are observing our project to see if it turns out well.

On the whole, this project could be considered a pilot study not only for France but also for other tropical islands with similar

Table 3.—*Land use change table between 1951 and 2005.**

Land use in 1951	Land use in 2005		
	Forest	Heath	Agriculture
Forest	1	1	0
Heath	1	1	0
Agriculture	1	1	0

**1 = mapped; 0 = not mapped.*

conditions. It is important for us to mutually communicate with the Regional Council of Martinique and the French NFI. This project will be useful to show to the decisionmakers the advantages offered by such a program in similar territories.

Literature Cited

Anon. 2001. Manuel d'exploitation forestière. Tome 2. Paris: Association Forêt Cellulose, Centre Technique du bois et de l'Ameublement, Mutualité Sociale Agricole. 415 p.

Boureau, J.G. 2002. Les photographies aériennes, leur utilisation par l'Inventaire Forestier National Manuel de photo-interprétation. Nogent-sur-Vernisson, France: Inventaire Forestier National. 152 p.

Fiard, J.P. 1994. Les forêts du nord de la montagne Pelée et des édifices volcaniques du Piton Mont Conil et du Morne Sibérie (Martinique). Pointe-à-Pitre, Guadeloupe, France: Université des Antilles et de la Guyane. 520 p.

Food and Agriculture Organization of the United Nations (FAO). 2000. The forest resources assessment programme: 2.1 Land cover. http://www.fao.org/docrep/007/ae217e/ae217e00.htm. (10 January 2007).

Inventaire Forestier National (IFN). 2003. Nomenclature de l' Inventaire Forestier National. Nogent-sur-Vernisson, France: Inventaire Forestier National. 46 p.

IFN. 2005. Instruction pour la cartographie et la photo-interprétation. Édition 2005. Nogent-sur-Vernisson, France: Inventaire Forestier National. 145 p.

Joseph, P. 1998. Contribution à la nomenclature de l'UNESCO pour les forêts de la Martinique et des Petites Antilles. Paris: Terre d'Amérique, GEODE Caraïbe-Karthala 1: 269-303.

Ministère de l'outre-mer. 2005. Géographie de la Martinique. http://www.outre-mer.gouv.fr/outremer/front?id=outremer/decouvrir_outre_mer/martinique/publi_P_presentation_geographie_1123062435443. (10 January 2007).

Palli, F. 2007. Arbres de la Martinique les arbres de la forêt martiniquaise—répertoire complet. http://www.palli.ch/~kapeskreyol/divers/arbres.html. (10 January 2007).

Portecop, J. 1978. Phytogéographie, cartographie écologique et aménagement dans une île tropicale: le cas de la Martinique. Grenoble, France: Thèse d'état, Université de Grenoble. 377 p.

Portecop, J.; Petit le Brun, T. 2003. A la découverte des Arbres des Antilles. Tomes I et II. Gosier, Guadeloupe, France: PLB Éditions. 64 p.

Poupon, J. 1982. Carte forestière de la Martinique. Martinique, France: Office National des Forêts, Direction Régionale de la Martinique. 24 p.

Rosette, C. 1985. Inventaire cartographique des zones à vocation forestière du sud de la Martinique. Fort de France, Martinique, France: Office National des Forêts. 78 p.

Additional Reading

Maillard, J.F. 2004. Orientations régionales de gestion de la faune sauvage et d'amelioration de la qualite de ses habitats. Fort de France, Martinique, France: Direction Régionale de l'Environnement—Martinique. 86 p.

The Finnish National Forest Inventory

Erkki Tomppo[1]

Abstract.—The National Forest Inventory (NFI) of Finland has produced large-area forest resource information since the beginning of 1920s (Ilvessalo 1927). When the 10th inventory (NFI10) started in 2004, the design was changed and the rotation shortened to 5 years. Measurements are done in the entire country each year through measuring one-fifth of the plots. About one-fifth of all plots are measured as permanent. Using field data only, it is possible to compute reliable estimates for large areas, the minimum size of the area is typically some hundreds of thousands of hectares. In practical forestry, estimates are also often required for smaller units such as municipalities with typical areas of tens of thousands of hectares. This is possible only if ancillary data are used in addition to sparse field data. The Finnish multisource NFI uses satellite images and digital map data, in addition to field data, and produces estimates for small areas and wall-to-wall maps. Information from the Finnish NFI has traditionally been used in large area forest management planning, such as planning regional and national level cutting, improving silviculture and forest regimes, making decisions concerning forest industry investments, and providing a basis for forest income taxation. The NFI also provides forest resource information for national and international forest statistics and processes such as the United Nations Food and Agriculture Organization's (FAO) Forest Resource Assessment process and the Land Use, Land-Use Change and Forestry reporting of the United Nations Framework Convention on Climate Change.. Sampling designs for the ninth inventory rotation (NFI9)—conducted from 1996 to 2003—and NFI10 are described, as well as the basic principles of estimation methods based on field data only.

Introduction

The sampling design and plot- and stand-level measurements have been changed over time to respond to contemporary requirements and to optimize the use of the available resources. The sampling system in the first NFI was line-wise survey sampling, introduced by Professor Yrjö Ilvessalo (Ilvessalo 1927). The line interval was 16 km in most parts of the country, but, for error estimation purposes, an interval of 13 km was used in one province and 10 km in the Åland Islands. Plot measurements were conducted in line strips 10-m wide. The plot length was 50 m and the interval between plots was 2 km. Similar sampling systems with different sampling intensities were used in the following three inventories up to 1963. Detached tracts have been used instead of continuous lines since the fifth inventory (1964–1970) (Kuusela and Salminen 1969). At the same time, the inventory became a continuous operation, proceeded by regions from south to north. The fixed-size sample plots were also changed to Bitterlich plots (angle gauge plots, or probability proportional to size sampling, determined the size of the plot based on the basal area of a tree at breast height). A new feature in the fifth, sixth, and seventh inventories was the use of aerial photographs in northern Finland (Poso 1972). Two-phase stratified sampling (stratification based on aerial photographs) was used in the fifth and sixth inventories and photo interpretation plots in the seventh inventory.

The ground sampling intensity has been adapted to the variability in forests, taking into account the necessary budget constraints. The sampling intensity in northern Finland has thus been lower than that in southern Finland. About one fifth of the sample plots have been made permanent since the eighth inventory in northern Finland (1992–94), and the establishment of such plots was completed for the entire country in NFI9. The aim is to be able to obtain information of a kind that cannot be derived from temporary plots (e.g., the amount and structure of the drain, detailed changes in

[1] Professor of Forest Inventory, Finnish Forest Research Institute, Vantaa Research Centre, Unioninkatu 40, FIN-00170 Helsinki, Finland. E-mail: erkki.tomppo@metla.fi.

land use, and other changes taking place), and also to reduce the standard error of some estimates. The length of each cycle, comprising one complete inventory, has been dependent on the funds granted in the national budget, the smallest area unit for which results are required, and the statistical precision of the estimates that is considered desirable. The first four inventory rotations took about 3 years each, while the next five took 6 to 9 years each. The rotation was shortened to 5 years starting from NFI10, which started in 2004. The sampling design was slightly modified at the same time.

The main administrative unit for forestry in Finland is the Forestry Center district, commonly comprising 0.8 to 5.0 million ha of forest land. The mainland is divided into 13 such districts, with the Åland Islands forming an additional district. The standard error of the estimated growing stock volume for these districts is between 2.7 and 1.9 percent, and that for the entire country is 0.6 percent (Tomppo *et al.* 1997, 2001; Tomppo 2006).

The information generated by the NFI has traditionally been used for large-area forest management planning (e.g., in the planning of cutting, silviculture, and forest improvement regimes at the regional and national levels), in decisions concerning forest industry investments, and as a basis for forest income taxation. It has also provided forest resource information for national and international statistics such as the FAO Forest Resource Assessment process and the Ministerial Conference on the Protection of Forests in Europe. It currently also produces information on forest health status and damage, biodiversity and carbon pools, and changes in these for the Land Use Land Use Change reports of the United Nations Framework Convention on Climate Change. The NFI covers all forests and the information has been used by all ownership groups for justifying and calibrating their own results. It serves as a central information source and tool for use in forestry, the forest industry, and forest environment decisions and policymaking.

Field Sampling System Used in NFI9 and NFI10

The sampling unit used in NFI9 was a cluster, also referred to as a tract. The sampling design was adapted to the variability in the forests, the distances between two tracts varying from 6 by 6 km in the southernmost part of the country to 10 by 10 km in Lapland. The sampling density regions, six regions, together with field plot clusters are shown in figure 1. The NFI9 sampling designs, cluster sizes, and distances between clusters, in the southernmost part of the country, central Finland, north-central Finland, south Lapland, and north Lapland are shown in figure 2 (figs. 2a, 2b, 2c, 2d). The distances between clusters were 10 by 10 km in the municipality of Kuusamo and in south Lapland, and 7 by 7 km elsewhere in north-central Finland.

Figure 1.—*The sampling density regions of the NFI9 and NFI10 with the field plot clusters of NFI9.*

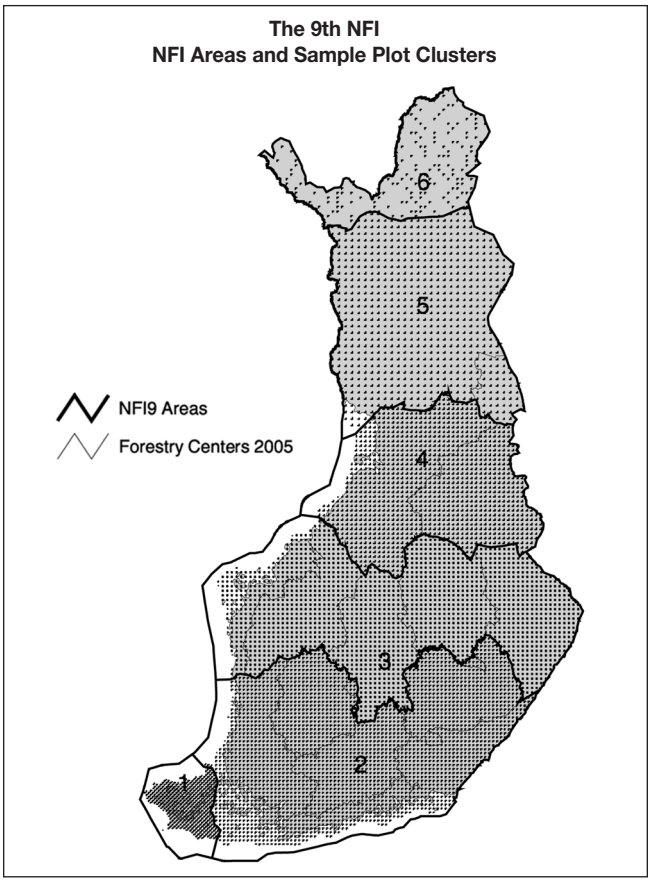
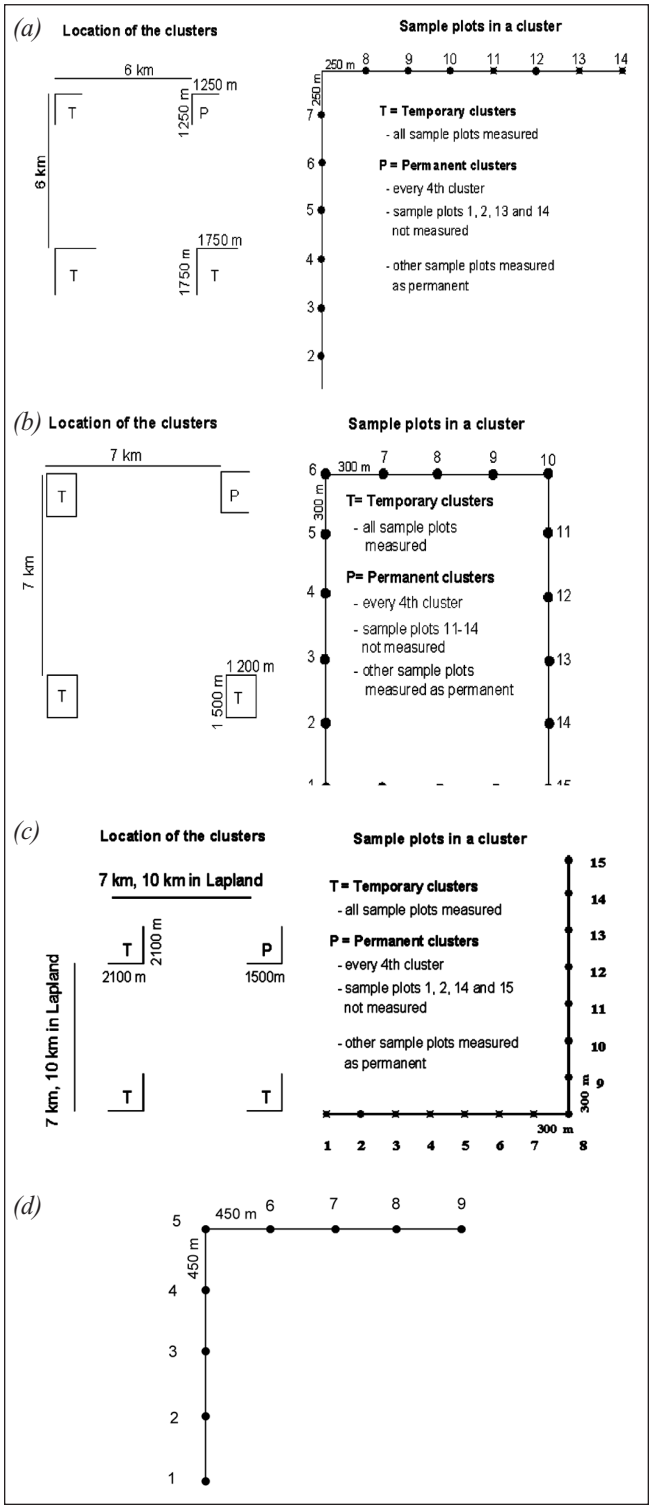


Figure 2.—*Sampling design of the NFI9 in different inventory regions: (a) region 2 (in region 1, the design is same but the distances are 3 by 3 km), (b) region 3, (c) region 4 (in region 5, the design is same but the distances are 10 by 10 km, (d) region 6. Stratified sampling was applied in region 6.*



The two-phase stratified sampling applied to the area of the three northernmost municipalities was based on three variables: (1) the percent of waste land (e.g., open bogs and very poor mineral sites such as open rocks), (2) the volume of growing stock, and (3) predicted cumulative day-time temperature. The two first variables were predictions of multisource forest inventory in a form of thematic maps.

Note that in the estimation, except error estimation, a field plot is a single observation; in the error estimation, a field plot is a cluster.

Satellite image-based digital volume maps and sampling simulations were used to evaluate different sampling designs. For each design tested, 1,000 samples were chosen and standard deviations for the mean volume were computed (Henttonen 1991) and assumed to represent the standard error in mean volume. Another quite important aspect was that a sampling unit (cluster) should represent 1 day's work on average. It was found that the optimum design depended on the distribution of forest land and the heterogeneity of the forests, for instance, and, therefore, varied from south to north and from east to west. The sampling intensity was adapted to the spatial variation in forests throughout the whole country, being lower in the north than in the south.

The progress of the inventory was changed somewhat for NFI10 (2004–08). Measuring field plots in the entire country each year was the biggest change compared to the NFI5 - NFI9. One-fifth of the clusters is measured annually. Exceptions are region 1, which was measured in 2007, and region 6, which will be measured next time during NFI11. At the same time, the inventory rotation was shortened to 5 years, a reduction of almost 5 years. Inventory progressed by regions from the fifth inventory (1964–70) until NFI9 (1996–2003). This change makes it possible to compute the basic forest resource estimates annually for the entire country. NFI9 data for permanent plots will also be used in estimation. The method is thus sampling with partial replacement.

The basic principles in sampling designs of NFI10 are similar to those of NFI9. The need to shift the locations of the clusters with temporary plots caused some changes. The

clusters with temporary plots were shifted 1 km west and 1 km north in this rotation (fig. 3—figs. 3a, 3b, 3c, 3d). Shortening of the inventory rotation also adds pressure to reduce the number of field plots slightly. New sampling simulation studies were carried out in all parts of the country for the forest area, mean volumes by tree species (m^3/ha), and total volumes (m^3) variables. The numbers of the field plots per cluster in different regions are shown in figure 3.

The sample plot was a Bitterlich plot (angle-gauge plot) in both NFI9 and NFI10. Tally trees are selected with a relascope, the basal area factor being 2 in southern Finland and 1.5 in northern Finland. The maximum radius is 12.52 m and 12.45 m, respectively (corresponding to breast height diameters of 34.5 cm and 30.5 cm, respectively). Where a relascope could not be used, inclusion was checked by measuring the distance and diameter of the tree at a height of 1.3 m. Reducing the radius of a sample plot detracts very little from the reliability of the estimates, but it does ease the amount of fieldwork noticeably in some cases, as the number of divided sample plots (i.e., sample plots belonging to two or more stands or strata) decreases. The use of maximum distance may also reduce errors caused by possible unobserved trees, usually located a long distance from the plot centre and behind other trees. Every seventh tally tree is measured as a sample tree (fig. 4).

In the Finnish NFI schema, forestry land is divided into productive forest land, poorly productive forest land, unproductive forest land (also called waste land), and forestry roads. Note that the national definitions of both forest land and poorly productive forest land deviate from the definitions of forest land and other wooded land of the FAO (2001), although the FAO definitions are currently used in parallel with the national definitions in the Finnish NFI.

The number of field plots on land in NFI9 was 81,249 in the entire country, of which 67,264 were on forestry land. Of the plots on forestry land, 62,266 were on combined forest land and poorly productive forest land, and 57,457 on forest land alone. Note that the land area and water area by municipalities are assumed to be known and the figures are based on the statistics of Land Survey Finland (Land

Figure 3.—*Sampling designs of the NFI10 in different inventory regions: (a) region 2 (in region 1, the design is same but the distances are 3 by 3 km), (b) region 3, (c) region 4 (in region 5, the design is same but the distances are 10 by 10 km,(d) region 6. Stratified sampling will be applied in region 6.*

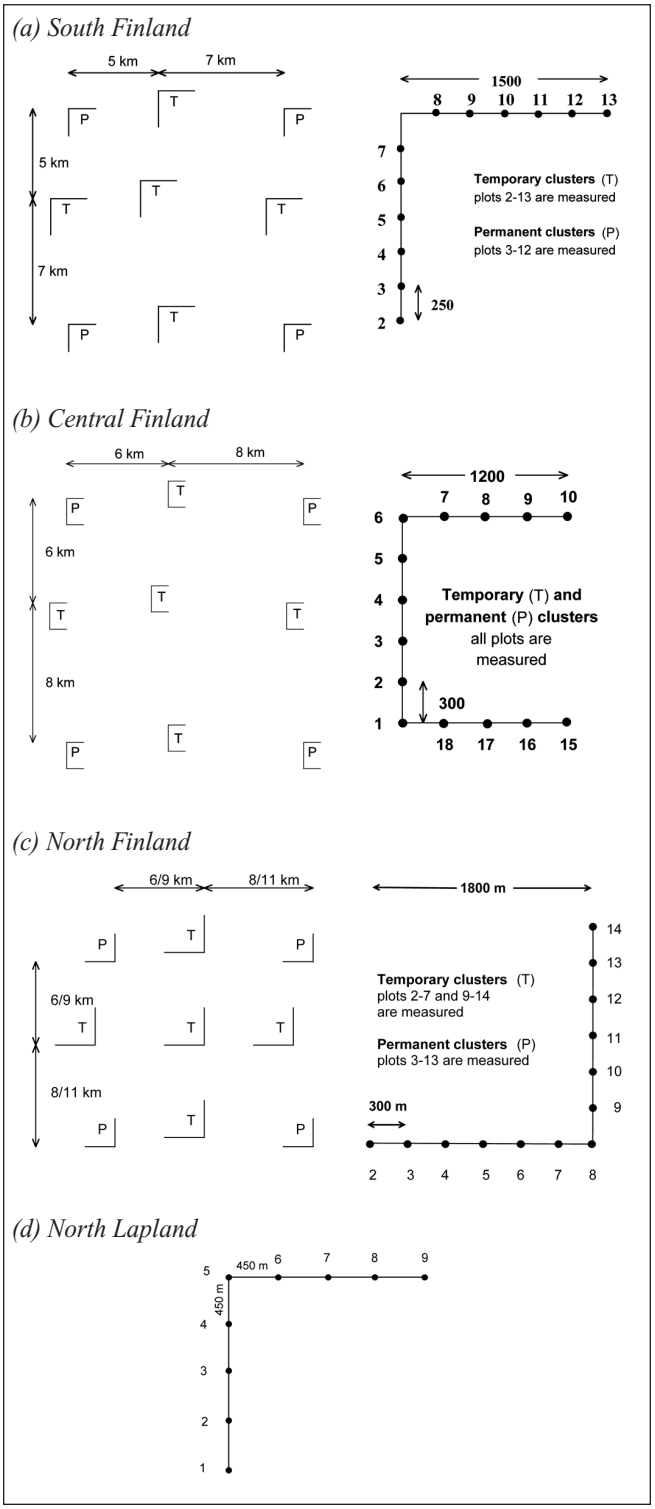
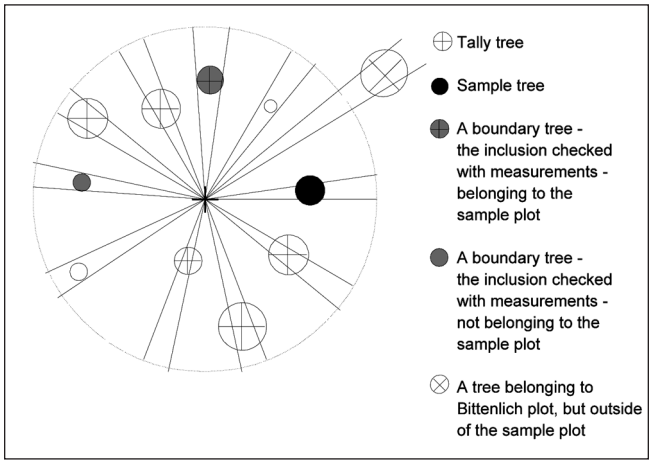


Figure 4.—*A sample plot as used in NFI9. The maximum radius for trees to be counted was 12.52 m in southern Finland (q = 2) and 12.45 m in northern Finland (q = 1.5). Every seventh tree is measured as a sample tree. The trees are counted by crews, starting at the beginning of the field season.*



Survey Finland 2003). The field plots were geolocated with geographic positioning system receivers and trees are measured on those plots that contained forest land and/or poorly productive forest land.

Estimation Based on Field Data

Some basic principles of the estimation in NFI9 are given. The major change in NFI10 and in the following inventories will be the use of permanent plots from earlier rotations and estimation based on framework of sampling with partial replacement. The final estimates are made based on (1) estimates derived from NFI10 data, and (2) estimates derived from NFI9 data and changes on NFI9 and NFI10 data. The two estimates are inversionally weighted proportionally to their estimated variances (Scott 1984, Ranneby 1995). The use of post-stratification based on multisource output map data will also be tested (McRoberts *et al.* 2002).

The NFI results can be divided into area, volume, and increment estimates. The NFI plots cover the entire land area of the country and its waterways, so that the inventory produces area estimates not only for forestry land strata but for all land use classes. Forest land and forestry land are divided into subcategories on the basis of site, ownership, silviculture, and cutting regimes, growing stock and needed treatments (e.g., tree species composition, age, and mean diameter of trees).

Area Estimation

Area estimation is based on the total land area and inland water areas that are known or assumed to be error free and on the number of center points of the plots. In brief, the area estimate of a land stratum is the number of plot centers in the stratum divided by the total number of plot centers and multiplied by the known land area. Due to the fact that the number of plot centers on land is a random variable (depending on the design), the area estimators are ratio estimators (Cochran 1977)

$$a_S = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i} A = \frac{\bar{y}}{\bar{x}} A, \qquad (1)$$

where:
a_s is the area estimate of the stratum s, A is the land area on the basis of the official statistics of the Finnish Land Survey (Land Survey Finland 2003), y_i is 1 when the center point of the plot belongs to the stratum in question and 0 otherwise, x_i is 1 when the center point is on land and 0 otherwise, and n is the number of center points on land (Tomppo *et al.* 1997, 1998, 2001; Tomppo 2006). Examples of land strata are forest land, spruce-dominated forest land, and forest land thinned during the past 10 years.

Volume Estimation

Volume in the Finnish NFI means tree stem volume over bark (i.e., with bark) from above the stump to the top of the tree, excluding branches. All trees of at least 1.3 m of height (i.e., breast height diameter > 0 cm) are included in the volume estimate. The volume estimators are ratio estimators similar to the area estimators (eq. 1). Briefly, to obtain the mean volume for a given stratum, the mean volumes of all trees belonging to that stratum are added and divided by the number of field plot center points in the stratum. The mean volume of a tree is the volume per hectare represented by the

tree (see equations [3a] and [3b]). The indicator variable y_i in the numerator of equation (1) is replaced with the mean volume represented by a tree, or the mean volume of timber assortment class of interest represented by the tree, on field plot i when computing mean volume or total volume estimates. For total volumes, the mean volumes have to be multiplied by the area estimate for the stratum in question.

The mean volumes (m^3/ha) and total volumes (m^3) are estimated as follows:

1. Volumes and volumes by timber assortment classes are predicted for sample trees (every seventh tally tree) using volume and taper curve models (Laasasenaho 1982) and sample tree measurements (Tomppo *et al.* 1997, 1998, Tomppo 2006).
2. The volumes of tally trees are predicted by strata using the volume predictions for the sample trees and measured and observed tally tree, stand, and site variables.
3. Mean volumes are tabulated by computation strata.
4. Area estimates are calculated for the volume strata.
5. Total volumes are tabulated by computation strata.

When using Bitterlich sampling (angle-gauge plots), each tree represents the same basal area per hectare. It is thus convenient to work with quantities called form heights rather than single tree volumes when computing mean volumes or total volumes. Form height is defined as

$$fh = \frac{v}{g}, \tag{2}$$

where v is the volume of a tree stem (or the volume of a timber assortment in a tree) and $g = \pi d_{1.3}^2 / 4$ is the intersectional area of the tree at breast height.

Form heights are predicted for tally trees by the nonparametric *k*-Nearest Neighbour (k-NN) estimation method. For each tally tree whose volumes are to be predicted, the *k*-nearest sample trees are sought, the distance metric applied being Euclidean distance in the space of tree-level variables, tree species, $d_{1.3}$, and tree quality class, and stand-level variables, region code, cumulative day time temperature, site fertility class, and stand establishment type.

The mean volume (m^3/ha) represented by a tree identified using angle-gauge sampling is

$$u = q\, fh. \tag{3a}$$

The maximum distance from the plot center assigned to tally trees is 12.52 m in southern Finland, where q = 2, and 12.45 m in northern Finland, where q = 1.5. Trees thicker than 34.5 or 30.5 cm, respectively, are counted in a fixed-radius plot of area $a = \pi R^2$, where R is the maximum distance. The mean volume represented by this type of tree is

$$u = \frac{g}{a} fh, \tag{3b}$$

where g is the basal area of the tree, $g = \pi d_{1.3}^2 / 4$.

The mean volume (m^3/ha) of a stratum is estimated in NFI9 using the formula

$$v_S = \frac{\sum_{i=1}^{n} \sum_{k=1}^{n_i} u_{i,k}}{\sum_{i=1}^{n} x_i}, \tag{4}$$

where v_s is the estimate for the mean volume of a stratum S, n is the number of center points of plots on land in the region, $u_{i,k}$ is the mean volume represented by tree k in stratum S on plot i, x_i is the number of trees in stratum S on plot i, and x_i is 1 if the center of plot i belongs to stratum S and 0 otherwise. The total volume estimate is

$$V_S = v_S a_S \tag{5}$$

where a_s is the estimate for the area of the stratum.

Note that the method takes into account plots shared between two or more calculation strata, so that trees belonging to the stratum in question in parts of a plot that do not include the center are also included in the sum in equation (4). It is assumed in volume estimation that the plot parts are distributed purely randomly between any two arbitrary strata s_1 and s_2. That is, for plots whose centerpoints belong to s_2 , the expected area of the plot parts belonging to s_1 is the same as the area of the plot parts belonging to s_2 whose center points belong to s_1.

Increment Estimation

Volume increment in the Finnish NFI means the increase in tree stem volume over bark, from above the stump to the top of the tree. The annual volume increment is calculated as an average over 5 years, based only on full growing seasons, assuming that tree growth has finished by August 1. Thus the increments in the 5 years preceding the inventory year are used for trees measured before August 1, and those in the inventory year and the 4 preceding years for trees measured on or after August 1.

The following phases are used in calculating the volume increment of a stratum:

1. Prediction of the annual increments in sample trees.
2. Calculation of the average increments for sample trees by diameter classes (at 1-cm intervals) and by strata (e.g., land use classes, site fertility classes, and tree species groups).
3. Calculation of the total increment for survivor trees in each stratum by diameter classes, by multiplying the average increment for trees in each diameter class by the number of tally trees in that class and summing the increments over the diameter classes.
4. Calculation of the final increment adding the drain increment to that for the survivor trees.

The details are given in Tomppo (2006) and Kujala (1980).

Conclusions

The sampling designs of NFI9 and the current NFI10 are described, as well as basic estimation principles. The NFI was changed from regionwise, progressing to a rolling system for NFI10, causing some minor changes in the design as well.

The sampling design was selected on and modified on the basis of experience and information gathered from the previous inventories. Sampling simulation studies were conducted in all the inventory regions to optimize the design, given acceptable maximum standard errors in the mean volume and total volume of growing stock and estimated measurement costs.

The estimation methods gained their current form during previous inventories and through experience accumulating since the 1920s. Some basic facts affecting the estimation methods are that NFI9 was based on temporary plots (permanent plots were established in the course of that survey, or in NFI8 in the case of northern Finland). In both NFI9 and NFI10, the land area is assumed to be known, and the tally tree plot is an angle-gauge plot (Bitterlich plot). Both the area and volume estimators are ratio estimators. Area estimation is based on the number of center points of plots.

In volume estimation, all trees belonging to the stratum in question are counted, including trees on parts of a plot that do not include the center point. All trees are assigned to calculation strata in the field measurements. Increment estimation is based on increment borings and height increment measurements performed on sample trees (i.e., height increment models in the case of broadleaved trees).

NFI10 began in 2004 and is proceeding in a different way from NFI9, with one-fifth of the plots in the entire country being measured each year. Thus country-level estimates can be updated annually and region-level estimates updated within 2 to 3 years of the start of the survey. An estimation method sampling with partial replacement will be used. The method used for estimating the standard errors in the area and volume calculations is based on the ideas presented by Matérn (1960) and is described and discussed in detail by Heikkinen (2006).

Literature Cited

Cochran, W.G. 1977. Sampling techniques. 3rd ed. New York: John Wiley. 428 p.

Food and Agriculture Organization of the United Nations (FAO). 2001. Global Forest Resource Assessment 2000. FAO Forestry Paper 140. Rome: FAO. 429 p.

Heikkinen, J. 2006. Assessment of uncertainty in spatially systematic sampling. In: Kangas, A.; Maltamo, M., eds. Forest inventory, methods and applications. Managing forest ecosystems Vol. 10. London: Springer Dordrecht: 155-176.

Henttonen, H. 1991. VMI8:n Pohjois-Suomen otanta-asetelmien vertailu satelliittikuvatulkinnan avulla. Comparing of sampling designs through simulation. Finnish Forest Research Institute. 10 p. In Finnish.

Ilvessalo, Y. 1927. The forests of Suomi Finland. Results of the general survey of the forests of the country carried out during the years 1921–1924. Communicationes ex Instituto Quaestionum Forestalium Finlandie 11. In Finnish with English summary.

Kujala, M. 1980. Runkopuun kuorellisen tilavuuskasvun laskentamenetelmä. Summary: a calculation method for measuring the volume growth over bark of stemwood. Folia Forestalia. 441: 1-8.

Kuusela, K.; Salminen, S. 1969. The 5th National Forest Inventory in Finland. General design, instructions for field work and data processing. Communnications Instituti Forestalis Fenniae. 69(4): 1-72.

Laasasenaho, J. 1982. Taper curve and volume functions for pine, spruce, and birch. Seloste: Männyn, kuusen ja koivun runkokäyrä- ja tilavuusyhtälöt. Communicationes Instituti Forestalis Fenniae. 108: 1-74.

Land Survey Finland. 2003. Finland's land area by municipalities on January 1, 2003. http://www.nls.fi/kartta/julkaisu/alat2003.html. (1 January 2003).

Matérn, B. 1960. Spatial variation. Meddelanden från statens skogsforskningsinstitut. 49: 1-144.

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

Poso, S. 1972. A method of combining photo and field samples in forest inventory. Communnications Instituti Forestalis Fenniae. 76(1).

Ranneby, B.; Rovainen, E. 1995. On the determination of time intervals between remeasurements of permanent plots. Forest Ecology and Management. 71: 195-202.

Scott, Ch. T. 1984. A new look at sampling with partial replacement. Forest Science. 30: 157-166.

Tomppo, E. 2006. The Finnish National Forest Inventory. In: Kangas, A.; Maltamo, M., eds. Forest inventory, methods and applications. Managing forest ecosystems. Vol. 10. London: Springer Dordrecht: 179-194.

Tomppo, E.; Henttonen, H.; Korhonen, K.T.; Aarnio, A.; Ahola, A.; Heikkinen, J.; Ihalainen, A.; Mikkelä, H.; Tonteri, T.; Tuomainen, T. 1998. Etelä-Pohjanmaan metsäkeskuksen alueen metsävarat ja niiden kehitys 1968-97. Julkaisussa: Etelä-Pohjanmaa. Metsävarat 1968-97, hakkuumahdollisuudet 1997-2026. Metsätieteen aikakauskirja - Folia Forestalia 2B/1998: 293-374. In Finnish.

Tomppo, E.; Henttonen, H.; Tuomainen, T. 2001. Valtakunnan metsien 8. inventoinnin menetelmä ja tulokset metsäkeskuksittain Pohjois-Suomessa 1992-94 sekä tulokset Etelä-Suomessa 1986-92 ja koko maassa 1986-94. Metsätieteen aikakauskirja - Folia Forestalia 1B/2001: 99-248. In Finnish.

Tomppo, E.; Varjo, J.; Korhonen, K.; Ahola, A.; Ihalainen, A.; Heikkinen, J.; Hirvelä, H.; Mikkelä, H.; Mikkola, E.; Salminen, S.; Tuomainen, T. 1997. Country report for Finland. In: Study on European forestry information and communication systems. Reports on forestry inventory and survey systems. Vol. 1. European Commission: 145-226.

The Reporting Revolution—The Southern Endeavor

Mark J. Brown[1]

Abstract.—The need for expeditious portrayal of statewide inventory findings is paramount. Demand is intensifying. Yet, to date, relaying data results and analysis through traditional publications has been extremely time consuming. To address this issue, southern forest inventory and analysis (FIA) reporting is in transition. This article discusses the evolution of authorship, reporting formats, and incorporating nontraditional topics and presents examples of new instruments, such as "Factsheets" and "Congressional Corner" Web sites, and the advantages to address timely output. This article summarizes ventures into wildland-urban interface analysis, National Woodland Owner Survey categorization, invasive/exotic species identification, National Forest System reporting, and nontimber-product use arenas beyond regular reporting. Finally, differences in key issues and other considerations identified that impede standardized format goals are discussed. Beyond the hurdle of transitioning from periodic to annual reporting, the future appears brighter; however, currently mandated timeframes and unresolved format consensus are leading to more succinct methods of output under present organizational staffing.

Introduction

As the population of the United States increases, pressures and demands on its forest resources rise as well. The need for information about the status of these forest resources is critical. Data from the forest inventories conducted by the Forest Inventory and Analysis (FIA) work units of the Forest Service, U.S. Department of Agriculture are sought by a wide variety of clientele. These end users highly anticipate new data after it is completed at the field level. Demand for access to the data and expectations of analytical findings reported are intensifying and time is of the essence. Simultaneously, FIA has implemented a plethora of changes to the inventory it conducts. These changes stemmed from Congressional mandates toward a national system and involved plot design, sample intensity, methods, cooperators, and frequency of the inventory. As a result of these changes, FIA has been in transition in its data acquisition, processing, and reporting. To accommodate mandated timeframes and more frequent data output, analytical reporting is evolving.

The purpose of this article is to summarize southern FIA endeavors to maintain traditional reporting while incorporating new topics and issues relevant to forest resources, all the while being more expedient in the process. This effort is a revolution at best and a challenge at the least. A synopsis is presented of where southern FIA has been in reporting, what southern FIA has done and is doing to address the evolving challenges, why they are doing these things, and when they expect to achieve the goals.

Evolution of Reporting (Traditional, Authorship, and Formats)

Where southern FIA has long focused its reporting revolves around traditional forestry related subjects. Southern FIA primarily reported on area, ownership, treatments, volume, growth, removals, mortality, and timber product output (TPO) for decades. Many reports were produced under sole authorship, but gradually, multiple authorship was used. Still, these authors were usually within the southern FIA unit. Now it is common, if not expected, to include a State cooperator in the authorship. In addition, other FIA units or other agencies may be included as coauthors.

Report formats changed gradually from black and white with typewritten text and tables to color covers with a few text-related graphs along with the tables. This format with a color cover continued to be used for several survey periods. Begin-

[1] Forester, U.S. Department of Agriculture, Forest Service, Southern Research Station, Forest Inventory and Analysis Unit, 4700 Old Kingston Pike, Knoxville, TN 37919. E-mail: mbrown03@fs.fed.us.

ning in the 1990s, color was incorporated throughout, including graphics. This era saw growing use of more graphs and maps within the publications. Historically, the reports were mostly text documents, whereas now a preponderance of the findings is presented through the use of maps, charts, and graphs, which often outweigh the text.

Evolution of Report Types

What southern FIA has done to address the reporting challenge is to deviate in the report types used and expand the topics covered. Formerly, report types largely adhered to unit, State, TPO, and analytical reports. The transition to annual inventories and the associated frequency of output has created or coincided with the need to produce or utilize other instruments for data reporting. These include National Forest System (NFS) reports, newsletters, quarterly magazines, brochures, fact sheets, catastrophic assessments, specialty reports, and Web links to quickly provide access to key data or explanatory caveats. The reporting types used varied depending on end-user needs to be addressed or for publication expediency requirements. Individual survey unit reports ceased along with the last periodic inventories. Interim reports were published as requested by a State, such as the "South Carolina's Forest Resources—2000 Update" (Conner and Sheffield 2001), which was based on three panels of annual inventory data.

Analytical reports continue to require the most effort and consume the most time for publication. They have expanded over time from reporting on traditional timber related subjects to encompassing many more topics related to the forest resource that will be discussed later in the article. TPO report formats (Bentley and Lowe 2006) have changed the least over time, as have their data collection. They also have remained the most consistent in delivery time, perhaps a testimony to the impacts of change or lack thereof on the overall reporting process.

The NFS reports (Oswalt 2005) are new to southern FIA. They are designed to report on all the national forests within a single State as opposed to individual forest reports. Plot intensification was necessary to achieve adequate sample sizes. Future inclusion of these reports as chapters within the analytical reports is being debated.

Southern FIA reinstated the use of a newsletter, titled *The Inventory* (USDA Forest Service 2006e), which is used to keep its constituents abreast of field inventory status, survey updates, data posting, and other FIA information and news. Southern FIA periodically uses the Station's quarterly magazine, *Compass* (USDA Forest Service 2006b), as an outlet for articles involving survey findings and other research derived from inventory data. More recently, brochures were developed to identify threats to southern forests, clarify data access, and summarize what FIA does and can provide. One brochure was produced in conjunction with the Southern Forest Research Partnership (Southern Forest Research Partnership, Inc. 2006) and the others were produced internally.

Factsheets (e.g., USDA Forest Service 2006a) are one of the most recent instruments created to quickly portray many key results from a State inventory. Factsheets for 10 of the 13 southern States have been completed and posted on the Web this year. This instrument has proven to be the quickest approach to data output thus far. They have been well received and feedback indicates high value for legislative and executive summaries. In-house production and reproduction of factsheets on an immediate or as-needed basis offer tremendous advantages in time, cost, and distribution. Initial attempts in forming factsheets sought a standard design and content, but differences between the States in issues and economies required flexibility in the type of data reported.

Catastrophic assessments (e.g., Glass and Oswalt, n.d.) and reports, although not new, are much improved. After hurricanes and major ice storms, urgent demand exists for data. With this emphasis and the advent of factsheet style in reporting, assessments are being produced in a timely manner. Southern FIA intermittently produces specialty reports, such as the avian diversity study conducted on St. John (Oswalt *et al.*, in press), but time spent on routine inventory reporting limits these opportunities.

Lastly, southern FIA is using Web sites or URL links to make key State-level data or explanations of use quickly accessible. On the Station's Congressional Corner (USDA Forest Service 2006c) for instance, totals for key data items are provided by clicking on an individual State. In Mapmaker (USDA Forest Service 2006d), a link was added in the selection window for

the State of Louisiana to explain differences in volume computation among surveys and to provide recomputed data.

Evolution of Report Topics

Southern FIA continues to report on traditional timber-related items of area, volume, growth, removals, mortality, TPO, treatments, harvest, and regeneration, among others. Although in some cases, these subjects are reported in lesser detail than in earlier reports. In essence, this trend has become somewhat necessary to capture discussion of a wider range of topics affecting the forest resources. Many of these topics are related to subjects that many perceive as threats to southern forests and are becoming increasingly important to assess. Urbanization, parcelization, ownership change, population shifts, and invasive species are on this list.

Forest health data are now incorporated into southern analytical reports. This incorporation first began with the 2001 South Carolina report (Conner *et al.* 2004), and forest health data are expected to remain a standard subject in analytical reports. The National Woodland Owner Survey findings are also new to southern analytical reports, again beginning with the 2001 South Carolina report. Although covered to some degree through other outlets, another topic recently added to analytical reports is the assessment of invasive/exotic species present in the State's forests. This assessment first appears in the 2002 North Carolina report (Brown *et al.* 2006). A discussion of the wildland-urban interface is new to analytical reporting for southern FIA and was first approached in the 2002 North Carolina report. Information gleaned from these last four topics could help evaluate the degree to which the perceived threats mentioned above are present in a State's forests.

Other new topics have begun to be included as well. An assessment of nontimber products from a State's forests was incorporated in the 2004 Kentucky report (Turner *et al.* n.d.) and the 2002 North Carolina report (Brown *et al.* 2006). These specialty forest products include edible, medicinal, dietary, floral, and crafts. Some analysts are investigating incorporation of ecoregion data, more detailed snag information (although this information is somewhat addressed in forest health sections), and possible inclusion of the NFS data in their analytical reports. An urban inventory pilot test may lead to future inclusion of this subject as well.

Considerations

The mandates for an analytical report every 5 years or five panels and the change from periodic to annual inventories requiring annual updates is why reporting is in transition. As a result, the need for expedient means of reporting data is paramount. Thus, the use or creation of alternative instruments and other outlets to make inventory results available as soon as possible are being explored. When will these feats be accomplished? Southern FIA's first obstacle was the transition from periodic to annual inventories. The differences in plot design, inventory methods and procedures, sample intensity, and data processing that often led to distorted data comparisons between inventories gradually have been addressed. Analysts' time spent struggling with these issues is on the decline as processing systems are edited. Analysts should eventually have more time to devote to the analysis and reporting process.

Other issues exist, however, that give pause to the reporting process. Consensus on reporting formats and types is still developing. Differences within and between State forests in the South can be significant enough to require varied content. Physiography, species diversity and range, and even climate can differ. For example, Florida contains tropical species to the south and temperate ones in the north. Texas has pine and cypress in the east and mesquite in the west. Another anomaly is that many States are not consistent in the collection rate of their panel data, and many are collecting at rates differing from neighboring States. Situations like these could complicate reporting at the State level and affect reporting on regional items. Reports could be every 5 years or every five panels or otherwise.

An important issue in the South is the changing ownership of the forests. Forest industry ownership has been declining and other private entities, such as timber investment management organizations and real estate investment trusts, are increasing. Tracking these changes is difficult on its own, and particularly so with the ceasing of enumerating ownership data. Another

issue in the South is the parcelization of the forests to smaller tracts, which complicates true determination of timberland status as well as ownership. High rates of urbanization in certain States are an issue complicating accurate assessment of diversions from forest land. Finally, the need to collaborate with State partners in the reporting process can lead to lengthier or even multiple reviews, which extends the timeframe of completing a publication.

Future

Under currently mandated timeframes for reporting the data, FIA analysts are properly investigating all avenues for streamlining the reporting process and getting findings into the public forum as expediently as possible. The analytical reports remain the most challenging to expedite and format consensus remains unresolved. Analysts continue to be faced with emerging and lingering considerations when analyzing the data. A growing list of forest-resource-related topics requiring expert address has created the need for multiple authors, and the association with State partners in conducting the inventory has established multiagency authorship. Under the limitations of present organizational staffing, southern FIA envisions smaller and shorter publications along with the use of more graphics and map products. Perhaps that is when the reporting challenges facing FIA will be met.

Literature Cited

Bentley, J.W.; Lowe, L. 2006. Kentucky's timber industry—an assessment of timber product output and use, 2003. Resour. Bull. SRS-105. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 50 p.

Brown, M.J.; New, B.D.; Oswalt, S.N. *et al.* 2006. North Carolina's forests, 2002. Resour. Bull. SRS-113. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 63 p.

Conner, R.C.; Adams, T.; Butler, B. *et al.* 2004. The state of South Carolina's forests, 2001. Resour. Bull. SRS-96. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 67 p.

Conner, R.C.; Sheffield, R.M. 2001. South Carolina's forest resources—2000 update. Resour. Bull. SRS-65. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 31 p.

Glass, P.A.; Oswalt, S.N. [N.d.]. Initial estimates of Hurricane Katrina impacts on Mississippi Gulf Coast forest resources. Manuscript in preparation. On file with: P.A. Glass, Mississippi Forestry Commission, P.O. Box 1667, Jackson, MS 39215–1667.

Oswalt, S.N. 2005. Forest resources of South Carolina's national forests, 2001. Resour. Bull. SRS-98. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 35 p.

Oswalt, S.N.; Brandeis, T.J.; Steadman, D.W.; Robinson, S.K. [In press]. Wintering neotropical migratory songbirds and resident landbirds on St. John, U.S. Virgin Islands. In: Proceedings, 11th Caribbean urban and community forestry conference. St. Croix, U.S. Virgin Islands: [Publisher unknown].

Southern Forest Research Partnership, Inc. 2006. Forest Resources—the heartbeat of the South. Bloomington, IL: Donan Communications Consulting. 25 p. http://www.sfrponline.org/forestresources.pdf.

Turner, J.; Oswalt, C.; Conner, R. *et al*. [N.d.]. Kentucky's forests, 2004. Manuscript in preparation. On file with: J. Turner, U.S. Department of Agriculture, Forest Service, Southern Research Station, 4700 Old Kingston Pike, Knoxville, TN 37919.

U.S. Department of Agriculture (USDA) Forest Service. 2006a. Forest Inventory and Analysis factsheets. Knoxville, TN: U.S. Department of Agriculture, Forest Service, Southern Research Station, Forest Inventory and Analysis. http://srsfia2.fs.fed.us/states/state_information.shtml. (10 September 2006).

USDA Forest Service. 2006b. Compass—perspectives and tools to benefit southern forest resources. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. www.srs.fs.usda.gov. (10 September 2006).

USDA Forest Service. 2006c. Congressional corner. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. http://www.srs.fs.fed.us/cc/#nogo. (10 September 2006).

USDA Forest Service. 2006d. Introduction to mapmaker [Database]. http://srsfia2.fs.fed.us/data/index.shtml. (10 September 2006).

USDA Forest Service. 2006e. The inventory [FIA newsletter]. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. www.srs.fs.usda.gov. (10 September 2006).

Forest Resources of the United States, 2002: Mapping the Renewable Resource Planning Act Data

Cassandra M. Kurtz[1], Daniel J. Kaisershot[2], Dale D. Gormanson[3], and Jeffrey S. Wazenegger[4]

Abstract.—Forest Inventory and Analysis (FIA), a national program of the Forest Service, U.S. Department of Agriculture conducts and maintains comprehensive inventories of the forest resources in the United States. The Forest and Rangeland Renewable Resources Planning Act (RPA) of 1974 mandates a comprehensive assessment of past trends, current status, and the future potential of forest resources across the United States. Data for the RPA represent the highest level of aggregation of timberland and forest land statistics generated by the FIA program. This article describes several forest attributes that make up the FIA 2002 RPA national map atlas series and presents several of the maps. The maps are a crucial tool for visualizing, analyzing, and presenting forest resource information essential to national and regional forest planning and policymaking.

Introduction

In response to the mandate in the Forest and Rangeland Renewable Resources Planning Act of 1974, P.L. 93-378, 88 Stat. 475, the 2002 Renewable Resources Planning Act Assessment (2002 RPA) was prepared. The Forest Inventory and Analysis (FIA) program currently provides updates of forest assessment data every 5 years as required by the Agriculture Research, Extension and Education Reform Act of 1998 (also known as the 1998 Farm Bill). Since the passage of the McSweeney-McNary Forest Research Act of 1928, FIA's emphasis has evolved from solely an economic focus on timber supply and demand to a concern about national resource conditions, ecosystem health, sustainability, and the international resource situation. RPA data are useful to gauge the status of forest resources and make informed strategic-level policy decisions. Currently, RPA is the only format available for exploring long-term trends in the forest resources of the United States across all forest ownerships. Data for the assessment represent the highest level of aggregation of timberland and forest land statistics generated by the FIA regional units.

In support of the 2002 RPA report, national atlas choropleth maps were produced. The maps display a variety of forest characteristics across the United States, including Alaska and Hawaii, and present FIA data in a more dynamic fashion than printed reports using only supporting tables. The project objective was to compile and present a series of thematic maps that evoke the overall county-level geographical distribution of a given forest attribute for the United States. More than 100 maps were created covering 17 attributes.

Data and Methods

FIA plot data, aggregated by county, were downloaded from the 2002 RPA plot summary database in the form of a dBase file (http://www.fia.fs.fed.us). Data were joined to the National Atlas of the United States County Boundaries 2001 shapefile (http://nationalatlas.gov/atlasftp.html) using Federal Information Processing Standards (FIPS) county codes. Data were not available for interior Alaska, the Pacific Basin, Puerto Rico, the U.S. Virgin Islands, and a number of counties in the lower 48 States.

In creating the map series it was important for each map to have the same format with components that meld into a coherent,

[1] Natural Resource Specialist, U.S. Department of Agriculture (USDA) Forest Service, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108. E-mail: cmkurtz@fs.fed.us.

[2] Natural Resource Specialist, USDA Forest Service, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108.

[3] Supervisory Forester, USDA Forest Service, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108.

[4] Forester, USDA Forest Service, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108.

consistent graphic design (Robinson *et al.* 1995). Thus, to standardize map layout and automate production, a map template was designed. The template incorporated normalizing data factors, symbolism, projection, class numbers and limits, color use, typographical relationships, general line weights, and lettering sizes.

Mapping procedures used PC–based Environmental Systems Research Institute ArcMap™ version 9.1 ArcGIS software.[5] Classes were based on natural groupings inherent in the data. To reveal underlying groupings and spatial patterns in the data, categories were first classified using natural breaks. Natural breaks identify breakpoints between classes using a Geographic Information System statistical (Jenks 1963) optimization method. The optimization method minimizes within-class variance and maximizes between-class variance in an iterative series of automated calculations. The breakpoints optimally group similar class values and maximize the differences between classes. Class breaks were manually rounded to ease interpretation. For a complete listing of data sources and estimation procedures (e.g., biomass and mortality) for the 2002 RPA database, see Smith *et al.* (2004) and Miles *et al.* (2004).

The suite of RPA maps include the following map attributes:

- Area of forest land (acres).
- Area of timberland (acres).
- Number of live trees 1.0 in diameter at breast height (d.b.h.) (4.5 ft) and larger on timberland.
- Number of live trees 5.0 in d.b.h. and larger on timberland.
- Volume of growing stock trees on timberland (ft^3).
- Volume of rough and rotten cull trees on timberland (ft^3).
- All live tree biomass on timberland (dry lbs).
- Proportion of land that is forested.
- Proportion of land that is timberland.
- Number of live trees 1.0 in d.b.h. and larger per acre of timberland.
- Number of live trees 5.0 in d.b.h. and larger per acre of timberland.
- Volume of growing stock trees per acre of timberland (ft^3/acre).
- Volume of rough and rotten cull trees per acre of timberland (ft^3/acre).
- Average annual net growth of growing stock trees on timberland (ft^3/acre/year).
- Average annual mortality of growing stock trees on timberland (ft^3/acre/year).
- All live tree biomass on timberland (dry lbs/acre).
- All live merchantable tree biomass on timberland (dry lbs/acre).

In addition, select attributes display hardwood and softwood categories.

The following definitions are adapted and expanded from documentation for the 2002 RPA assessment (Smith *et al.* 2004). Implicit or explicit in these categories is the concept of a size threshold that defines the lower size limit for a tree to be included in the estimate (Birdsey and Schreuder 1992). *Forest land* is defined as land at least 10-percent stocked by forest trees of any size, or formerly having such tree cover, and not currently developed for nonforest uses, with a minimum area classification of 1 acre. *Timberland* is defined as forest land capable of producing crops of industrial wood (20 ft^3 per acre per year in natural stands) that is not withdrawn from timber utilization by statute or administrative regulation. *Growing stock trees* are live trees of commercial species which meet specified standards of quality and vigor. Trees classified as growing stock are commercial species of good form that are at least 5.0 in d.b.h. Trees not classified as growing stock are classified as cull trees. *Cull volume* is defined as the volume of rotten or missing wood in a live or dead tally tree. *Net annual growth* is the average annual net growth in tree volume between remeasurement periods. The components of net growth include the increment in net volume of live remeasure trees, plus the net volume of trees that reached a minimum size threshold during the remeasurement period, minus the volume of trees that died during the period. *Net annual mortality* is the net volume of sound wood in growing stock trees that died during the remeasurement period. *Biomass* is a measure of the total aboveground, oven-dry weight

[5] Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

for live sample trees 1.0 in in diameter or larger, including all tops and limbs (excluding foliage).

Results

Forest Land

Figure 1 shows the spatial distribution of the nation's forest land in 2002 as a percentage of total census land area for each county. Nationwide, an average of 37 percent (749 million acres) of all land is forest land. The accompanying graph (fig. 2) displays the number of counties in relation to the proportion of land that is forested.

Figure 1.—*Proportion of forest land across the United States, 2002.*

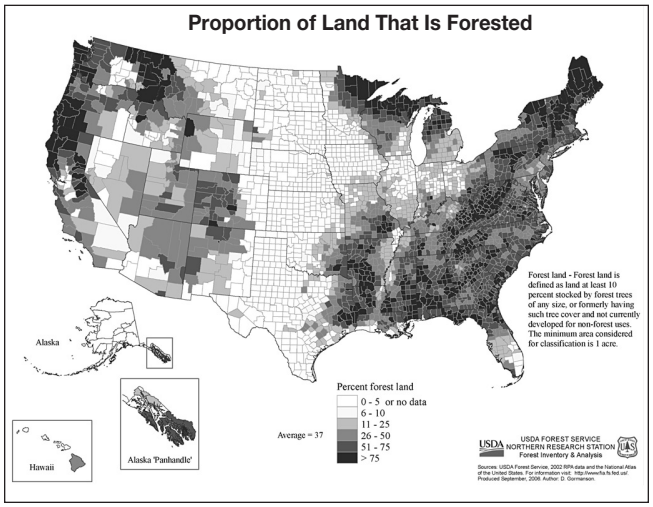
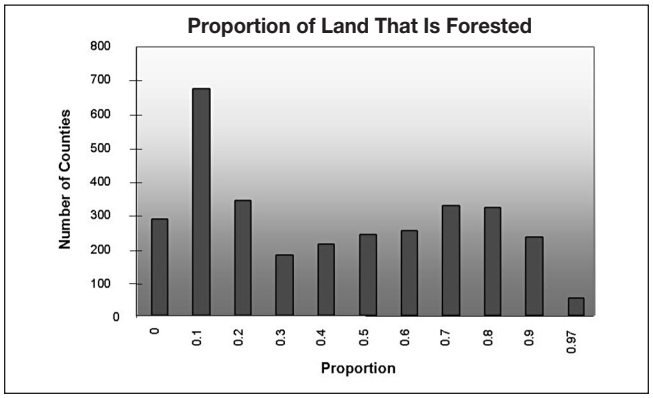


Figure 2.—*Distribution graph of forest land across the United States, 2002.*



Annual Growth

Figures 3 and 4 show side-by-side comparisons of annual growth to mortality, implying that counties with negative growth are likely to have increased annual mortality. Potential causes of negative growth are biological (e.g., insect infestations) or natural disasters (e.g., fire). Several issues must be understood, however, to prevent spurious conclusions (Luppold and McWilliams 2004), especially when examining forest change characteristics such as net growth and mortality.

Figure 3.—*Average annual net growth of growing stock trees per acre of timberland across the United States, 2002.*

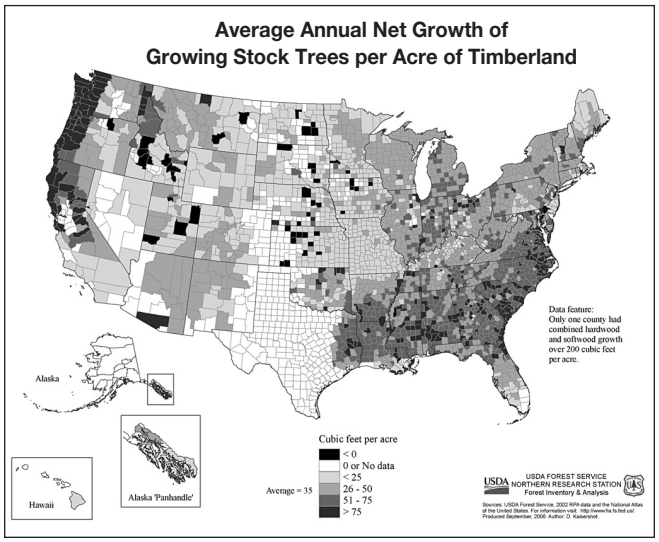
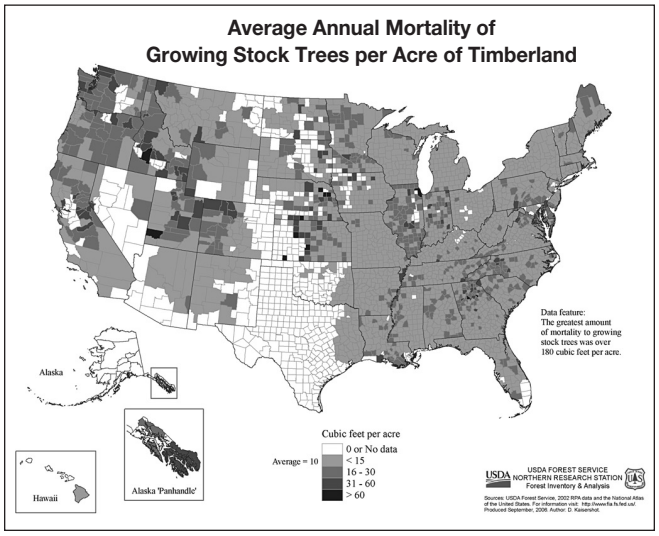


Figure 4.—*Average annual mortality of growing stock trees per acre of timberland across the United States, 2002.*



FIA statistics are developed using sampling procedures and are subject to sampling error; see Bechtold and Patterson (2005) for details of FIA estimation. As the size of the forest area under scrutiny decreases, the number of observations used to develop FIA statistics decreases. In smaller inventory units, the number of observations used to calculate FIA statistics may be particularly inadequate when examining forest components of change; i.e., net growth, removals, and mortality. A small sample size can thus dramatically influence the outcome of the data. This observation may be evidenced by counties in the plains States that have little forest land, particularly Kansas and Nebraska. Overall, 54 counties show negative growth. The smallest negative growth, -.3 ft^3 per acre, is equivalent to the loss of half the volume of a 5-in tree.

Annual Mortality

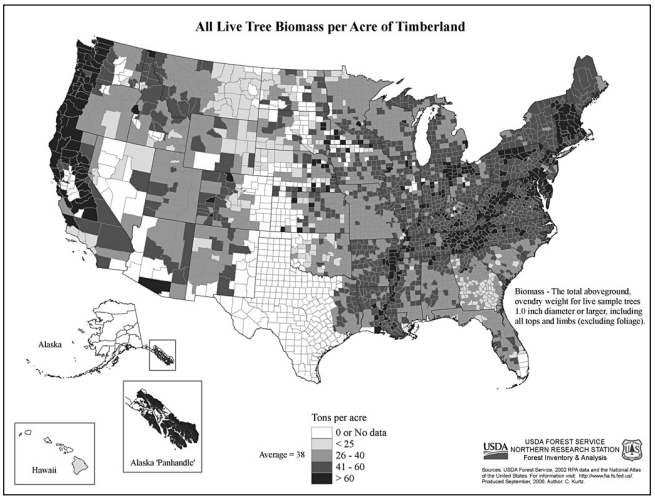
Figure 4 displays average annual mortality of growing stock trees per acre of timberland. The range is from 0 (or no data available at the time of the assessment) to 187 ft^3 per acre with an average of 10 ft^3 per acre. The greatest amount of mortality to growing stock trees occurred in Cuming County, NE. Because Nebraska is sparsely forested, this amount of mortality is likely the result of a small sampling size and further investigation is recommended.

Biomass

RPA inventory data are used in the estimation of live tree biomass (an indicator of carbon) and applied to a number of studies of national and international importance. Biomass data are used for reports on greenhouse gas emissions as documented by the Environmental Protection Agency (EPA 2006) and wildland fuel management as seen in the LANDFIRE Prototype Project (Rollins *et al.* 2006). In addition, live tree biomass has an important role in the global carbon cycle (Brown 2001).

Live tree biomass per acre of timberland (fig. 5) ranges from 0 (or no data) to 210 dry tons per acre with an average of 38 dry tons per acre. The top five counties are all located in California; San Benito County has the greatest live biomass per acre of timberland. In the United States, the highest average biomass per acre is predictably along the Pacific Coast, where larger trees reside.

Figure 5.—*All live tree biomass per acre of timberland across the United States, 2002.*



Past, Present, and Future

The original RPA Plot Summary database format was developed in 1987 to meet RPA reporting requirements (Waddell *et al.* 1989). FIA has developed the only forest resource database of sample plots across the United States. This unique database is continually evolving and is used by a wide range of researchers both within and outside the Forest Service, natural resource agencies, State, regional economic development and special interest groups, industry, and landowner associations.

Forest inventory data are available to create custom tables and maps using FIA's online services at www.fia.fs.fed.us/tools-data/tools. More than 100 maps created with the 2002 RPA data are available at http://www.fia.fs.fed.us/program-features/rpa/. Previous RPA reports can be found at http://svinet2.fs.fed.us/pl/rpa/list.htm. Future plans include mapping the 2007 RPA data and providing enhanced Web services.

Acknowledgments

The authors thank Pat Miles and John Vissage for their assistance with data acquisition and analysis and thank John Vissage, Chris Woodall, and Leah Rathbun for reviewing this article.

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.

Birdsey, R.A.; Schreuder, H.T. 1992. An overview of forest inventory and analysis estimation procedures in the eastern United States. Ft. Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station. 11 p.

Brown, S. 2001. Measuring carbon in forests: current status and future challenges. Environmental Pollution. 116(3): 363-372.

Jenks, G.F. 1963. Generalization in statistical mapping. Annals of the Association of American Geographers. 53: 15-26.

Luppold, W.G.; McWilliams, W.H. 2004. Avoiding spurious conclusions from Forest Service estimates of timber volume, growth, removal, and mortality. Northern Journal of Applied Forestry. 21(4): 194-199.

Miles, P.D.; Vissage, J.S.; Smith, W.B. 2004. The 2002 RPA plot summary database user's manual. Gen. Tech. Rep. NC-251. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station. 80 p.

Robinson, A.H.; Morrison, J.L.; Muehrcke, P.C.; Kimerling, A.J.; Guptill, S.C. 1995. Elements of cartography. 6th ed. New York: John Wiley. 674 p.

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, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 416 p.

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

U.S. Environmental Protection Agency (EPA). 2006. Inventory of U.S. greenhouse gas emissions and sinks: 1990–2004. EPA 430-R-06-002. http://www.epa.gov/climatechange/emissions/downloads06/06_Complete_Report.pdf. (21 December 2006).

Waddell, K.L.; Oswald, D.D.; Powell, D.S. 1989. Forest statistics of the United States, 1987. Res. Bull. PNW-168. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 106 p.

Additional Reading

U.S. Department of Agriculture (USDA) Forest Service. 1980. An assessment of the forest and range land situation in the United States. Forest Res. Rep No. 345. Washington, DC: U.S. Department of Agriculture, Forest Service. 641 p.

USDA Forest Service. 1989. RPA assessment of the forest and rangeland situation in the United States, 1989. Forest Res. Rep. No. 26. Washington, DC: U.S. Department of Agriculture, Forest Service. 72 p.

USDA Forest Service. 1994. RPA assessment of the forest and rangeland situation in the United States, 1993 update. Forest Res. Rep. No. 27. Washington, DC: U.S. Department of Agriculture, Forest Service. 73 p.

USDA Forest Service. 2001. 2000 RPA assessment of forest and range lands. FS-687. Washington, DC: U.S. Department of Agriculture, Forest Service. 78 p.

U.S. Department of the Interior. 2006. National atlas of the United States, county boundaries of the United States 2001. http://nationalatlas.gov. (5 March 2006).

The Influence of International and Domestic Events in the Evolution of Forest Inventory and Reporting Consistency in the United States

W. Brad Smith[1]

Abstract.—This article takes a brief chronological look at resource inventory and reporting and links to international influences. It explores events as drivers of more consistent data within the United States and highlights key dates and events in the evolution of inventory policy and practice. From King George to L'Ecole nationale forestiere to the Food and Agriculture Organization of the United Nations to the Montreal Process criteria and indicators, Forest Service, U.S. Department of Agriculture inventories have been shaped by events beyond our shores and national assessments have been a focal point for cross-regional data consistency and program changes. Each event has brought us closer to a more consistent and compatible inventory to share with the Nation and world.

Introduction

You may be surprised to discover the role that global events have played in the historic development of forest inventory in the United States and, more specifically, the Forest Inventory and Analysis (FIA) program in the United States in terms of standardizing and reporting data. It is a story that begins nearly 300 years ago, before the American Revolution. You will see a convergence of events that you may wish to call mere coincidence, suggesting that forest inventory has evolved in a vacuum devoid of outside influence; in a more likely scenario, it has influenced and has been influenced by global events at key points in history.

The quest for knowledge of the quantity, quality, and location of natural resources and access to them permeates world history. At the core, the key measures to delineate forests include area, ownership, density, size, age, quality and volume. Time-series measures of these key items provide data on growth, mortality, removals, and land use change. These trends, in turn, have provided measures of potential growth and yield to enable predictive forest management planning. Although the key measures have remained relatively static over time, a tremendous amount of energy has been expended on improving their consistency and accuracy across larger and larger landscapes.

This article is not intended as an exhaustive review of how global and domestic events have affected forestry in the United States; instead, it contains a series of chronological vignettes and observations that highlight events that have played a role in shaping strategies for monitoring forests and, in turn, how we report and analyze their condition and trends.

1729–1850
The Early Demand for Data

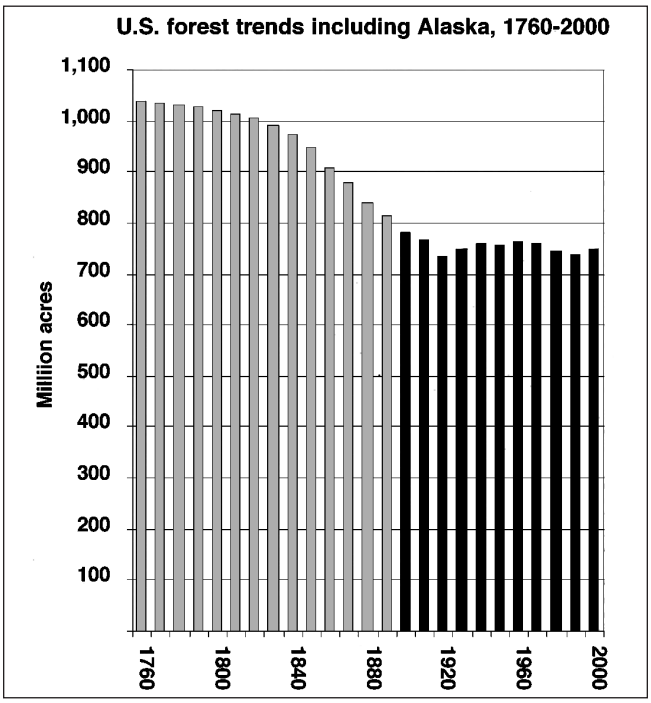
Much of the early story revolves around the exploitation of U.S. forests for products, beginning with the early European settlers. Before the American Revolutionary War, conflicts involving England, France, and Spain had a profound impact on U.S. forests. Particularly notable were conflicts with England, who had cut the larger part of her forests to build a navy over the centuries. During European settlement in America in the mid-18th century, people appointed by the British Royal Navy as Surveyors General of His Majesty's Woods were commissioned to survey forests to locate suitable trees for ship masts (Albion 1926). The core variables of these surveys were species (white

[1] Associate National Program Manager, U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis program, Washington DC 22209. E-mail: bsmith12@fs.fed.us.

pine), diameter (more than 24 inches diameter at breast height), suitable form and height, quantity, and location (not too far from a navigable water). The report of the survey, based on the British king's "Broad Arrow" policy, went to the headquarters of the Royal Navy; it also went to the local American press as fodder for the coming revolution.

Volume was added to the list of core variables shortly after the fledgling United States began to set aside thousands of acres of live oak and pine for ship timbers and naval stores (Albion 1926). It would be more than 100 years later before the U.S. Census Office (now the U.S Census Bureau) would begin to envisage how to amass some form of data for the entire Nation and another 50 years before it would have the statistical means to begin a true field inventory of the Nation's forests. For the time being, the landscape was too large and the manpower and technology too sparse to operate on anything but an "as needed" basis for inventory.

Figure 1.—*Forest area in what is now the United States, from 1760 through 2000.*



1850–1900
First National Estimate of Forest Cover

By 1850, forest clearing in the United States had reached 90 million acres (fig. 1) and moved into high gear as immigration surged into the vast midcontinent forests, which were cleared for farmland. Two hundred million more acres of forest fell to the axe by the end of the century, a pace of nearly 13 square mi a day for 50 years (MacCleery 2002). The earliest attempts at large-scale inventory of forests occurred in Massachusetts in 1830, but they occurred in only one State and were very limited in scope. Other anecdotal attempts were made, but nothing covering any significant area was accomplished.

In the midst of this massive forest clearing, concerned conservationists like Franklin Hough, Charles Sargent, Bernard Fernow, and even Henry David Thoreau began to sound alarms.[2] Meanwhile, the chaos of the Civil War scarcely slowed down the forest clearing rate. The first serious attempt at estimating forest resources nationwide came out of the shadows of the Civil War in the 1870 Decennial Census, as described by William Brewer (1873): "The ratio of woodland to other land was calculated for each county, and made up the first basis for the (forest) map. … No published map of any considerable area in our country is known to us, on which the woodlands are laid down from actual surveys." (Brewer 1873) Only general information was provided about the geographic reference of species and type distributions, but it was a beginning. This early work of the Census Office, although crude, incited the imagination of those wishing to extract even more information about the Nation's forests in a consistent way. The core variables for this effort were estimates of "wooded land" on census forms sent to private individuals and estimates from public land managers, primarily in the West. These estimates and anecdotal information from surveyors' notes on various species sizes and distributions were used to develop a density map of forests (fig. 2). It would be another 50 years before organized forest surveys would begin in earnest, but this early effort helped proponents of a national forest inventory visualize where and how much forest was present in the United States.

[2] An address on the Succession of Forest Trees read to the Middlesex Agricultural Society in Concord, MA, in September 1860.

In 1882, Franklin Hough's *Report Upon Forestry* set forth basic principles that encouraged the gathering of knowledge from all sources, domestic and foreign, to address the forestry situation in America:

> "The collection, determination, and diffusion of facts having practical application to forest culture,…and the promotion of researches tending to enlarge the boundaries of our knowledge in sciences that concern this subject. Under this head we would include careful examination of the results of experiments and observations as they are published from time to time in Europe, with the view of availing ourselves of so much of these as may appear applicable within our own country." (Hough 1882: 3)

Hough (1882) also suggested that future surveyors keep notes on the character of the topography and vegetation near the survey lines, noting, "Had this information been provided for in former surveys, the government would have had at hand information that would have been of inestimable value upon points where it can now only be obtained by agents sent upon particular occasions, and at considerable expense." (Hough 1882: 21) Although Hough clearly saw the value of integrating a "value-added" survey in ongoing census work, the idea did not take hold just yet and 1880 would be the last decennial census to have a focal area on forests (Sargent 1884). A formal survey would have to wait.

In 1876, at the Philadelphia Centennial Exposition, the chief forester of the Prussian Forest Service, Richard von Steuben, conferred with Secretary of the Interior Carl Schurz about forest

Figure 2.—*The density and distribution of woodlands in the United States (compiled in 1873).*



conservation and the need to institute governmental forestry practice in the United States (Rodgers 1951). Schurz addressed the American Forestry Congress in Cincinnati that year about the need for a Federal forestry program.

In 1889, Gifford Pinchot observed, "Without being himself a forester, my Father understood the relationship between forests and national welfare…. He was sure that forestry must come to America…." (Pinchot 1947: 2) With little support, even from Charles Sargent or Bernard Fernow, for the forestry profession in America, Gifford went to France to attend L'Ecole nationale forestiere, where he learned the basics of forestry and forest management. Sir Dietrich Brandis, Pinchot's most cherished mentor, noted, "Nothing general can be done until some State or large individual owner makes the experiment and proves for America what is so well established in Europe, that forest management will pay." (Pinchot 1947: 15)

Located in Asheville, NC, the Biltmore Estate was the property of the wealthy Vanderbilt family and spread across 7,000 acres of mountains and valleys, half of which was wooded. The estate provided Pinchot the opportunity to show off what he had learned in Europe, and American forestry was born—of clearly international lineage. Although Pinchot's tenure at Biltmore was an entrée to forestry in America, once again, a survey of the Nation's forests on anything but this anecdotal scale would have to wait.

When Congress passed the Organic Administration Act of 1897, the mission of forestry in the United States finally had solid legal footing. This act provided inventory guidance before harvesting, as detailed in Section 23 (core items are underlined for emphasis):

> "While sales of timber may be directed by this Department without previous request from private individuals, petitions from responsible persons for the sale of timber in particular localities will be considered. Such petitions must describe the land upon which the timber stands by legal subdivisions, if surveyed; if unsurveyed, as definitely as possible by natural land marks; the character of the country, whether rough, steep or mountainous, agricultural or mineral, or valuable chiefly for its forest growth; and state whether or not the removal of the timber would result injuriously to the objects of forest reservation. If any of the timber is dead, estimate the quantity in feet, board measure, with the value, and state whether killed by fire or other cause. Of the live timber, state the different kinds and estimate the quantity of each kind in trees per acre. Estimate the average diameter of each kind of timber, and estimate the number of trees of each kind per acre above the average diameter." (Organic Act 1897, Section 23).

Although targeted to timber for harvest, the act required an inventory of the location, type (species), quantity (volume), diameter, and number of trees.

1900–1950
Momentum Finally Produces Action

As the 19th century turned, the United States still had no comprehensive picture of its forests. In 1909, however, R.S. Kellogg patched together all the best estimates available into a national view, but in the end he lamented, "The estimates of the original and present forest areas and stands are at best only approximate. They are offered tentatively, and any information which will make them more accurate will be gladly received. Great as is the need for it, there has never been a timber census of the United States, nor, with one or two exceptions, any close estimate of the forest resources of any individual State. Such a census must eventually be taken to furnish the basis for permanent forest conservation." (Kellogg 1909: 3)

In the heady Teddy Roosevelt era at the beginning of the 20th century, America pushed its way onto the global political stage and a 35-year-old Forest Service researcher named Raphael Zon from the Lake States Forest Experiment Station began working with William Sparhawk, a Forest Service economist, and global colleagues on an assessment of the forests of the world. In the foreword of *Forest Resources of the World*, Pinchot noted, "International organizations, such as the International Institute of Agriculture in Rome, are maintained for the collection of world agriculture statistics. Fairly accurate statistics of the world's mineral resources are also available. Forest resources, however, although basic for industrial

development of nations, are known least of all. There is no international organization for the systematic collection of forest statistics." (Zon and Sparhawk 1923: vii) "What gives particular interest to the book is the fact that the forests are treated not merely as available exploitable materials, but as living, renewable resource not to be destroyed by use, but to be regrown, perpetuated, and improved." (Zon and Sparhawk 1923: viii) Pinchot also noted that these volumes would likely serve as a standard reference source for many years to come. This early work, still relying mostly on expert opinion rather than on field surveys, and the work of the International Institute that Pinchot mentioned became the forerunners to the post-World War II work of the Food and Agriculture Organization (FAO) of the United Nations, in which the Forest Service played a prominent role. Pinchot approached every president after Woodrow Wilson to push for a world conservation conference; his persistence eventually paid off, as we shall see later.

By the mid 1920s, Scandinavian countries had started inventorying national forest resources using statistical methods. In 1925, a Canadian volume entitled *Statistical Methods in Forest Investigative Work* clarified concepts of statistical error and how to correct for it (Wright 1925). In 1928, Dr. Yrgo Ilvessalo of the Finnish national forest inventory met with President Calvin Coolidge and discussed these inventories and their methodology (Van Hooser *et al.* 1993). Shortly thereafter, the McSweeney-McNary Act of 1928 was passed, which directed the Secretary of Agriculture "to make and keep current a comprehensive survey of the present and prospective requirements for timber and other forest products in the United States, and of timber supplies, including a determination of the present and potential productivity of forest land therein, and of such other facts as may be necessary in the determination of ways and means to balance the timber budget of the United States." (RPA 1978). This one-sentence directive was the legal mandate for the Nation's forest survey, which today we call FIA. FIA was only the fourth national-scale forest inventory in the world, following forest inventory initiatives in Germany, Sweden, and Finland. In the early 1930s, FIA began fieldwork in the South, Lake States, and Pacific Northwest. President Franklin Roosevelt's New Deal, which initiated and supported conservation work throughout the country, supplied much of the manpower for this fieldwork.

World Conservation Union—A Sidebar to History

Starting in 1909, Gifford Pinchot attempted, through the administrations of U.S. presidents, to convene a world conservation conference. Finally, after 30 years of trying to persuade elected officials, Pinchot convinced Franklin Roosevelt of the need for such a conference. Pinchot summed up his philosophy by saying, "International co-operation to inventory, conserve, and wisely utilize natural resources to the mutual advantage of all nations might well remove one of the most dangerous obstacles to all nations to a just and permanent world peace…." (Pinchot 1947: 369)

Roosevelt favored the world conservation conference idea but died before it came to pass. Pinchot, ever persistent, put the idea before President Harry Truman before he himself died in 1947 without knowing its fate. President Truman, however, took up the cause and the conference was put on the agenda of the United Nations Educational, Scientific and Cultural Organization (UNESCO) in 1948.

Ken McDonald of the University of Toronto provides a further tantalizing insight into Pinchot's long-sought-after conservation conference, which was known as the International Union for the Conservation of Nature and Natural Resources (IUCN), or today's World Conservation Union, noting,

> "To understand the contemporary form of IUCN and to recognize the degree of continuity in institutional goals and objective it's important to engage in this kind of analysis and to remember the institutional origins of IUCN. Rather than emerging from the ether of international concern in 1948, IUCN has its genealogical roots in earlier organizations that presented themselves as devoted to the cause of nature preservation. In particular, many individuals involved in the establishment of IUPN, as IUCN was known until 1956, were leading figures in the Society for the Preservation of the Wild Fauna of Empire, established in 1903 (and which continued on to become the Fauna Preservation Society)." (McDonald 2003)

UNESCO and the embryonic International Union for the Protection of Nature (IUPN) came into being at approximately the same time (1946). The latter, however, was not formally constituted until 1948, by which time UNESCO was already planning a technical conference on the conservation of nature to be held in 1949 in New York. IUPN was brought under the wing of UNESCO and charged with preparing the agenda for the conference.

The International Institute of Agriculture in Rome published forest estimates supplied by national governments in a series of yearbooks that appeared from 1933 to 1938. After the outbreak of World War II temporarily suspended forest surveys, available information on forests was summarized in 1946, when a report was submitted to the second session of the FAO Conference under the title *Forestry and Forest Products—World Situation, 1937–1946.* These investigations made valuable additions to knowledge of forests but suffered from a lack of reliable information, including common definitions of important terms. We still struggle with these issues today. In 1947, FAO subsumed the work of international forest statistics under its program of work. By the late 1940s, forest surveys resumed in the United States.

1950–1980
First National Assessment and Standards

It is interesting to note that FAO was initially housed in Washington, DC, from 1946 through October 1951, when it was permanently moved to its present location in Rome. In November 1951, Lyle Watts, then Chief of the Forest Service, chaired an FAO forestry panel at the sixth session of the FAO Conference in Rome that made recommendations to promote national forest inventories worldwide. The conference empowered the FAO Secretariat to collect and publish at 5-year intervals available information on all the forest resources of the world. Watts became a driving force in initiating *Timber Resources for America's Future,* first national-scale forest assessment (USDA Forest Service 1958). Forest Service Chief Richard McArdle oversaw the completion and publication of this comprehensive report in 1958. A young forester with a prominent role in that effort was John McGuire, who later became Chief of the Forest Service in 1974, when the Resources Planning Act was passed; the act mandated periodic national assessments of America's forests.

Following the 1951 FAO Conference in Rome, FIA completed 42 State inventories in 10 years —nearly double the pace of the inventories that had occurred during the previous 20 years. The national assessment that was compiled to provide leadership in national reporting was a big success. McArdle noted, "We hope that this study will add to America's leadership in forestry, that it will be useful to other nations of the world in relating their timber situation to ours, and that it will serve as the basis for long-range forestry planning for progressive forest landowners and for State and Federal Governments." (USDA Forest Service 1958: iii) McArdle gave special praise to State forestry agencies and forest industries for their assistance and input, which were critical in preparing the report.

The 1950s was the most active decade in beginning the national survey. Field locations were established in States where forest survey had been inactive before World War II. As the decade closed, the first cycle of the forest survey was nearly complete.

In 1954, the United Nations Economic Commission for Europe (UNECE) Joint Committee was founded as a forum for the exchange of information and experience among Europe and North America countries about forest working techniques and the training of forestry workers. The Timber Section of the UNECE Trade and Timber Division, the Industrial Activities Branch of the International Labour Organization, and the Forest Harvesting and Transport Branch of the FAO Forestry Department's Forest Products and Industries Division produced annual reports of timber production. Providing data for these reports was the underpinning of FIA's Timber Products Output data and reports.

Back in 1928, the initial mandate for forest survey had been very specific: timber was the main concern—timber requirements, timber supply, and balancing the two. Early surveys tracked these elements but in different ways across the country. The difficulty and frustration of pulling together the disparate data from surveys designed and administered regionally for the 1953 assessment would set into motion the development of FIA's first national Forest Survey Handbook (USDA Forest Service 1967). The last major revision of the handbook was signed in 1975 by John McGuire, who had been responsible for the accuracy of the first national assessment. The handbook contained 87 definitions, dozens of variables, and 26 tables that later became core requirements for all State reports. The handbook was used as an FIA guide until Congress's approval of the 1999 FIA Strategic Plan as a provision of the 1998 Farm Bill. Even today, the contents of 22 of the 26 core tables required by the 1967 handbook are present in each State report.

During the 1970s, growing awareness of the complex interactions among the many forest uses and recognition of acute problems in the budgeting process led Congress to pass the Forest and Rangeland Renewable Resources Planning Act (RPA) of 1974 and the Forest and Rangeland Renewable Resources Research Act of 1978. RPA has been described as a bold new experiment in resolving resource issues. It directed the Secretary of the U.S. Department of Agriculture to prepare an assessment of the Nation's forest and rangeland resources every decade and develop a long-range program to guide the orderly development of the natural resources on the national forests. Most importantly for FIA, both laws expanded the inventory mandate to all renewable resources on the Nation's forests and rangelands. The Honorable Hubert H. Humphrey, former Vice President and Senator of the United States, summed up the issues in remarks during a debate of the National Forest Management Act of 1976: "The days have ended when the forest may be viewed only as trees and trees only as timber." (USDA 1983)

1980–1990
Blue Ribbon Panels and Sustainability Take Center Stage

During the 1980s, the FIA program continued to develop a strong regional flavor and drifted a bit from the standards laid down in the 1967 Forest Survey Handbook as new demands were placed on forest inventories to deliver more than just timber data. Although the regions addressed many of the same issues that the national program did, they did it separately. Forest industry organizations with national interests found the shift toward regional programs frustrating and began to lobby for more consistency. Budgets languished during the 1980s, which spurred even more creative local solutions to address issues.

In the early 1990s, stakeholders who needed more timely information on the Nation's resources thought that FIA's capacity to report data within a given timeframe was falling behind. These stakeholders convened a series of Blue Ribbon Panels (AFC 1992; AFPA 1998, 2001) to focus on the issue. Intense and widespread interest accompanied the convening of the First Blue Ribbon Panel on Forest Inventory and Analysis in 1992. The second Blue Ribbon Panel was convened in 1997 to expand on the progress of the First Blue Ribbon Panel's recommendations. A followup to the second panel was convened in 2001. Each Blue Ribbon Panel included high-level leaders from the entire forestry community, including Federal and State agencies, industry organizations, environmental organizations, academia, and other user groups. Their mission was to develop a national vision and strategy, as well as goals and objectives, for meeting the present and future needs for forest inventory information. Around the time of the first panel, the FIA program developed a strategic outlook report (USDA Forest Service 1993).

Key initiatives from the First Blue Ribbon Panel included the following:

- Improving and expanding information on ecosystems and noncommodity values.
- Recognizing and identifying ownership, regulatory, and social impacts on forest productivity.
- Producing the most current resource data possible.
- Implementing a uniform approach for all ownerships.
- Increasing consistency and compatibility among FIA units.
- Enhancing coordination between FIA and public agencies.
- Improving service to user groups.
- Expanding clientele.

Key initiatives from the Second Blue Ribbon Panel included the following:

- Initiating an annual inventory and supporting analysis.
- Fulfilling the mandate of reporting on all forest lands.
- Concentrating on core ecological and timber data.
- Developing a strategic plan.

After 10 years of effort, three Forest Service Chiefs had been receptive, but resources were not moving to sufficiently support the Panel's recommendations. At this point, a convergence of events sparked a major change. First, in a letter to the Council on Environmental Quality in March 1998, the National Association of State Foresters (NASF) took an early stand on the value of the international processes stemming from the United

Nations Conference on Environment and Development in Rio de Janeiro, Brazil:

> "As you know, the C&I were developed and agreed to by the United States and 12 other countries making up what is known as the Montreal Process, a group of nations representing 90 percent of the world's temperate and boreal forests. Within the C&I framework, the Criteria represent broad level values or general categories of attributes associated with forests. …There was also fairly strong consensus among the stakeholders that the C&I are a good framework to guide future forest resource assessments and inventories. …The White House should direct the other Federal agencies with natural resources inventory and monitoring responsibilities to work with the Forest Service so that all data gathered is compatible, timely, and informs the Montreal Process C&I. … Adequate funding must be made available within the Forest Service budget to support the Forest Inventory and Analysis Program. The Forest Health Monitoring Program (which cover inventory on a substantial portion of the nation's private forests), and the expansion of the appropriate inventory, monitoring, and assessment protocols to all Federal public lands." (Imbergamo 1998)

A similar letter was sent to the Chief of the Forest Service. At the same time, constituents banded together to draft legislation that would legally define the mandate of FIA to provide support for these efforts. The outcome was the Farm Bill of 1998, which redefined the scope and process for how FIA conducted inventories. The key elements of the Farm Bill that amended the Forest and Rangeland Renewable Resources Research Act 1978 include the following:

Sec 3(e) Forest Inventory and Analysis.—

(1) Program required.—In compliance with other applicable provisions of law, the Secretary shall establish a program to inventory and analyze, in a timely manner, public and private forests and their resources in the United States.

(3) 5-year reports.—Not more often than every 5 full fiscal years after the date of enactment of this subsection, the Secretary shall prepare, publish, and make available to the public a report, prepared in cooperation with State foresters, that—

  (a) contains a description of each State inventory of forests and their resources, incorporating all sample plot measurements conducted during the 5 years covered by the report;

  (b) displays and analyzes on a nationwide basis the results of the annual reports required by paragraph (2); and

  (c) contains an analysis of forest health conditions and trends over the previous 2 decades, with an emphasis on such conditions and trends during the period subsequent to the immediately preceding report under this paragraph.

(4) National standards and definitions.—To ensure uniform and consistent data collection for all forest land that is publicly or privately owned and for each State, the Secretary shall develop, in consultation with State foresters and Federal land management agencies not under the jurisdiction of the Secretary, and publish national standards and definitions to be applied in inventorying and analyzing forests and their resources under this subsection. The standards shall include a core set of variables to be measured on all sample plots under paragraph (2) and a standard set of tables to be included in the reports under paragraph (3).

(6) Strategic plan.—Not later than 180 days after the date of enactment of this subsection, the Secretary shall prepare and submit to Congress a strategic plan to implement and carry out this subsection, including the annual updates required by paragraph (2) and the reports required by paragraph (3).

Long Lead up to the Simple Point

A common thread in all the preceding dissertation is the repeated call for uniform and consistent data. As local constituents drove the inventory regionally from the bottom up, the demand for consistent global and national information about techniques and resource data drove change from the top down.

Those entrenched in the day-to-day operations at the provincial level rarely saw the value of producing consistent data for the next level up; they saw no client there, and, if the consistency cost money, they saw no compelling rationale to change. This consistency implied a scrutiny that, in turn, implied higher quality data at all levels.

After all is said and done, national and international reporting is a cross-regional quality assurance process. It is through these periodic exercises that we check ourselves for compliance across regional boundaries. Standards that we assume are uniformly applied often are nudged and wiggled for local needs and it is no wonder that, following the Forest Service's first national assessment, a grand push was made to standardize plots, variables, and definitions for regional inventories.

We like to think we have a "check as we go" continuous quality assurance system and perhaps we do regionally, but not nationally. We only check the regional outputs as they come off the end of the line to see if the parts are interchangeable with outputs from other regions. Today's system makes it imperative to check within regions as well. In the old days of the "gypsy crews" that went State to State by region, continuity within regions was a given—the same crews and the same trainers were used. Now, because we operate in nearly all States simultaneously, that is no longer the case and we begin to see anomalies even State to State within regions. Now, instead of tracking 4 inventory systems, we must track 50 for continuity. Because of the energy involved getting on the same page, we alternately enter periods of standardization and then drift. Standardization usually is driven by trying to gather data for national or international assessments.

As the 2010 national assessment takes shape, we are in the throes of several changes that are causing regional angst as we try to improve cross-regional consistency. Because the changes will have little impact at the national scale of the data, the locals say, "Why bother?" In the end, it is accountability based on attention to details that makes programs great. Historically, this process has been strong at the regional level, but the end game has changed. To compete effectively in a global society, we cannot just "sell them what we got" regionally; we have to deliver what they want globally.

Although accuracy is generally a bottom-up issue that is addressed from the design phase, consistency is more often a top-down issue that is only addressed when multiple surveys need to be merged to achieve a common goal. The most significant issue arises when inventory designers do not consider, or are unaware of, consistency issues pertinent to design elements. We continue to struggle with this problem to this day, but progress is being made. The primary driver for that progress is the continued demand for national and global resource reports.

The processes that have done the most to shape national consistency in our inventories have been inspired by people thinking at the national and international scales. No one saw this better than Pinchot, the Forest Service's first Chief, who understood the value of current, consistent data about our resources and how they fit into the global community. The demand for consistency has not changed all that much in the past 100 years. Bertram Husch authored one of the primary texts on forest mensuration that has been used in forestry schools for decades. In a paper derived form a speech given at the World Forestry Congress in Spain in 1965, Husch (1966) noted:

> "Anyone interested in more that the results of one single inventory will be immediately struck by the diversity of inventory methods and even more so by the variety in the form of information obtained. This variability is a carryover from the earlier days of forestry when there was little incentive for any kind of standardization. …There is little incentive and indeed a feeling of wasted effort and funds to achieve comparability or combinability which is seemingly only of external benefit with little internal use to the organization carrying out the inventory. …The result is that few inventories are executed using standards or specifications allowing combining or comparing the results. …I would suggest that the benefits of standardization to make forest inventory information comparable and compliable with other inventory results would be advantageous to all foresters and organizations with dynamic interests. … It should be obvious that the goal of comparability and compatibility is not beneficial only on the broad national or international level but is of vital importance to all forest enterprises…." (Husch 1966)

In today's terms, all ecological enterprises, it is all good. Tear a page from Pinchot's and Husch's books and let persistence be a virtue when it comes to consistent and compatible inventory data.

Literature Cited

Albion, R.G. 1926. Forests and sea power. Cambridge, MA: Harvard University Press. 483 p.

American Forest & Paper Association (AFPA). 1998. Forest Inventory and Analysis program: the report of the second blue ribbon panel. Washington, DC: American Forest & Paper Association. 17 p.

AFPA. 2001. Status report on the U.S. Forest Service Forest Inventory and Analysis program: update of the findings of the 1998 Blue Ribbon Panel. Washington, DC: American Forest & Paper Association. 9 p.

American Forest Council (AFC). 1992. Report of the blue ribbon panel on forest inventory and analysis. Washington, DC: American Forest Council. 14 p.

Brewer, W.H. 1873. The woodlands and forest systems of the United States. U.S. Census Statistical Atlas. New York: J. Bien.

Forest and Rangeland Renewable Resources Research Act (RPA) of 1978; P.L. 95-307, 92 Stat. 353, as amended.

Hough, F.B. 1882. Report upon forestry. Vol. III. Washington, DC: U.S. Government Printing Office. 318 p.

Husch, B. 1966. Standardization in forest resource surveys. In: Proceedings, sixth world forestry congress. Rome, Italy: Food and Agriculture Organization of the United Nations: 3933-3940.

Imbergamo, B. 1998. Written correspondence with the Forest Service and Office of Management and Budget in support of Montreal Process criteria and indicators. On file with: National Association of State Foresters, 444 North Capitol Street, NW, Washington, DC 20001.

Kellogg, R.S. 1909. The timber supply of the United States. Circ. 166. Washington, DC: U.S. Government Printing Office. 24 p.

MacCleery, D.W. 2002. American forests: a history of resiliency and recovery. Durham, NC: Forest History Society. 58 p.

McDonald, K.I. 2003. IUCN: a history of constraint. Text of an address given to the Permanent Workshop of the Centre for Philosophy of Law Higher Institute for Philosophy of the Catholic University of Louvain. Toronto, Ontario, Canada: University of Toronto, Department of Geography. 25 p. http://perso.cpdr.ucl.ac.be/maesschalck/MacDonaldInstitutional_Reflexivity_and_IUCN-17.02.03.pdf.

Organic Administration Act of 1897 (Organic Act); 16 U.S.C. § 478, 551.

Pinchot, G. 1947. Breaking new ground. New York: Harcourt Brace. 1987 reprint. Washington, DC: Island Press. 522 p.

Rodgers, A.D. 1951. Bernhard, Eduard Fernow: a story of North American forestry. Princeton, NJ: Princeton University Press. 623 p.

Sargent, C.S. 1884. Report on the forests of North America (exclusive of Mexico). Vol. 9. Washington, DC: U.S. Government Printing Office. 6 p. http://www.census.gov/prod/www/abs/decennial/1880.htm.

U.S. Department of Agriculture (USDA). 1983. Yearbook of agriculture. Washington, DC: U.S. Government Printing Office. 158 p.

USDA Forest Service. 1958. Timber resources for America's future. FRR No. 14. Washington, DC: U.S. Department of Agriculture, Forest Service. 713 p.

USDA Forest Service. 1967. Forest survey handbook. FSH 4809. Washington, DC: U.S. Department of Agriculture, Forest Service. 253 p.

USDA Forest Service. 1993. A blueprint for forest inventory and analysis research and vision for the future. Program Aid No. 1512. Washington, DC: U.S. Department of Agriculture, Forest Service. 20 p.

Van Hooser, D.D.; Cost, N.D.; Lund, H.G. 1993. The history of the forest survey program in the United States. In: Preto, G.; Koch, B., eds. Proceedings, 1992 International Union of Forest Research Organizations centennial meeting. Tokyo, Japan: Society of Forest Planning Press: 19-27.

Wright, W.G. 1925. Statistical methods in forest-investigative work. Forestry Branch Bull. 77. Ottawa, Canada: Department of Interior.

Zon, R.; Sparhawk, W. 1923. Forest resources of the world. New York: McGraw-Hill. 997 p.

Additional Reading

Hough, F.B. 1878. Report upon forestry. Vol. I. Washington, DC: U.S. Government Printing Office. 650 p.

Hough, F.B. 1880. Report upon forestry. Vol. II. Washington, DC: U.S. Government Printing Office. 618 p.

Thoreau, H.D. 1906. The writings of Henry David Thoreau. Boston: Houghton Mifflin: 184-204.

Forest Inventories Generate Scientifically Sound Information on the Forest Resource, But Do Our Data and Information Really Matter?

Christoph Kleinn[1] and Göran Ståhl[2]

Abstract.—Current research in forest inventory focuses very much on technical-statistical problems geared mainly to the optimization of data collection and information generation. The basic assumption is that better information leads to better decisions and, therefore, to better forest management and forest policy. Not many studies, however, strive to explicitly establish the relationship between information quality and decision quality. In this article, we discuss this issue and suggest that forest inventory research should include more studies on the immediate and indirect effects of results and findings of forest inventories.

Introduction

Forest inventories are carried out to collect scientifically sound and defendable data and generate equally sound and defendable information. This information is demanded and required by decisionmakers and policymakers responsible either for the management of the resource itself or for defining the regulatory framework for resource usage and management. The information requirements are relevant in particular when the resource becomes scarce: wise and target-oriented management is then required in particular, and sustainable usage strategies are in demand so as to secure the long-term provision of goods and services and to balance conflicting user interests.

Forest inventories have a long history; virtually no forest exists worldwide that has not experienced some sort of inventory exercise. Forests are ecosystems and serve as a resource at the same time. They are as variable as the concrete management objectives. Forest inventories are adapted to these very conditions and objectives, be it a mere resource inventory, a forest health assessment, an inventory of nonwood goods and services, an inventory of a small property's or of an entire country's forest resource—to name just a few types of forest inventories.

To adapt to such manifold situations, forest inventory researchers and scientists have developed a highly versatile toolbox of techniques of data collection, making use of many data sources, analysis techniques, and modeling techniques. Given the crucial role of forests in currently intensively debated issues such as climate change, combating desertification, and biodiversity conservation, it is not surprising that forest inventory research is continuing to be intense and that forest information is demanded from many parties.

Browsing through the forest inventory research agenda, however, it calls one's attention that practically all forest inventory research is on technical issues of optimization of efficiency, be it statistical or economic efficiency or both. Much less research is being carried out about questions such as "What minimum information is needed for the sustainable management of a forest resource?" and, maybe even more relevant, "What is the role of scientific information in decisionmaking and policymaking processes?" and "What can be learned from these two questions for forest data acquisition practices?"

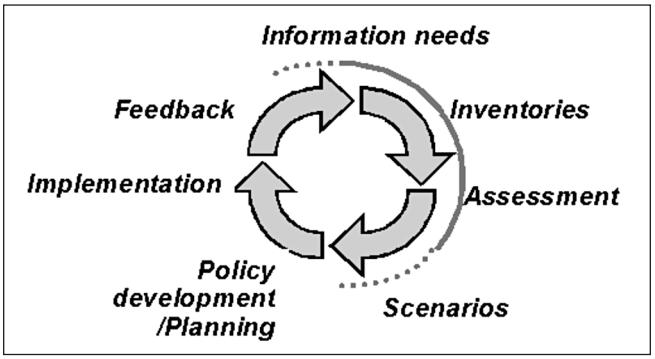
It is somewhat surprising that these fundamental questions are not as widely addressed and discussed in forest inventory research as other topics—although they are probably more relevant than the optimization of sampling techniques because they are the very fundamental questions; optimization

[1] Institute of Forest Management, Georg-August-Universität, Göttingen Büsgenweg 5, D-37077 Göttingen, Germany. E-mail: ckleinn@gwdg.de.
[2] Department of Forest Resource Management and Geomatics, Swedish University of Agricultural Sciences, Umeå, Sweden. E-mail: Goran.Stahl@resgeom.slu.se.

of forest inventories can only take place in the framework defined by the answers to these questions.

The questions are very complex, however. In this article we try to discuss them mainly from the point of view of large-area forest inventories (such as national forest inventories [NFIs]), and address the issue also for management-oriented inventories at the level of forest estates. The overall goal is to contribute to better-focused forest inventory activities in planning, implementation, reporting and communication. This is a discussion article and, as such, rather than giving answers, it poses the questions and discusses plausible paths toward the solutions.

Figure 1.—*In this illustration of the decisionmaking and policymaking processes, data collection takes place in the "Inventories" stage and information generation occurs in the "Assessment" and "Scenarios" stages. The outer line indicates where forest inventory may play a major role (from FAO 2000).*



Forest Inventories in Decision and Policy Processes

Forests are complex systems, whether we look at them as a resource or as an ecosystem. If such a system is to be sufficiently understood for specific management purposes, information is required. Here, we will not expand on the multitude and complexity of definitions of the term "information" but restrict ourselves to the simple and basic view of information as *interpreted data*, where interpretation is meant to be *knowledge based* and data is meant to be the sort of sound and scientifically defendable data generated by inventories on scientific grounds.

In Agenda 21, an entire chapter is dedicated to "Information for Decision Making" (Chapter 40). There, paragraph 40.1 reads in part, "In sustainable development, everyone is a user and provider of information considered in the broad sense. That includes data, information, appropriately packaged experience and knowledge. The need for information arises at all levels, from that of senior decision makers at the national and international levels to the grass-roots and individual levels."

Information generation and provision is but one part in the policymaking and decisionmaking processes, as illustrated in figure 1. There, the outer circle describes the stages of the process where forest inventory expertise plays a major role.

Forest inventory is the process of gathering data. Through analysis and assessment, these data are processed, interpreted, and turned into information that can be used by decisionmakers in forest management or can be used by other interested parties. Forest inventory is usually about key questions such as the following:

- How much is out there at a given point in time (growing stock) and what is the quality and/or value?
- Where is it?
- What are the changes and trends (of the stock, of the uses, of the functions, etc.)?
- How much may sustainably be harvested (accessibility, ownership restrictions, etc)?

In addition to the question of "What and how much," interest also exists in the question, "How precise is the estimation?"

Occasionally, criticism is expressed, in particular, in the context of large-area forest inventories, that the right information is not generated because "… the information presented [from forest inventories] is supply-driven …" and "… the mechanisms that formulate policy-relevant questions are lacking" (Janz and Persson 2002). Even if the right information has been generated, "… too often, scientific information is available, yet policy makers do not use it" (Guldin 2003). This observation is, of course, a phenomenon that is

not restricted to forest-related policymaking but has been stated by Gordillo and Andersson (2004) in more general terms: "… the link between policy evaluation and policy action is often quite weak or entirely missing" (Gordillo and Andersson (2004).

Although forest inventory is probably the scientifically best developed and most systematically organized data collection exercise in the field of renewable natural resources, obviously some basic issues still need to be worked on because "our scientific and technical abilities far outstrip our decision making methods and ability to understand the relationship between science and its many outcomes" (Crow 2000, cited in Guldin 2003).

Much of what we discuss here is about the efficiency of forest inventory data collection. An inventory strategy is efficient, in simple terms, if the defined objectives are achieved with low input. In this context, those who design, implement, and analyze forest inventories are sometimes confronted with seemingly simple questions on the scope of their activities.

When, for example, planning in inventory in a classroom manner, we usually resort to a sample-size formula like $n = t^2 s^2 / A^2$, where A is half the width of the confidence interval (as a measure of precision) and t comes from the t-distribution. On the right-hand side of this equation, the (estimated) population variance s^2 is the only characteristic that has to do with the resource forest itself. The value of A, describing the target precision, needs to be defined on the basis of strategic or demand criteria. While usually even numbers such as 5, 10, 15, or 20 percent are defined as desirable precision levels, the authors of this article do not know scientific (or other) studies that give proof, evidence, and justification that these precision levels are adequate for a specific inventory. It appears largely to be based on tradition and convention. The same holds for the level of statistical significance. An α=5-percent level of error probability is defined by default—without giving more thought to it or a justification about whether these 5 percent are appropriate or not, and why so. In order to be able to define scientifically an optimal level of α and of precision, it would be necessary to formally establish a relationship between them and the objectives of the study, namely the target attributes.

Some Thoughts on Decisionmaking

When the goal is to raise the overall efficiency of forest inventory data provision, it is necessary not only to look at the sampling and data processing stages but also in more general terms to look at the ways and mechanisms how data and information are eventually used and how they affect decisions.

"Information about the subject/matter problem is a basis for good decisionmaking." Many forest inventory reports and many scientific articles on forest inventory optimization begin with that or similar statements that are usually accepted undisputedly adhering to the following two implicit assumptions:

- It pays to spend money for information provision.
- Better information leads to better decisions and hence, to better management.

While assumption 1 is frequently addressed and sometimes questioned when funds are to be allocated to an inventory study, usually no doubt occurs about assumption 2. The relationship between information quantity/quality and decision quality is all but clear, however, and certainly not a simple and "linear" one. In fact, surprisingly few scientific studies attempt to establish or test this relationship. Hardly any study that begins with a statement like the above that "good information is required for good decision making" ends with a critical evaluation of whether the information provided by that specific study actually made the decisions better.

Not only the "objective" information but also the decisionmakers themselves must be taken into account in that context. Data and information are but two factors in decisionmaking; information is only helpful if it is presented in an adequate manner to the target group, decisionmakers, or stakeholders. A certain level of education and expertise is usually required to make rational use of information. The more professional expertise exists, the less —but the more

specific—is the information that is required. More nontechnical factors may affect decisionmaking, including the type and extent of professional experience of the decisionmakers, their position within the institution, their motivation, their cognitive capacities, social and cultural norms, and their advisors. In addition, we may expect interactions between these factors.

Inventories on Different Geographical Scopes

The question of the relevance of information about the forest resource occurs at all geographical levels of forest inventories (stand, enterprise, national, global, etc.), where, however, it is not predominantly the geographical scope but the overall setting of people and institutions that determine the role of scientific information. We briefly discuss three cases:

Community Forestry

In an instructive study on information, communication, and decisionmaking in community forest user groups, Banjade *et al.* (2006) analyze, among other points, the factors that make the users make use of information and the contribution of different types/qualities of information in community forest decisionmaking in a case study in Nepal. Not really surprisingly, scientific information (the type of data-based information as provided by forest inventories) does not play a major role, but the "experiences of community members and stories coming from within the community" outweighed that scientific information by far (Banjade et al. 2006). In an attempt to quantify the contributions of different types/qualities of information in community forest decisionmaking, the following observations resulted: experience (47 percent), stories (18 percent), enthusiasm (14 percent), scientific information (12 percent), and images and representation (9 percent) (Banjade *et al.* 2006).

The fact that knowledge and experience (and also rumors and stories) are relevant to decisionmaking is probably known to everybody also in the context of everyday decisions. It has also to do with the question of to what extent an expert may replace data and scientific evidence. A well-trained forest officer with a longstanding experience in local forest management will probably demand only very specific inventories (if at all). Such a decision system is based completely on trust and belief in the expert, however, and decisions are not necessarily transparent or replicable.

In addition, the "traditional know how" of forest officers becomes less and less valid as the system becomes very complex. "Decision-support systems" are needed that account for all kinds of goods and services and allows different types of analyses to be made. To feed and update these decision-support tools and to generate realistic scenarios, we do obviously need data.

It is not suggested here that expert decisions are all bad, but, if independent monitoring is an issue and external experts want to have insight into the decisions and their background, the scientifically sound and defendable data and information provided from forest inventories on statistical grounds are probably indicated. In regulatory frameworks for community forestry, for example, forest inventories along statistical principles are sometimes required for exactly that reason. Such an inventory needs then to be properly done and used in order to fulfill its functions.

Forest Enterprises

Forest enterprises have clear economic objectives and planning follows usually straightforward managerial paths. Resources (in terms of manpower, time, money) will only be invested if it is economically reasonable; i.e., if a return may be expected that exceeds the investment.

Although generally not used in practice, studies trying to incorporate the "decisionmaking effects" of using different types of data in planning forest resource usage do exist. Most of these studies are based on the minimization of inventory cost plus expected loss due to nonoptimal decisions ("cost-plus-loss analysis"); e.g. Hamilton (1978) and Ståhl (1994). The studies are few, however, and in general they are based on assumptions of what data should be acquired and explicit knowledge on how the quality of the information is related to the loss due to nonoptimal decisions.

In their review article on the influence of data quality on planning processes in forestry, Duvemo and Lämås (2006) state that "in general, a more accurate description of the state of the forest leads to more accurate forecasts … and, hence, to better decisions" (Duvemo and Lämås 2006). They also state, "It is concluded that research in this area is scarce" and that, "… those who seek to evaluate forestry data often oversimplify the problems," which is probably due to the inherent complexity of forest planning (Duvemo and Lämås 2006).

National Forest Policy Formulation

In contrast to forest enterprises where a fairly direct link exists between information procurement and decisionmaking and where economic analyses can possibly be done, NFIs aim at supporting, formulating, and monitoring of forest and related policies—without immediate economic implications. NFIs are carried out for many decades in some countries (but not at all in other countries). Their frequent repetition in some countries may be evidence enough for their usefulness. Studies that establish a clear link between policy decisions and availability of scientifically sound information are scarce. This research question, however, appears to receive more and more attention currently: the Food and Agriculture Organization of the United Nations' (FAO's) project to support national forest assessments convened recently an expert consultation on "Generating knowledge through National Forest Assessments —Towards improved forest, land use and livelihood policies" (FAO 2007). Linking forest assessments better to forest and related policies was a major topic.

Data and information generated by forest inventories for larger areas (including NFIs) serve a multitude of functions, the benefits of which undisputedly exist but are difficult to rate and quantify. Among those benefits are the following:

- The development and evaluation of general forest and environmental policy.
- Supporting the allocation decisions of larger wood-based industries.
- Documenting the state and trends in the development of the forest resource.
- Generating a database and information base for scientific research into forest uses and forest development.
- Informing the public about the state of the forest resource.
- Raising public awareness about the forests and their functions.
- Reporting to international conventions such as the Climate Convention (and the Kyoto Protocol) and the Convention on Biological Diversity as a means to cooperate on the global scale toward sustainable development.

It is suggested to evaluate the benefits of forest inventory information not only in a "one-dimensional manner" as input into immediate decision and policymaking processes but to take the whole of the benefits into account. It may well be, for example, that the direct effect of the provided information on policymakers is much less than the indirect effect that is provoked by a clearly expressed and informed public opinion. An interesting study would, therefore, be to systematically research questions of dissemination and use of the information generated by NFIs: who knows about the results and who uses them for what purpose?

Conclusions

We join the statement of Duvemo and Lämås (2006) that, "evaluation of forest data should also include its usefulness in the forest management and decision process," (Duvemo and Lämås (2006) and would like to extend that also to large-area forest inventories that do not have an immediate forest management objective. It is suggested that forest inventory planners and scientists do also put the lesser technical-statistical topics more seriously on their research agendas, including the following questions:

- *How is forest inventory data and information (and which part of it) being used and for what purposes?* This requires that an inventory does not end with the publication and dissemination of the report, but that some followup is being done: do the results meet the needs and expectations of the users? Are there potential users who miss information that could have been generated? Essentially, we should do an "inventory" of uses and users and of the effects of our forest inventory results and findings.

- *What data is required for different users?* This question refers to the variables of interest, their precision of estimation, the spatial resolution (geographic unit of reference), the periodicity, etc. This is actually a question that comes close to the technical-statistical optimization issues in forest inventory research, because precision is immediately linked to inventory design and inventory cost.
- *How do information requirements and information usage interact with other factors* such *as professional experience, academic and professional education, and position and power within the institution?* It is likely that the information requirements cannot be formulated in absolute terms but must be seen in the very concrete context not only of the forest resource and the biophysical conditions but also in the context of other determinants like the organizational setting, the decision structures, the decisionmakers themselves, and their motivation and agenda.
- *How to optimize the communication strategy?* The role of communication is sometimes underestimated in forest inventory reporting. Whether forest inventory data and information are eventually being used by potentially interested stakeholders depends mainly on the communication strategy. Generating information is not enough. This is not only the case in forestry; Brewer (2006) states the same for conservation biology: "The data we continue to collect and report on … may make no difference … if this information is not translated into meaningful stories …." (Brewer 2006).

To work on these questions, forest inventory experts need to resort to and integrate expertise from various disciplines from the social sciences, such as sociology, psychology, cultural anthropology, etc. It is suggested that the development of the discipline forest inventory will benefit very much if we work on the "information procurement and decisionmaking" topic with an integrated interdisciplinary approach. It is further suggested to integrate such elements also into university teaching: forest inventory must not predominantly be seen as a technical-statistical field that requires quantitative skills above all, but equally as a discipline that is embedded in a multitude of other disciplines and requires a considerable amount of communication and analysis skills beyond sampling.

It is expected that this approach will help to even better guide the technical-statistical optimization of forest and natural resource inventories.

Coming eventually back to the question posed in the article's title: "Do our data and information really matter," the authors clearly believe that yes, scientific information does matter very much, in particular when credibility and transparency are ranked high. In science, "belief" is not enough, however, and forest inventory experts must put more emphasis in adding tangible evidence to this belief.

Literature Cited

Banjade, M.R.; Schanz, H.; Leeuwis, C. 2006. Discourses of information in community forest user groups in Nepal. International Forestry Review. 8(2): 229-240.

Brewer, C. 2006. Translating data into meaning: education on conservation biology. Conservation Biology. 20(3): 689-691.

Crow, M.M. 2000. Linking scientific research to societal outcomes. In: Proceedings, American Association for the Advancement of Science Colloquium on Science and Technology Policy. Washington, DC: American Association for the Advancement of Science: 129-131.

Duvemo, K.; Lämås, T. 2006. The influence of forest data quality on planning processes in forestry—a review article. Scandinavian Journal of Forest Research. 21(4): 328-339.

Food and Agriculture Organization of the United Nations (FAO). 2000. Global forest survey concept paper. For. Res. Assess. Working paper 28. Rome, Italy: Food and Agriculture Organization of the United Nations. 41 p.

FAO. 2007. Generating knowledge through national forest assessments—towards improved forest, land use and livelihood policies. NFM Working paper 05. Rome, Italy: Food and Agriculture Organization of the United Nations. 67 p.

Gordillo, G.; Andersson, K. 2004. From policy lessons to policy actions: motivation to take evaluation seriously. Public Administration and Development. 24: 305-320.

Guldin, R.W. 2003. Forest science and forest policy in the Americas: building bridges to a sustainable future. Forest Policy and Economics. 5: 329-337.

Hamilton, D.A. 1978. Specifying precision in natural resource inventories. In: Integrated inventories of renewable resources. Gen. Tech. Rep. RM-55. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experimentation Station: 276-281.

Janz, K.; Persson, R. 2002. How to know more about forests? Supply and use of information for forest policy. CIFOR Occas. Pap. 36. Bogor, Indonesia: Center for International Forestry Research. 28 p.

Ståhl, G. 1994. Optimal stand level forest inventories under deterministic and stochastic stumpage value assumptions. Scandinavian Journal of Forest Research. 9: 405-412.

A Method for Examining the Impacts of Oregon Land Use Laws on Forest Lands and Farmlands

David L. Azuma[1], Gary Lettman[2], and Erica Hanson[3]

Abstract.—Over the past 8 years, the Pacific Northwest Research Station Forest Inventory and Analysis unit, in conjunction with the Oregon Department of Forestry, Oregon Department of Agriculture, and Oregon Department of Land Conservation and Development, has researched the effect of Oregon's land use laws on the conversion and development of land. The studies have used aerial photography photopoints to classify land use into broad categories and to count the number of structures in an 80-acre circle around the photopoint. Several studies have looked at the probability of management on forest land as a function of increased structures and proximity to more developed zones. A recent State ballot measure has established a system to challenge some of the land use restrictions, and a new study is under way to establish a new baseline for evaluating future development within the State. Using digital imagery, we will be capturing a digital version of the land use polygons, enabling ease of future assessments.

Introduction

Oregon's forest, range, and agricultural lands are extraordinary in their diversity and contribution to the State's economy. Three of the top four industries in the State are timber, agriculture, and tourism. Other industries can tout the quality of life partially based on the forest-farm image to attract the best employees. Maintaining and enhancing the contribution of nonurban areas is essential to the well-being of Oregonians.

In 1973, the Oregon Legislative Assembly passed the Land Conservation and Development Act. The act required counties and cities to prepare comprehensive land use plans in accordance with statewide goals; Goals 3 and 4 sought to preserve forest lands and farmlands. By the 1980s, most of the comprehensive plans were completed. Each of the plans identified lands that were already built on and committed to residential uses. These areas were zoned for continued development, and expansion into other areas was limited.

By 1996, the land use laws had been in effect for 16 years; however, no definitive study existed that showed the impacts of the laws. The Pacific Northwest Research Station Forest Inventory and Analysis (FIA) unit, in conjunction with the Oregon Department of Forestry, Oregon Department of Agriculture, and Oregon Department of Land Conservation and Development, designed a study to investigate the effects of the land use laws on non-Federal lands in western Oregon. In this study, 24,000 photopoints were interpreted from 3 sets of aerial photographs taken in 1974, 1982, and 1994. In 2002, these same locations were interpreted on aerial photographs taken in 2000. The principal finding was a decrease in the rates of conversion between the 1974–82 period and the 1982–94 period and a continued reduced rate of conversion in the 1994–2000 period (Azuma *et al.* 1999. Lettman *et al.* 2002). Together, these studies suggest that most forest land and farmland development over the past few decades has tended to occur near existing urban areas and especially within urban growth boundaries implemented under Oregon's land use laws. Where rural development has occurred, it has been correlated with reductions in private forest management. Although the conversion of land to more urbanized conditions slowed, the average number of developments per point continued to increase through the 2000 period. A similar study was conducted in eastern Oregon (Lettman *et al.* 2004).

[1] Researcher Forester, U.S. Department of Agriculture (USDA) Forest Service, Pacific Northwest Research Station, Forest Inventory and Analysis (FIA) Unit, Portland Forestry Sciences Laboratory, 620 S.W. Main Street, Suite 400, Portland, OR 97205. Phone: 503–808–2047. E-mail: dazuma@fs.fed.us.

[2] Principal Forest Economist, Oregon Department of Forestry, 2600 State Street, Salem, OR 97310. Phone: 541–945–7408. E-mail: GLETTMAN@ODF.STATE.OR.US.

[3] Forestry Technician, USDA Forest Service, Pacific Northwest Research Station, FIA Unit, Portland Forestry Sciences Laboratory, 620 S.W. Main Street, Suite 400, Portland, OR 97205.

The data from these studies was also utilized to investigate other questions. One study looked at the effect of population growth and urban expansion on private forestry practices (Kline *et al.* 2003a). Another study used structure count data to model the spatial distribution of humans in the Oregon Coast Range (Kline *et al.* 2003b). These studies have informed the debate about land use planning in Oregon.

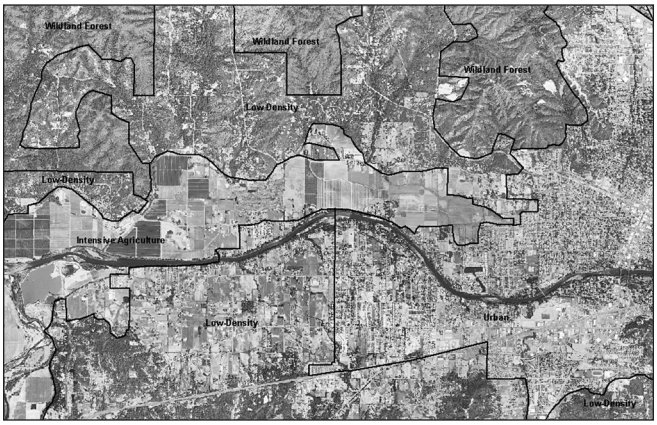
In 2004, Oregon voters passed Ballot Measure 37. The measure provides that a private landowner is entitled to receive just compensation when a land use regulation, implemented after the landowner obtains the property, restricts the use of the property and reduces its fair market value. Alternatively, Measure 37 would allow governments to remove, modify, or not apply the regulation rather than paying compensation. As of December 1, 2006, around 3,600 claims existed of which 2,360 were in the northern Willamette Valley. More than 100,000 acres would be affected by these claims, with the majority of the effect being on farmland. With the prospect of modified land use planning going into effect, we decided to initiate a new study to capture another snapshot of land use in 2005. This article reports on the pilot portion of the new study. New techniques and imagery were used in Josephine County in western Oregon.

Methods

This study uses 2005 National Agriculture Imagery Program (NAIP) imagery (0.5-m resolution) as the base layer for evaluating land use zones. The land use zone polygons from the 2000 National Aerial Photography Program (NAPP) aerial photographs were hand digitized onto the new NAIP imagery, photopoints were located, and zones were labeled (fig. 1). During the transference of 2000 NAPP polygons onto the 2005 NAIP imagery, any changes to the 2000 polygons were made on the 2005 imagery.

The coordinates from the 1990 systematic aerial photography grid (0.85-mi scale) were overlaid onto the NAIP imagery. The coordinates had varying levels of accuracy depending on how they were acquired. As a first check, 30 aerial photo images with a pinprick for a point were compared to the coordinates located on the digital imagery by visual estimate of where

Figure 1.—*Example of land use zone polygons on National Agriculture Imagery Program imagery for Josephine County, OR.*



the pinprick was on the imagery. Most of the coordinates corresponded well, within 30 m; in 3 out of 30 test points, a difference of more than 30 m occurred, with one point being 1,100 m off. Because of the location differences, all points that had large changes in either the number of structures, distance to other zones, or the zone call itself were checked for location correspondence between the digital imagery and the aerial photography. In cases where the point on the imagery differed by more than 30 m, the point was moved to correspond to the pinprick on the aerial photograph.

At each of the 741 locations in Josephine County, three attributes were collected; the land use zone, distance to other zones, and number and location of structures within an 80-acre circle around the point. The land use zones collected included the following:

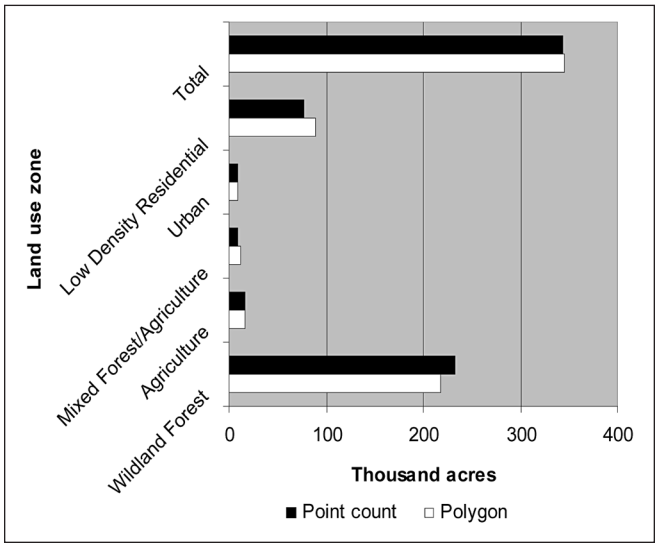
- Wildland forest.
- Intensive agriculture.
- Mixed forest/agriculture.
- Urban.
- Low-density residential/commercial.

(Lettman *et al.* 2002)

A comparison of the area in various land use zones was made between the expanded point estimate and the area from the polygons created on the NAIP imagery.

Structure counts were taken at all nonurban classified points and consisted of a count of structures within an 80-acre template centered over a point. When a structure was identified, the location was recorded. Structure (or "dwelling") counts include individual buildings and/or clusters of buildings, and these buildings may or may not be related to the management of the land they are on. For example, a farmer's house would be included in the count. A cluster of barns, storage sheds, and other associated buildings on rural land would be considered one "structure" for the count. Where four houses might appear clustered, however, such as on the corners of a road intersection, they would be counted as four separate structures because they likely represent four different parcels of ownership, each with a house built on it.

Figure 2.—*Area comparison of expanded photo-interpreted points and Geographic Information System polygons.*



Results

As one might expect, very little change occurred between the 2000 photography and the 2005 imagery. A total of 6 points out of 741 changed land use zones during the period, all due to low density or urban zones expanding on the 2005 imagery. Four other points also changed land use zones, but closer examination showed a difference in how the polygon was interpreted or, in one case, that the point had been moved. The acreages computed by point expansion and summation of polygon area were remarkably close; the totals differed by less than 0.5 of a percentage point (fig. 2).

In about 200 points, the distance to other zones changed between the 2000 and 2005 estimates. A closer look at these differences shows that actual change occurred on only 35 of these points. The rest of the changes were attributed to a slightly different photo scale in the NAPP imagery, 1:43,000 instead of 1:40,000, and to areas where the zones lines were drawn differently even though no actual change occurred.

The structure count data is the most subjective piece of data because it depends on the ability of the photo interpreter to be able to identify the structures. The new Geographic Information System imagery allows a user to "zoom" in for a closer look at what could be a structure. While more than half the points had the same 80-acre count, many had changed, mostly due to an increase in the count by one to four structures but in about one-third of the cases a decrease in count occurred. Some of these differences maybe accounted for by the up-to-30-m difference in the location of the coordinates compared with the pinprick location. It is very unlikely that structures were removed; the count decreases were mostly due to the ability of the interpreter to better distinguish clusters of structures. In areas where structures are more abundant, individuals are more likely to have differing counts.

Discussion

With new ballot measures taking effect in Oregon, it is critical that we track changes in forest, range, and agricultural lands. This study provides an approach that can be easily repeated in the future to detect change. The added location information on the structure count data will provide ease of comparison with future time periods. To evaluate differences in structure counts from past studies, it may be necessary to adopt some threshold for what we will consider to be real change. In areas where structures are more abundant, it may not be feasible to obtain precise comparisons with results from previous studies. In other areas, such as wildland forest, increases of more than two structures may be sufficient to represent real change. Further

investigation as to whether we can stratify the points to obtain the actual change in numbers of structures is needed. This data will also provide distance-specific information for structures when used with FIA's plot data and may provide some assessment of the effects of population density on forest management.

Literature Cited

Azuma, D.L.; Birch, K.R.; DelZotto, P.; Herstrom, A.A.; Lettman, G.J. 1999. Land use change on non-federal land in western Oregon, 1973-1994. Salem, OR: Oregon Department of Forestry. 55 p.

Kline, J.D.; Azuma, D.A.; Alig, R. 2003a. Population growth, urban expansion, and private forestry in western Oregon. Forest Science. 50(1): 33-43.

Kline, J.D.; Azuma, D.A.; Moses, A. 2003b. Modeling the spatially dynamic distribution of humans in the Oregon (USA) coast range. Landscape Ecology. 18: 347-361.

Lettman, G.J.; Azuma, D.L.; Birch, K.R; Herstrom, A.A.; Kline, J.D. 2002. Land use change on non-federal land in western Oregon, 1973-2000. Salem, OR: Oregon Department of Forestry. 48 p.

Lettman, G.J.; Azuma, D.L.; Birch, K.R.; Herstrom, A.A.; Kline, J.D. 2004. Land use change on non-federal land in eastern Oregon, 1975-2001. Salem, OR: Oregon Department of Forestry. 42 p.

Research Applications of Ecosystem Patterns

Robert G. Bailey[1]

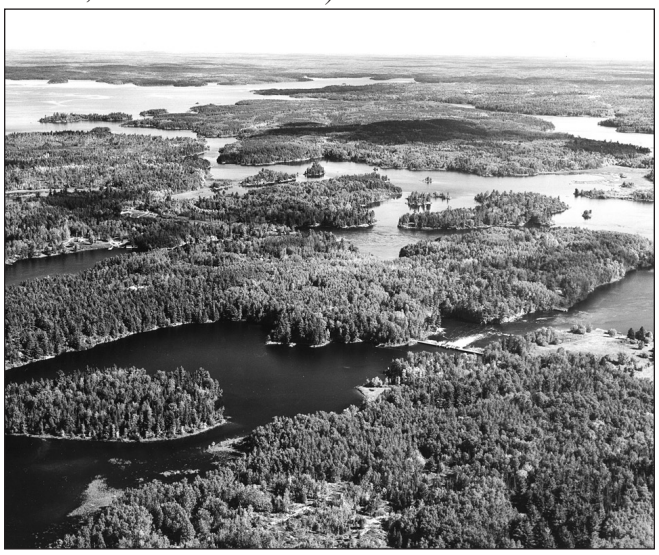
Abstract.—This article discusses the origins of natural ecosystem patterns from global to local scales. It describes how understanding these patterns can help scientists and managers in two ways. First, the local systems are shown within the context of larger systems. This perspective can be applied in assessing the connections between action at one scale and effect at another, the spatial transferability of models, and the links between terrestrial and aquatic systems. Second, scientists and managers can benefit because they gain information about the geographic patterns in ecosystems. Consequently, they are in a better position to design sampling networks, transfer knowledge, and analyze ecosystem diversity. The usefulness of multiscale analysis of ecosystem patterns suggests new scientific directions for research and points the way for restructuring the Forest Service, U.S. Department of Agriculture research programs.

Introduction

The public land-management agencies have been phasing in a new approach to land management. They are shifting from their focus on individual resources to a more holistic approach of managing *whole ecosystems*. This newer approach examines the vertical structure, which is how the ecosystem components (e.g., climate, biota, soil) are integrated at a site. The current reductionist, phenomena-based research is replaced by a *synthesis* approach (cf. Silbernagel 2005).

Ecosystems come in different scales that are nested within each other. The boundaries are open and permeable, leading to interaction, or linkages, between systems. These fluid boundaries can cause identification problems because ecosystems related by geography are not necessarily related by taxonomic properties (Bailey 1996). For example, an area of spruce forests and glacially scoured lakes in Voyageurs National Park, Minnesota constitutes a single system, linked by flows of water and nutrients through podzolized soils toward oligotropic lakes (fig. 1). These are not solely terrestrial or aquatic systems, but geographical ecosystem units. Because of these linkages, modification of one system affects surrounding systems, sometime adversely. Most research provides only data on the vertical structure of the components. There is, however, a need to consider horizontal structure, or connections between systems, because systems are linked to form larger systems. Altering larger systems may affect smaller ones. We must examine these relationships between systems of different scales to analyze the effects of management.

Figure 1.—*A landscape of spruce forests and glacially scoured lakes in Voyageurs National Park, MN.(Photograph by Jack Boucher, National Park Service.)*



[1] U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, 240 W. Prospect Road, Fort Collins, CO 80526. E-mail: rgbailey@fs.fed.us.

Scale of Patterns

Repeated relationships or patterns emerge at varying scales. For example, in the Boulder, CO area, the rocky, forested Front Range slopes of the Rocky Mountains that rise abruptly from the grassy plains are the most prevalent patterns in that region. In the mountains, trees that respond to additional moisture are seen on north-facing slopes. Numerous rocky outcrops appear on the nearby Great Plains grasslands, which act as reservoirs to support islands of trees and shrubs.

Causes of Ecosystem Pattern at Varying Scales

To delineate these patterns and to understand how and why they are distributed, we must understand the processes of how they form, which is important in understanding their dynamics and how they respond to management.[2] Although the ecosystem concept implies equality among all the components of the system, all the components are not equally significant. Climate largely determines ecosystem boundaries. As it changes, the other components change in response.

Controls over the climatic effect change with scale. At the macroscale, ecosystem patterns are controlled by the macroclimate (i.e., the climate that lies just beyond the local modifying irregularities of landform and vegetation). The effects of latitude, continental position, and elevation combine to form the world's ecoclimatic zones, also know as *ecoregions*.

On a mesoscale within the same macroclimate, we commonly find several broad-scale landform patterns that break up the zonal pattern. Hammond's (1954) landform classification (based on surface geometry) is useful in capturing this effect on zonal climate and provides a basis for further differentiation of ecosystems, known as *landscape mosaics*. The automatic classification of maco morphological landforms using geographic information systems (GIS) and digital elevation models has been investigated by Brabyn (1998). He applied a process developed by Dikau *et al*. (1991), which automates Hammond's manual procedure, to the South Island of New Zealand.

A landform unit consists of different ecosystem patterns depending upon the way it breaks up the zonal climate. For example, the Idaho Mountains and Yellowstone Plateau, Wyoming, are both highlands in a temperate steppe climate. The Idaho Mountains are made up of various site-specific ecosystems in a complex pattern, including riparian, forest, and grassland. Deep dissection of the mountain range has resulted in variously oriented slopes with varying local climates. Steep slopes oriented at different angles to the sun add complexity to the otherwise simple arrangement of elevational zones. The Yellowstone Plateau, however, does not have these spotty distribution patterns because its landform is relatively uniform.

At the microscale, local variation in topography will cause small-scale variations in the amount of solar radiation received, create topoclimates, and affect the amount of soil moisture. There are three classes of topoclimate: normal, hotter than normal, and colder than normal. Normal topoclimates occur on flat slopes. In the northern hemisphere, steep, south-facing slopes are hotter than normal, with north-facing slopes colder than normal. In differentiating local sites within topoclimates, soil moisture regimes have been found to be the feature that most significantly segregates plant communities. This feature is the edaphic (related to soil) factor. A common division of the soil moisture gradient is normal, drier, and wetter. Normal soil moisture regimes are found on well-drained uplands. The other regimes occur in topographic positions ranging from drier to wetter, from the top to the bottom of a slope, as well as in bottomlands where the water lies near the surface. These variations subsequently affect the biota, creating ecosystem *sites* as subdivisions of a larger zone (ecoregion). New remote sensing data and GIS analysis methods offer some promise to predicting the distribution of these sites (cf. Franklin 1998).

Ecosystem sites recur in predictable patterns within an ecoregion. The temperate continental zone of southern Ontario, Canada, which also extends south into the northeastern United States, provides a good example of the pattern of sites formed by edaphic-topoclimatic differentiation. On level or moderately rolling areas where the soil is well drained but moist, a maple-

[2] We do not necessarily need more data, which allows an understanding of pattern, because such data cannot generate an understanding of the processes that create these patterns.

beech community is the terminal succession, known as the climatic climax. Where the soil remains wetter or drier than normal, a somewhat different end community occurs. The climatic climax theoretically would occur over the entire ecoregion, but topography leads to different local climates, which partially determine edaphic conditions.

Disturbance and Succession

Disturbance and the subsequent development of vegetation are key contributors to pattern on the landscape at a variety of spatial and temporal scales. The composition of the vegetation of the ecosystem changes with time in a sequence from pioneer vegetation through successional series of intermediate steps to a relatively stable state called late successional vegetation. The late successional types are used to characterize ecosystems because they tend to be far more site specific than pioneer types, which might occur over a wider range of conditions. Furthermore, they are used as baselines for contending with the temporal variability associated with disturbance regimes and attending successional states of vegetation.

Applications

The study of ecosystem patterns has many applications, such as research and management. Ecosystem patterns must be analyzed at multiple levels: local (site) and groups of geographically related sites know as landscapes or ecoregions. This section describes how understanding these patterns can help scientists and managers in two ways.

Local Systems Within Context

First, local systems (sites) are seen within the context of the larger system.

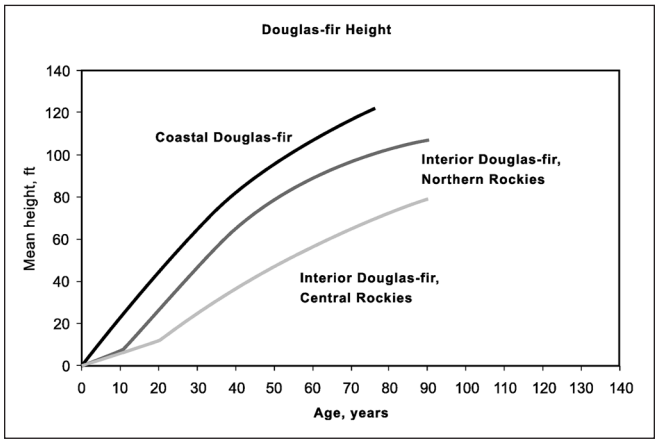
Connections Between Action at One Scale and Effect at Another. The perspective of seeing context can be applied in assessing the connections between action at one scale and effect at another. For example, logging on upper slopes of an ecological unit may affect downstream riparian and meadow habitats.

With the ecosystem approach, the interaction between sites can be understood. Processes emerge that are not evident at the site level. An example is a snow-forest landscape that includes dark conifers that cause snow to melt faster than either a wholly snow-covered or a wholly forested basin. Landscapes function differently as a whole than would have been predicted by analysis of the individual elements (cf. Marston 2006).

Spatial Transferability of Models. Predictive models differ between larger systems. The same type of forest growing in different ecoregions will occur in a different position in the landscape and have different productivity. For example, figure 2 shows that the height-age ratio of Douglas-fir (*Pseudotsuga mensiesii*) varies in different climatically defined ecoregions. The ecoregion determines which ratio to apply to predict forest yield. This is important, because if a planner selects the wrong ratio, yield predictions and the forest plans upon which they are based will be wrong. The ecoregion map is helpful in identifying the geographic extent over which results from site-specific studies (such as growth and yield models) can be reliably extended. Thus the map identifies areas for the spatial transferability of models.

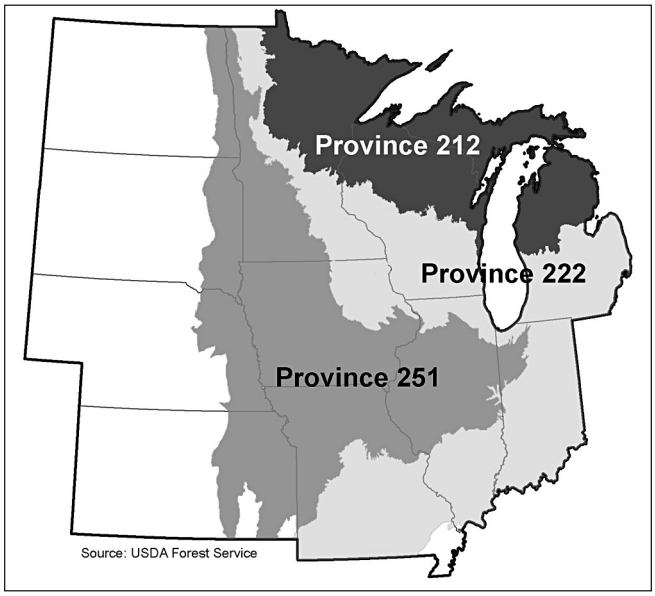
In Canada, studies have found that the height-diameter models of white spruce (*Picea glauca*) were different among different ecoregions (Huang *et al*. 2000). Incorrectly applying a height-diameter model fitted from one ecoregion to different ecoregions resulted in overestimation between 1 and 29 percent, or underestimation between 2 and 22 percent.

Figure 2.—*Height-age ratio of Douglas-fir varies in different climate-defined ecoregions. (Source: Chad Keyser, Forest Service, National Headquarters, Fort Collins, CO.)*



There is another, even more compelling, example. Each of five previous regional Forest Inventory and Analysis (FIA) programs has developed its own set of volume models, and the models have been calibrated for regions defined by political boundaries corresponding to groups of States rather than ecological boundaries. Further, the regional models sometime bear little resemblance to each other. The same tree shifted a mile in various directions to move from southwest Ohio (previous Northeastern FIA) to southeast Indiana (previous North Central FIA) to northern Kentucky (Southern FIA) could have quite different model-based estimates of volume. Growth estimates are likely improved if growth models are calibrated by ecoregions rather than States or FIA regions.[3] Figure 3 shows the ecoregions within the former North Central FIA Region, which were used for the application of volume and biomass equations (Hansen 2002). Statistical analysis of additional FIA data could further test the value of this approach.

Figure 3.—*Ecoregion provinces for the North Central Forest Inventory and Analysis Region of the United States: 212 = Laurentian Mixed Forest; 222 = Eastern Broadleaf Forest (Continental); 251 = Prairie Parkland Parkland (Temperate). (Source: Ronald McRoberts, Forest Service, Northern Research Station, St. Paul, MN.)*



Links Between Terrestrial and Aquatic Systems. Streams are dependent on the terrestrial system in which they are embedded. They therefore have many characteristics in common, including biota. Delineating areas with similar climatic characteristics makes it possible to identify areas within watersheds with similar aquatic environments. A good example is the distribution of the northern hog sucker (*Hypentelium nigricans*) in the Ozark Uplands of Missouri, which covers several watersheds. This species of fish is widespread but not uniformly distributed throughout the Mississippi River basin. In Missouri, it is found almost exclusively in the Ozark Uplands ecoregion.

Geographic Patterns in Ecosystems

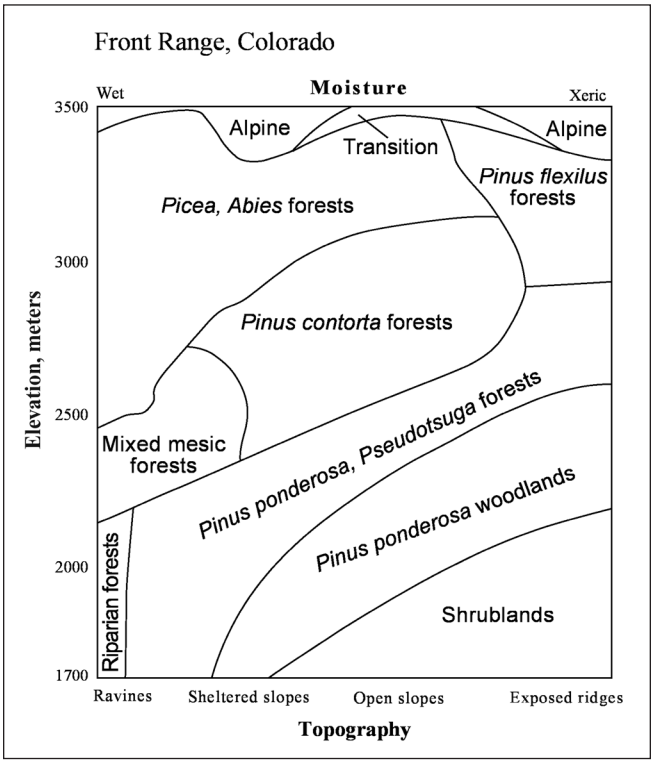
Second, multilevel analysis of ecosystems also provides information about the geographic patterns in ecosystems.

Design of Sampling Networks. Users are thus in a position to design efficient sampling networks. As we have seen, ecosystems recur in predictable patterns within an ecoregion. Data from representative sampled sites can be applied to analogous (unsampled) sites with a high degree of reliability (Bailey 1991). Identification of sites based on ecoregional classification could be used to impute their characteristics from sampled FIA sites, for example, using *k*-Nearest Neighbors or similar techniques (McRoberts *et al*. 2002).

Another example comes from the Rocky Mountains, a temperate steppe mountains ecoregion. This ecoregion, like all ecoregions, is a climatic region within which specific plant successions occur upon specific landform positions. The most likely successional series growing on a site within an ecoregion can be predicted from landform information if the vegetation-landform relationships are known in a particular ecoregion. For example, Douglas-fir forests occur on moist, mid-elevation sites within the Front Range of Colorado. Figure 4 shows the relationships between elevation-topography and climax plant communities. Understanding these relationships, vertically and horizontally, within ecoregion delineations allows the transfer of knowledge from research sites (or inventory plot) to

[3] See McRoberts (2005).

Figure 4.—*Relationships between elevation-topography and climax plant communities, Front Range, CO. (Source: Peet (1981) in Bailey (1996).)*



like sites within the same ecoregion. In fact, O'Brien (1996) found that surveys involving comprehensive sampling efforts will more accurately characterize unmonitored sites (plots) when samples are stratified according to ecologically similar areas such as ecoregions. Unfortunately, we often do not understand the spatial relationships between the FIA plots and the landform-vegetation types within a particular ecoregion. If these were developed, we could likely produce better small-area estimates of vegetation conditions.

Transfer Information. Ecoregions show areas that are analogous with respect to ecological conditions. This makes it possible to transfer knowledge gained from one part of a continent to another and from one continent for application in another; however, this approach should be used with caution. Because of compensating factors, the same forest type can occur in different ecoregion divisions. It has no regional alliance. For example, ponderosa pine (*Pinus ponderosa*) forest occurs in the northern Rockies and the southwest United States. This distribution does not imply that the climate, topography, soil, and fire regime are necessarily the same. The climate of Williams, AZ, for instance, is characterized by a late spring drought, whereas Colorado Springs is moist throughout the year.

This distinction is important because some ecoregions have a tendency for large wildfires; the ratio of large to small wildfires decreases from east to west (Malamud *et al*. 2005). Also, fire recurrence interval differs markedly between ecoregions with greater frequency in the West. This information suggests that the results of these studies can be used to assess burn probability across the nation to identify areas of high risk. Then government agencies could plan better for wildfire hazards. Ecoregions can also be used as a baseline from which to assess natural fire regimes, which can be used to abate the threat of fire exclusion and restore fire-adapted ecosystems.

Analyze Ecosystem Diversity. Maps of landscape mosaics reveal the relative diversity of ecosystems. For example, maps from the Coastal Plain landscape in North Carolina, which is flat and relatively undissected, show relatively few boundaries between landscape patterns. In contrast, maps from the Piedmont landscape display just the opposite character, reflecting greater ecosystem diversity.

Significance to Forest Service Research

It is important to link the ecosystem hierarchy with the Forest Service research hierarchy (table 1). In so doing, research structures and ecosystem hierarchies correlate such that research information, mapping levels, and research studies work well together. Comparison of research structures and ecosystem levels can identify gaps in the research network.

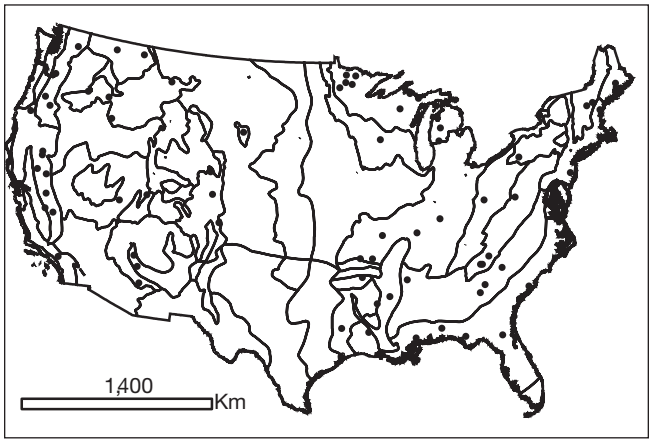
Table 1.—*Relationship between the ecosystem and research hierarchies.*

Ecosystem hierarchy*	Research hierarchy
Ecoregion	Research Station (multi-ecoregions)
Landscape mosaic	Experimental Forest/Range, Watershed
Site	FIA plot, LTER site, Research Natural Area

FIA = Forest Inventory and Analysis. LTER = Long Term Ecological Research.
**Source: Bailey (1988a)*

At the ecoregional scale, existing research locations can be compared with ecoregion maps to identify underrepresented regions or gaps in the network (fig. 5). For example, experimental forests or ranges occur in only 26 of 52 ecoregion provinces (Lugo *et al*. 2006). Several ecoregions have no research facilities while others have more than one. The greatest number (14) falls within the Laurentian mixed forest ecoregion of the Lake States and Northeast. A more comprehensive analysis could include other types of similar research sites, such as Long Term Ecological Research sites, Research Natural Areas, and the like. This analysis could reveal gaps in coverage both across and within ecoregions.

Figure 5.—*Approximate boundaries of ecoregion provinces within the conterminous United States and locations of experimental forests and ranges administrated by the Forest Service. (Source: Olga Ramos, Forest Service, International Institute of Tropical Forestry, San Juan, PR.)*



Restructuring Forest Service Research Programs

The many useful applications of the study of ecosystem patterns suggest new scientific directions for research and point the way for restructuring Forest Service research programs. To address critical ecological issues, it is essential to move from the traditional single-scale management and research on plots and stands to mosaics of ecosystems (landscapes and ecoregions), and from streams and lakes to integrated terrestrial-aquatic systems (i.e., geographical ecosystems). FIA thematic maps (e.g., biomass, forest types) could assist with this.

Some Research Questions

These studies reveal useful applications of ecosystem patterns. Many relevant research questions associated with these patterns still remain, including the following: What are the natural ecosystem patterns in a particular ecoregion? What are the effects of climatic variation on ecoregional patterns and boundaries? And, what are the relationships between vegetation and landform in different ecoregions? While some have suggested that GIS analysis can assist in answering these questions, that approach should be used with caution because it will help identify pattern, but it cannot generate an understanding of the processes that create these patterns (Bailey 1988b).

Natural Ecosystem Patterns

Historically, a high level of landscape heterogeneity was caused by natural disturbance and environmental gradients. Now, however, many forest landscapes appear to have been fragmented due to management activities such as timber harvesting and road construction. To understand the severity of this fragmentation, the nature and causes of the spatial patterns that would have existed in the absence of such activities should be considered. This analysis provides insight into forest conditions that can be attained and perpetuated (Knight and Reiners 2000).

Effects of Climatic Variation

Current climate exerts a very strong effect on ecosystem patterns, and climate change may cause shifts in those patterns (Neilson 1995). Anthropogenic and climatic change could yield ecoregions that are much different, or less useful, after many years; therefore, temporal variability is an important research issue. While several researchers are doing work on the effect of climate change on tree species distribution (cf. Iverson and Prasad 2001), others are working on the impact of climatic change on the geography of ecoregions. For example, Jerry Rehfeldt[4] has predicted the potential distribution of the American (Mojave-Sonoran) Desert ecoregion under the future climate scenario produced by the IS92a[5] scenario of the Global Climate Model, with about 5 °F warming and 50 percent increase in precipitation. He has produced maps that show a greatly

[4] Rehfeldt, J. 2005. Personal communication, retired, USDA Forest Service, Rocky Mountain Research Station, 1221 S. Main Street, Moscow, ID 83843.

[5] This is one of the emissions scenarios developed in 1992 under the sponsorship of the Intergovernmental Panel on Climate Change. IS92a has been widely adopted as a standard scenario for use in impact assessments.

expanding desert under this scenario. Despite the percentage increase in precipitation, the amount of rainfall fails to keep pace with the increase in temperature, so the climate becomes more arid.

There are limits to the number of monitoring sites that can be established for monitoring changes in the global environment. Obviously, sites should be representative. Stations also should be located where they can detect change. The boundaries between climate-controlled ecoregions are suitable for this purpose. FIA has roughly 160,000 forested sample sites. This criterion could identify a subset of these sites that could be more intensively sampled to provide needed monitoring information.

Relationships Between Vegetation and Landform

The relationships between vegetation and landform position changes from ecoregion to ecoregion, reflecting the effect of the macroclimate. A species occupies different positions in the landscape. For the same soil moisture condition but with different topoclimates, tree species change their positions in different regions (table 2). With these changes, related changes occur in the vigor of other tree species, ecosystem productivity, and so on. Knowledge of these differences is important for extending results of research and management experience, and for designing sampling networks. These relationships have been extensively studied in some regions (cf. Odom and McNab 2000) but, unfortunately, not in others. Where sufficient studies have been done, these relationships might be modeled and mapped to improve understanding of these ecosystems.

Table 2.—*Relationships between vegetation and landform in various ecoregions in Ontario, Canada. (From Burger (1976)).*

Ecoregion	Topoclimate		
	Hotter	Normal	Colder
1	P		
2	P		P
3	P		P
4	A	P	P
5	A		A,P
6	C		A,P
7			C,A

P = Picea glauca (white spruce)*; A = Acer saccharum* (sugar maple)*; C = Carya ovata* (shagbark hickory).

Relationship to Ecomap

The Forest Service Ecomap program (Cleland *et al*. 1997), which was initiated in 1993, follows the same ecosystem hierarchy presented here; however, it recognizes more levels but on the same principles. It has been adopted by the Forest Service as part of an ecological mapping system to support ecosystem management. Note that the finer scale delineations are not generally available beyond the boundaries of the national forests.

Acknowledgments

The author thanks Eric Smith, Ron McRoberts, and Jerry Rehfeldt, who suggested ideas to include in this article. The author also expresses his appreciation for the helpful criticism of Ray Czaplewski and Keith Moser.

Literature Cited

Bailey, R.G. 1988a. Ecogeographic analysis: a guide to the ecological divisions of land for resource management. Misc. Pub. 1465. Washington, DC: U.S. Department of Agriculture, Forest Service. 16 p.

Bailey, R.G. 1988b. Problems with using overlay mapping for planning and their implications for geographic information systems. Environmental Management 12: 11-17.

Bailey, R.G. 1991. Design of ecological networks for monitoring global change. Environmental Conservation. 18(2): 173-175.

Bailey, R.G. 1996. Ecosystem geography. New York: Springer. 216 p.

Brabyn, L. 1998. GIS analysis of macro landform. In: Proceedings, spatial information research centre's 10th colloquium. New Zealand: University of Otago: 35-48.

Burger, D. 1976. The concept of ecosystem region in site classification. In: Proceedings, international union of forest research organizations XVI world congress, division 1. Oslo, Norway: Norwegian Forest Research Institute: 213-218.

Cleland, D.T.; Avers, P.E.; McNab, W.H.; Jensen, M.E.; Bailey, R.G.; King, T.; Russell, W.E. 1997. National hierarchical framework of ecological units. In: Boyce, M.S.; Haney, A., eds. Ecosystem management. New Haven, CT: Yale University Press: 181-200.

Dikau, R.; Brabb, E.E.; Mark, R.M. 1991. Landform classification of New Mexico by computer. U.S. Geological Survey open-file rep. 91-634. Denver, CO: U.S. Geological Survey. 26 p.

Franklin, J. 1998. Predicting the distribution of shrub species in California chaparral and coastal sage communities from climate and terrain-derived variables. Journal of Vegetation Science. 9: 733-748.

Hammond, E.H. 1954. Small-scale continental landform maps. Annals Association of American Geographers. 44: 33-42.

Hansen, M. 2002. Volume and biomass estimation in FIA: national consistency vs. regional accuracy. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C.; Moser, J.W., eds. Proceedings, third annual forest inventory and analysis symposium. Gen. Tech. Rep. NC-230. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station: 109-120.

Huang, S.; Prince, C.; Titus, S.J. 2000. Development of ecoregion-based height-diameter models for white spruce in boreal forests. Forest Ecology and Management. 129: 125-141.

Iverson, L.R.; Prasad, A.M. 2001. Potential changes in tree species richness and forest community types following climate change. Ecosystems. 4: 186-199.

Knight, D.H.; Reiners, W.A. 2000. Natural patterns in southern Rocky Mountain landscapes and their relevance to forest management. In: Knight, D.H.; Smith, F.W.; Buskirk, S.W.; Romme, W.H.; Baker, W.L., eds. Forest fragmentation in the southern Rocky Mountains. Boulder, CO: University Press of Colorado: 15-30.

Lugo, A.E.; Brown, S.L.; Dodson, R.; Smith, T.S.; Shugart, H.H. 2006. Long-term research at the USDA Forest Service's experimental forests and ranges. Bioscience. 56(1): 39-48.

Malamud, B.D.; Millington, J.D.A.; Perry, G.W. 2005. Characterizing wildfire regimes in the United States. Proceedings of the National Academy of Sciences. 102: 4694-4699.

Marston, R.A. 2006. President's column. Ecoregions: a geographic advantage in studying environmental change. AAG Newsletter. 41(3): 3-4.

McRoberts, R. 2005. Personal communication, Mathematical Statistician. U.S. Department of Agriculture, Forest Service, North Central Experimental Station, 1992 Folwell Avenue, St. Paul, MN 55108.

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

Neilson, R.P. 1995. A model for predicting continental-scale vegetation distribution and water balance. Ecological Applications. 5: 362-385.

O'Brien, R.A. 1996. Forest resources of northern Utah ecoregions. Res. Bull. INT-RB-87. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station. 43 p.

Odom R.H.; McNab, W.H. 2000. Using digital terrain modeling to predict ecological types in the Balsam Mountain of western North Carolina. Res. Note SRS-8. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 11 p.

Peet, R.K. 1988. Forests of the Rocky Mountains. In: Barbour, M.G.; Billings, W.D., eds. North American terrestrial vegetation. Cambridge, England: Cambridge University Press: 63-102.

Silbernagel, J. 2005. Bio-regional patterns and spatial narratives for integrative landscape research and design. In: Tress, B.; Tres, G.; Fry, G.; Opdam, P., eds. From landscape research to landscape planning: aspects of integration, education, and application. Wageningen UR Frontis Series. Vol. 12. Dordrecht, The Netherlands: Wageningen UR: 107-118.

Developing Survey Grids To Substantiate Freedom From Exotic Pests

John W. Coulston[1], Frank H. Koch[2], William D. Smith[3], and Frank J. Sapio[4]

Abstract.—Systematic, hierarchical intensification of the Environmental Monitoring and Assessment Program hexagon for North America yields a simple procedure for developing national-scale survey grids. In this article, we describe the steps to create a national-scale survey grid using a risk map as the starting point. We illustrate the steps using an exotic pest example in which the purpose of the survey is to substantiate freedom from the pest.

Introduction

Exotic insects and pathogens pose a serious threat to forests in the United States. Damage and control costs for these pests have been estimated to exceed $4.3 billion annually (Pimentel *et al.* 2006). Moreover, these pests may cause severe ecological effects, in some cases virtually eliminating important tree species. In a recent example, the emerald ash borer (*Agrilus planipennis*), first detected in the United States in 2002, has killed more than 20 million ash (*Fraxinus* sp.) trees in Michigan, Ohio, and Indiana and continues to expand its range (USDA Forest Service *et al.* 2006). Exotic forest pests are commonly introduced to the United States through imports of raw wood products, live plants, and other commodities or in packing materials. If an introduced pest becomes established in a particular location, it may expand to new areas by a variety of human-mediated pathways (e.g., interstate shipment of contaminated nursery plants) and by its natural dispersal ability. Pests that are currently found only in limited distribution in portions of the United States but with the potential for major impact if spread elsewhere, include *Phytophthora ramorum* (which causes sudden oak death) and the sirex woodwasp (*Sirex noctilio*).

The 1999 Presidential Executive order establishing the National Invasive Species Council stated that Federal agencies have several critical duties in minimizing the impacts of exotic pests. In particular, these agencies—including the Forest Service, U.S. Department of Agriculture—are mandated to detect and rapidly respond to populations of invasive species in a cost-effective, environmentally sound manner and to perform reliable and accurate monitoring of invasive species populations (Clinton 1999). Fulfilling these mandates, especially at broad geographic scales, requires the establishment of protocols that address the detection, assessment, and monitoring phases of exotic pest surveillance. In the detection phase, we are generally interested in determining whether the exotic pest exists outside its original introduction area. If the pest exists outside the area of original infestation, then it may be important to assess the infestation's geographic extent; in particular, this assessment may determine if quarantine or other regulatory protocols are required. In areas where the pest of interest is established, monitoring to determine whether the prevalence level is increasing or decreasing is essential. Focusing specifically on the detection phase, key protocols include determining the sample size and reliability of surveys to substantiate freedom from an exotic pest and designing the exotic pest survey in a consistent manner once the sample size is known.

Substantiating freedom from an exotic pest requires that two conditions be defined *a priori*. First, a prevalence threshold must be set. The prevalence threshold sets the detection limit, meaning the minimum detectable level of infestation in any given landscape (e.g., 1 percent, 5 percent). If the prevalence

[1] Supervisory Research Forester, U.S. Department of Agriculture (USDA) Forest Service, Southern Research Station, 4700 Old Kingston Pike, Knoxville, TN 37919. E-mail: jcoulston@fs.fed.us.

[2] Research Associate, Department of Forestry and Environmental Resources, Box 8008, North Carolina State University, Raleigh, NC 27695. E-mail: fkoch@fs.fed.us.

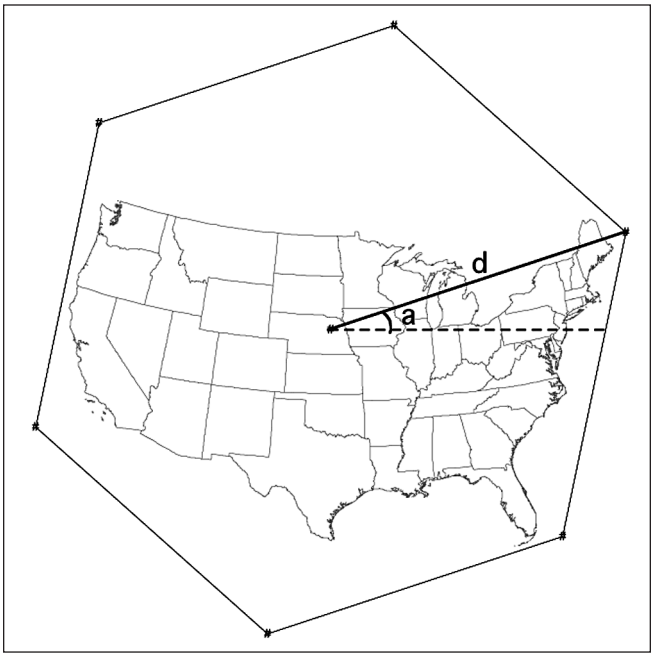
[3] Assessment Coordinator, USDA Forest Service, Southern Research Station, 3041 Cornwallis Road, Research Triangle Park, NC 27709. E-mail: bdsmith@fs.fed.us.

[4] Forest Health Technology Enterprise Team Director, 2150 Centre Avenue, Building A, Suite 331, Fort Collins, CO 80526. E-mail: fsapio@fs.fed.us.

threshold is zero, then a complete enumeration is needed. Second, the desired confidence must be identified. The selected confidence should reflect the required level of certainty (e.g., 95 percent, 99 percent). Once these conditions are defined, "substantiating freedom from disease" has context, and an appropriate method can be developed to estimate the sample size required to, for example, substantiate that an exotic pest of interest does not exist outside its original introduction area, above a 1-percent prevalence threshold, with 99 percent confidence. Coulston *et al.* (2008) developed techniques to estimate required sample size to substantiate freedom from exotic pests based on desired confidence levels and detection thresholds. These sample size estimation techniques were based on existing epidemiological approaches (Cameron and Baldock 1998) extended to the spatial domain. Here we do not address sample size estimation; rather, we describe techniques to develop survey grids in a consistent fashion, based on a global sampling design, once the required sample size has been identified.

White *et al.* (1992) developed a global Environmental Monitoring and Assessment Program (EMAP) sampling grid which serves as the basis for the Forest Service Forest Inventory and Analysis (FIA) Phase 2 (Forest Mensuration) and Phase 3 (Forest Health). The EMAP sampling grid was developed from a truncated icosahedron made up of 20 hexagons and 12 pentagons covering the planet, with one hexagon advantageously placed to cover North America (fig. 1). A noteworthy aspect of the EMAP grid's configuration is that this hexagon can be systematically intensified, yielding a wide range of potential sample frames. In short, this process provides a straightforward framework for creating systematic survey grids. The chief objective with this article is to demonstrate intensification of the EMAP North American hexagon to create a survey grid for substantiating freedom from a hypothetical exotic pest. In support of this objective, we develop a nonlinear regression model to estimate the appropriate intensification factor, provide a table with all possible intensification factors of the EMAP North American hexagon, and describe spatial intersection techniques to implement the sampling grid based on a risk map.

Figure 1.—*The North American hexagon used to generate the Ecological Mapping and Assessment Program sampling grid and other sampling grids.*



Methods

We developed a simulated risk map of susceptibility to a hypothetical forest pest for the conterminous United States. The risk map divides the country into two strata: low risk and high risk; areas within each stratum have equal probability of exposure to a hypothetical exotic pest. For this demonstration, we assumed that 500 sample areas throughout the high-risk stratum would be adequate to substantiate freedom from our hypothetical exotic pest above a 5-percent prevalence threshold with 90 percent certainty. We then intensified the EMAP North American hexagon to develop a sample grid with the appropriate number of points. We identified a final set of sample areas for our survey by spatial intersection of the sample grid and the risk map.

Risk Maps

We constructed the raster risk map (25 km^2 resolution) for our hypothetical pest via spatial overlay of three data layers, representing host species distribution (oak basal area), favorable recent weather conditions (mean January 2003–2005

temperatures), and potential pest spread (based on distance to significant ports). We developed the oak basal area data layer from recent FIA Phase 2 tables for each State, calculating total oak basal area (in m^2ha^{-1}) per Phase 2 plot from corresponding tree tables. We then used inverse distance weighting interpolation to generate a raster map from the (perturbed and swapped) plot points, masking nonforest areas with a forest map developed from MODIS imagery by the Forest Service Remote Sensing Applications Center. We developed the map of mean January temperatures (ºC) using weather observation station data from the National Oceanic and Atmospheric Administration's National Climatic Data Center. For each station in the conterminous United States, we calculated the mean January temperature for 2003–2005. We then created a raster map from the station data using gradient-plus-inverse distance squared interpolation, which combines multiple linear regression (with x, y, and elevation as potential explanatory variables) and distance weighting (Nalder and Wein 1998). For the potential pest spread layer, we first identified a set of "high-risk" U.S. marine ports (i.e., ports receiving raw wood products or packing materials from East Asia) based on U.S. Army Corps of Engineers foreign marine cargo data. We then constructed a national raster map depicting, for each pixel, the Euclidean distance to the closest port. By fitting a beta distribution (α=1.25, β=4) to the full range of distances in this map, we generated a second raster map of simulated introduction probabilities, with probability peaking at ~100 km from the closest port. In the risk map created from these three layers, the high-risk stratum was defined as forested areas with greater than 2.296 m^2ha^{-1} of oak basal area, mean January temperature greater than –9.44 °C, and simulated introduction probability greater than 0.2.

Estimating the Intensification Factor

The EMAP North American hexagon has an area of ~ $1.79 \cdot 10^7$ km^2 and is defined by six vertices and a center point. The hexagon can be intensified by 3, 4, and 7 or any product of these factors (see the Intensifying the North American Hexagon section for details). By intensifying in this manner, tessellations of smaller, equal-area hexagons are created and the required survey grid density can be achieved (table 1). Based on table 1, we developed a model to estimate the intensification factor required to meet the desired number of sample areas for a given survey. We used the following equation and nonlinear regression to develop the model:

Table 1.—*The associated hexagon area for each intensification factor of the North American hexagon (Spence and White 1992).*

Intensification factor	Hexagon area (km^2)
144	40,600
324	18,000
576	10,200
1,296	4,500
2,304	2,500
5,184	1,100
9,216	640
11,664	500
16,384	350

$$X = e^{\left(a - \ln\left(\frac{A}{na}\right)\right)} + \varepsilon, \quad (1)$$

where

X = the nominal intensification factor,

a = estimated parameter,

A = the total area (km^2) of the risk stratum of interest from the risk map,

na = the number of sample areas required, and

ε = error.

The value of A/na is the area that each intensification point represents and is analogous to hexagon area in table 1.

The idea behind equation (1) is to answer the following question: if one point represents $1.79 \cdot 10^7$ km^2 (this is the center point of the North American hexagon), then how many times should the sample be intensified so that each point represents A/na? By rearranging equation (1) we have simplified the relationship to the following:

$$X = e^a\left(\frac{na}{A}\right) + \varepsilon = a\left(\frac{na}{A}\right) + \varepsilon.$$

To apply this model, once the nominal intensification factor is calculated, it is compared to the intensification factors in table 2, which lists all possible factor sequences (i.e., all possible products of the factors 3, 4, and 7) up to an intensification factor of 50,176. The user then selects an 'actual' intensification factor from table 2, typically the one that is closest to the nominal intensification factor.

Table 2.—*Intensification factor and sequence used to generate sampling grids from the North American hexagon.*

Intensification factor	Sequence	Intensification factor	Sequence	Intensification factor	Sequence	Intensification factor	Sequence
3	–	448	4•4•4•7	3,969	3•3•3•3•7•7	16,384	4•4•4•4•4•4•4
4	–	567	3•3•3•3•7	4,032	3•3•4•4•4•7	16,464	3•4•4•7•7•7
7	–	576	3•3•4•4•4	4,096	4•4•4•4•4•4	16,807	7•7•7•7•7
9	3•3	588	3•4•7•7	4,116	3•4•7•7•7	19,683	3•3•3•3•3•3•3•3•3
12	3•4	729	3•3•3•3•3•3	5,103	3•3•3•3•3•3•7	20,412	3•3•3•3•3•3•4•7
16	4•4	756	3•3•3•4•7	5,184	3•3•3•3•4•4•4	20,736	3•3•3•3•4•4•4•4
21	3•7	768	3•4•4•4•4	5,292	3•3•3•4•7•7	21,168	3•3•3•4•4•7•7
27	3•3•3	784	4•4•7•7	5,376	3•4•4•4•4•7	21,504	3•4•4•4•4•4•7
28	4•7	972	3•3•3•3•3•4	5,488	4•4•7•7•7	21,609	3•3•7•7•7•7
36	3•3•4	1,008	3•3•4•4•7	6,561	3•3•3•3•3•3•3•3	21,952	4•4•4•7•7•7
48	3•4•4	1,024	4•4•4•4•4	6,804	3•3•3•3•3•4•7	26,244	3•3•3•3•3•3•3•3•4
49	7•7	1,029	3•7•7•7	6,912	3•3•3•4•4•4•4	27,216	3•3•3•3•3•4•4•7
63	3•3•7	1,296	3•3•3•3•4•4	7,056	3•3•4•4•7•7	27,648	3•3•3•4•4•4•4•4
64	4•4•4	1,323	3•3•3•7•7	7,168	4•4•4•4•4•7	27,783	3•3•3•3•7•7•7
81	3•3•3•3	1,344	3•4•4•4•7	7,203	3•7•7•7•7	28,224	3•3•4•4•4•7•7
84	3•4•7	1,372	4•7•7•7	8,748	3•3•3•3•3•3•3•4	28,672	4•4•4•4•4•4•7
108	3•3•3•4	1,701	3•3•3•3•3•7	9,072	3•3•3•3•4•4•7	28,812	3•4•7•7•7•7
112	4•4•7	1,728	3•3•3•4•4•4	9,216	3•3•4•4•4•4•4	34,992	3•3•3•3•3•3•3•4•4
144	3•3•4•4	1,764	3•3•4•7•7	9,261	3•3•3•7•7•7	35,721	3•3•3•3•3•3•7•7
147	3•7•7	1,792	4•4•4•4•7	9,408	3•4•4•4•7•7	36,288	3•3•3•3•4•4•4•7
189	3•3•3•7	2,187	3•3•3•3•3•3•3	9,604	4•7•7•7•7	36,864	3•3•4•4•4•4•4•4
192	3•4•4•4	2,268	3•3•3•3•4•7	11,664	3•3•3•3•3•4•4	37,044	3•3•3•4•7•7•7
196	4•7•7	2,304	3•3•4•4•4•4	11,907	3•3•3•3•7•7	37,632	3•4•4•4•4•7•7
243	3•3•3•3•3	2,352	3•4•4•7•7	12,096	3•3•3•4•4•4•7	38,416	4•4•7•7•7•7
252	3•3•4•7	2,401	7•7•7•7	12,288	3•4•4•4•4•4•4	45,927	3•3•3•3•3•3•3•3•7
256	4•4•4•4	2,916	3•3•3•3•3•3•4	12,348	3•3•4•7•7•7	46,656	3•3•3•3•3•3•4•4•4
324	3•3•3•3•4	3,024	3•3•3•4•4•7	12,544	4•4•4•4•7•7	47,628	3•3•3•3•3•4•7•7
336	3•4•4•7	3,072	3•4•4•4•4•4	15,309	3•3•3•3•3•3•3•7	48,384	3•3•3•4•4•4•4•7
343	7•7•7	3,087	3•3•7•7•7	15,552	3•3•3•3•3•4•4•4	49,152	3•4•4•4•4•4•4•4
432	3•3•3•4•4	3,136	4•4•4•7•7	15,876	3•3•3•3•4•7•7	49,392	3•3•4•4•7•7•7
441	3•3•7•7	3,888	3•3•3•3•3•4•4	16,128	3•3•4•4•4•4•7	50,176	4•4•4•4•4•7•7

Intensifying the North American Hexagon

To develop a method for systematically constructing survey plot networks, we first generated a triangular grid of points from the original seven points defining the North American hexagon (fig. 1). The geometric properties of the triangular grid allow for intensifications of 3, 4, 7, or any product of these factors. For implementation purposes, the North American hexagon can also be directly intensified by a factor of 9, which is a special case of the 3x3 intensification. The intensification grids are created by spawning additional points at regularly spaced distances from points in the initial grid. Conceptually, if an initial point is located at (0, 0) on the unit circle in Cartesian space, then additional points are spawned as follows:

1. 3X: add 2 points at locations $(0, -1/\sqrt{3})$ and $(-1/2, -1/2\sqrt{3})$.
2. 4X: add 3 points at locations $(1/4, -\sqrt{3}/4)$, $(-1/4, -\sqrt{3}/4)$, and $(-1/2, 0)$.
3. 7X: add 6 points at locations $(5/14, \sqrt{3}/14)$, $(2/7, -\sqrt{3}/7)$, $(-1/14, -3\sqrt{3}/14)$, $(-5/14, -\sqrt{3}/14)$, $(-2/7, \sqrt{3}/7)$, and $(-2/7, \sqrt{3}/7)$.
4. 9X: add 8 points at locations $(1/6, -1/2\sqrt{3})$, $(1/3, -1/\sqrt{3})$, $(0, -1/\sqrt{3})$, $(-1/3, -1/\sqrt{3})$, $(-1/2, -1/2\sqrt{3})$, $(-2/3, 0)$, $(-1/3, 0)$, and $(-1/6, -1/2\sqrt{3})$.

As the factor sequences in table 2 suggest, grids for larger intensification factors are created by first spawning points from the original seven-point grid at one of the basic factor levels (i.e., 3X, 4X, or 7X), then intensifying the resulting point grid again at one of these factor levels, and so on.

Operationally survey plot networks should only be developed using projected coordinates. Spence and White (1992) suggested using Lambert's azimuthal equal-area projection when working with the EMAP grid since it minimizes scale distortion. For our hypothetical pest example, we assumed Lambert's azimuthal equal-area projection with the following parameters: North American Datum of 1983, GRS80 spheroid, distance units in meters, and 6,378,137 m sphere radius; the center of this projection is –96° 00' 00" longitude and 37° 30' 00" latitude.

Two additional pieces of information are necessary for intensification: the angle α at which the grid deviates from the East-West direction and the distance δ between grid points. For the seven-point grid that defines the North American hexagon, given the aforementioned projection parameters, $\alpha \approx 18.8817°$ and $\delta \approx 2628774.8$ m (fig. 1).

In the following example of 3X intensification, steps that are ultimately applied to every point in an initial grid are illustrated in terms of one point. The point (x_0, y_0) is first centered on the unit circle in Cartesian space at (x_c, y_c) by $x_c = x_0 - x_0$ and $y_c = y_0 - y_0$. Two new points, $(x_{c1'}, y_{c1'})$ and $(x_{c2'}, y_{c2'})$, are then added at $\delta \cdot (0, -1/\sqrt{3})$ and $\delta \cdot (-1/2, -1/2\sqrt{3})$, respectively. The point $(x_{c1'}, y_{c1'})$ is rotated to match α by adjusting the coordinates as follows: $x_{c1} = x_{c1'}\cos(\alpha) - y_{c1'}\sin(\alpha)$ and $y_{c1} = x_{c1'}\sin(\alpha) + y_{c1'}\cos(\alpha)$. The point $(x_{c2'}, y_{c2'})$ is rotated in the same manner. The points are then shifted back to geographic space by $x_{01} = x_{c1} + x_0$ and $y_{01} = y_{c1} + y_0$ and $x_{02} = x_{c2} + x_0$ and $y_{02} = y_{c2} + y_0$. (See White *et al.* 1992 for additional information.) Once the grid has been intensified to the target density, hexagons are created by Thiessen expansion (i.e., points in the grid serve as hexagon centroids).

Spatial Intersection

After determining the intensification factor and creating an intensified point grid corresponding to the high-risk stratum for our hypothetical pest, we used spatial intersection to extract only those grid points that coincided with the high-risk stratum in our raster risk map. We accomplished this using the "SAMPLE" command in ArcInfo®, which extracts the value of a raster map pixel for each sample point. If a sample point from the high-risk grid fell within a high-risk pixel in the raster map, it was kept; otherwise, it was dropped from the final set of points. The final sampling hexagons were then the set of hexagons derived from Thiessen expansion of the retained sample points.

Results

Equation (1) was fit to the data in table 1 using PROC NLIN (SAS Institute, Inc. 2004). The final form of the model was

$$X = 5783883\left(\frac{na}{A}\right) \qquad (2)$$

The only estimated parameter was *a*, which was 5,783,883 with a standard error of 23,123 and a 95-percent confidence interval of 5,730,560 to 5,837,206. Figure 2 displays the model and original data points; notably, this model can be applied to estimate the nominal intensification factor for any future survey grid design efforts.

The final risk map had 920,975 km^2 of high-risk area and 6,857,225 km^2 of low-risk area. As previously noted, our goal was approximately 500 plots in the high-risk stratum. By applying equation (2), the nominal intensification factor for the high-risk stratum was 3,140. The actual intensification factor was 3,136 (table 2), which was implemented by intensifying $4^3 \cdot 7^2$. This intensification resulted in a systematic triangular grid of 4,093 points in the North American hexagon (fig. 3a). After a risk map value was extracted for each point and all points that did not fall in high-risk areas were deleted, 493 grid points remained (fig. 3b). The stage one sampling hexagons were the Thiessen polygons representing the 493 grid points (fig. 3c).

Figure 2.—*Results from the nonlinear regression. The squares represent data points and the solid line represents equation (2).*

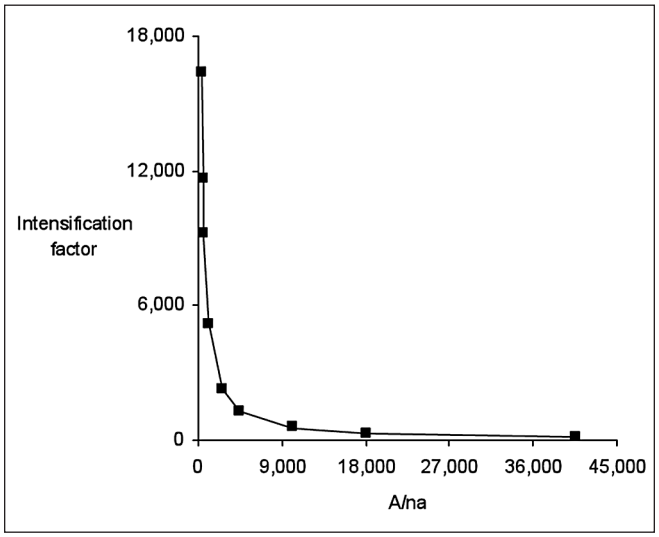
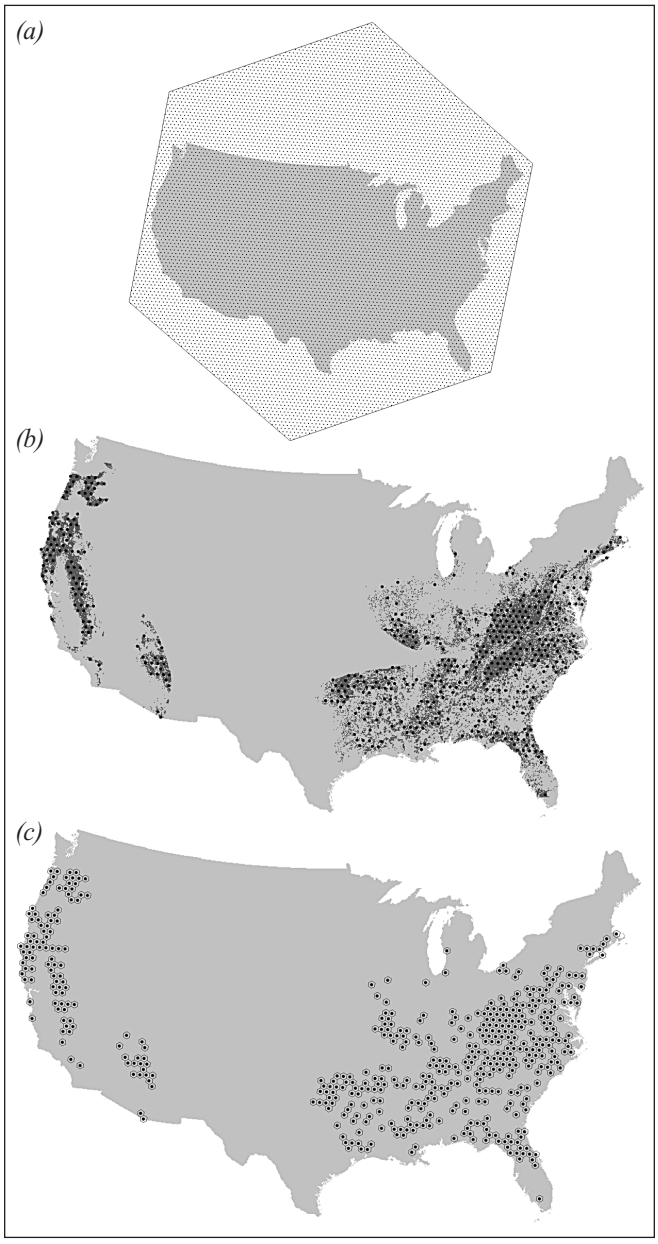


Figure 3.—*(a) The North American hexagon intensified 3,136 X; (b) the intersection of the high-risk stratum and hexagon centers resulting from the intensification; (c) final stage one sampling hexagons for the example survey.*



Discussion

The purpose of the survey grid methodology described here is to substantiate freedom from exotic pests. In general, a systematic grid will be adequate for determining whether an exotic pest of interest exists at specified levels outside the area of original detection. Certain situations exist, however, where a systematic grid may not be adequate. For example, when an exotic pest is expected to be associated with riparian areas in the conterminous United States, the approach presented here may not yield an appropriate distribution of sample areas. In situations such as this, we recommend following the guidelines of Stevens and Olsen (2004) to develop stage one sample areas based on restricted randomization.

In this article, we focused on the development of stage one sample areas for a single stratum survey. The outlined techniques are easily adaptable, however, to multistage samples where two or more risk strata exist. Suppose that our goal was to develop a survey grid that had 500 sample areas in the high-risk stratum and 50 sample areas in the low-risk stratum. To develop the multistage survey, we would estimate the intensification factor for each stratum using equation (2) and then select the actual intensification factor from table 2. A grid would be created for each of the two strata and intersected with the risk map. The final survey grid would then be the compilation of points from each grid that intersected the corresponding stratum in the risk map.

The spatial structure of risk maps plays a vital role in the development of useful survey grids. Risk maps with little within-strata spatial autocorrelation have a "salt-and-pepper" appearance. This autocorrelation will result in a survey grid with a similar salt-and-pepper appearance. Risk maps with high within-strata spatial autocorrelation, however, will have a smooth appearance where the likelihood that areas from the same stratum are adjacent is high. The appearance of both the risk map and the survey grid is partially related to the spatial resolution (i.e., pixel size) of the risk map and the data layers that were used to develop it. For the purposes of designing a national scale survey to substantiate freedom from an exotic pest, we suggest a pixel size of approximately 25 km^2 or larger to limit the influence of pixel size on survey grid design.

As global trade and the influx of foreign goods to the United States escalate, the probability that new exotic pests will be accidentally introduced also increases. If unchecked, some of these exotic pests are likely to become established and, in turn, impact the Nation's forest resources. The techniques described here offer a relatively simple way to monitor such emerging forest pest threats. Significantly, the foundation for these techniques is a global sampling design that is already being used by FIA; indeed, its capacity for systematic, hierarchical intensification may already be familiar to many scientists. Furthermore, adoption of standardized approaches like the one outlined here enables us to rapidly and more efficiently respond to exotic pest threats and, thus, better fulfill our current mandate in this regard.

Literature Cited

Cameron, A.R.; Baldock, F.C. 1998. Two-stage sampling in surveys to substantiate freedom from disease. Preventive Veterinary Medicine. 34: 19-30.

Clinton, W.J. 1999. Executive Order 13112 of February 3, 1999: Invasive Species. Federal Register. 64(25): 6183-6186.

Coulston, J.W.; Koch, F.H.; Smith, W.D.; Sapio, F.J. 2008. Exotic pest surveillance: survey development and reliability. Canadian Journal of Forest Research: 2422-2433.

Nalder, I.A.; Wein, R.W. 1998. Spatial interpolation of climatic normals: test of a new method in the Canadian boreal forest. Agricultural and Forest Meteorology. 92: 211-225.

Pimentel, D.; Zuniga, R.; Morrison, D. 2006. Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecological Economics. 52: 273-288.

SAS Institute, Inc. 2004. SAS/STAT® 9.1 User's Guide. Cary, NC: SAS Institute, Inc.

Spence, M.H.; White, D. 1992. EMAP sampling grid technical report. Corvallis, OR: ManTech Environmental Technology, Inc.; U.S. Environmental Protection Agency, Environmental Research Laboratory. 64 p.

Stevens, D.L.; Olsen, A.R. 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association. 99(465): 262-278.

U.S. Department of Agriculture (USDA) Forest Service; Michigan Department of Agriculture; Michigan Department of Natural Resources; USDA Animal and Plant Health Inspection Service; Michigan State University; Purdue University; Ohio State University. 2006. Emerald ash borer. http://www.emeraldashborer.info/. (07 November).

White, D.; Kimerling, J.A.; Overton, W.S. 1992. Cartographic and geometric components of a global sampling design for environmental monitoring. Cartography and Geographic Information Systems. 19(1): 5-22.

Predicting the Ability To Produce Emerald Ash Borer: A Comparison of Riparian and Upland Ash Forests in Southern Lower Michigan

Susan J. Crocker[1,2], Deborah G. McCullough[3], and Nathan W. Siegert[4]

Abstract.—Concern for the future of ash trees in the United States has risen since the 2002 discovery of emerald ash borer (EAB) (*Agrilus planipennis* Fairmaire) in southeastern Michigan. The ability of ash forests in the Southern Lower Peninsula of Michigan to produce EAB was compared by physiographic class and stand size. Results showed that EAB production potential was significantly higher in riparian forests and in large- and medium-diameter ash stands. The potential to support dense populations of EAB suggests that riparian forests may have higher ash mortality and a greater ability to influence the rate of EAB spread.

Introduction

Four years after being identified in the United States, emerald ash borer (EAB) (*Agrilus planipennis* Fairmaire, Coleoptera: Buprestidae) has killed an estimated 15 million ash trees in southeastern Michigan (Cappaert *et al.* 2005). Discovered near Detroit, MI, in June 2002, EAB is an Asian native that is believed to have entered the country in solid-wood packing material in the early nineties (Cappaert *et al.* 2005). EAB is a wood boring beetle that feeds in the inner bark of ash trees. It attacks native species of ash (*Fraxinus* spp.), including white (*F. americana*), black (*F. nigra*), and green ash (*F. pennsylvanica*) and many planted cultivars (Cappaert *et al.* 2005). Primary insect damage is caused as larvae feed and produce galleries within the phloem and outer sapwood. Tree mortality occurs within 1 to 3 years of initial attack (Haack *et al.* 2002, McCullough and Katovich 2004).

The known United States distribution of EAB is centered in southeastern Michigan. Outlier infestations have also been detected throughout the Lower Peninsula of Michigan, northwestern Ohio, northern Indiana, northeastern Illinois, and southwestern Ontario, Canada (Cappaert *et al.* 2005, Illinois Department of Agriculture 2006). In addition, transportation of infested firewood and infested nursery stock has introduced EAB to Maryland and the Upper Peninsula of Michigan (Maryland Department of Agriculture 2006, Michigan Department of Agriculture 2005).

Ash species are found on a variety of sites, from moist, well-drained uplands to partially inundated swamps and floodplains (Barnes and Wagner 2004). Because of its tolerance to flooding, ash is a dominant species in riparian forests and makes up a substantial portion of the overstory canopy. Riparian forests have a strong interaction with surface waters and are vulnerable to disturbance. Changes that occur in riparian forests have significant impacts on forest structure and composition, water quality, and stream and terrestrial habitat (Tepley *et al.* 2004). Because ash is a principal component of riparian forests, EAB presents a major threat to this vast network of resources.

In this study, the potential abilities of ash forests in different physiographic regions were compared to stand-size classes to produce adult EAB. Specifically, this study was designed to test the hypothesis that riparian ash forests have a greater potential to produce EAB compared to upland ash forests. The motivating factor was to assess the potential impact of EAB establishment in riparian forests and to evaluate the role of riparian ash in EAB dispersal.

[1] Forester, U.S. Department of Agriculture, Forest Service, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108. E-mail: scrocker@fs.fed.us.
[2] Graduate Student, State University of New York, College of Environmental Science and Forestry, 1 Forestry Drive, Syracuse, NY 13210.
[3] Professor, Departments of Entomology and Forestry, Michigan State University, 243 Natural Science Building, East Lansing, MI 48824.
[4] Postdoctoral Research Associate, Department of Entomology, Michigan State University, 243 Natural Science Building, East Lansing, MI 48824.

Methods

This study was conducted in the Southern Lower Peninsula (SLP) of Michigan (fig. 1). To estimate phloem area and determine EAB production potential for riparian and upland forests in the SLP, plot- and tree-level data were selected from the U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis (FIA) database. Plots were selected from the most current inventory and were measured between 2000 and 2004. Under the national FIA plot design, all trees greater than 12.5 cm in diameter at breast height[5] (d.b.h.) are measured on four 7.32 m radius subplots and saplings with d.b.h. between 2.5 and 12.5 cm are measured on four 2.07 m radius microplots. Seedlings (d.b.h. is less than 2.5 cm) are counted but detailed, individual measurements are not recorded (Miles *et al.* 2001).

Figure 1.—*Study area, Southern Lower Peninsula, Michigan.*



Sixty plots with an ash forest type were randomly selected to analyze EAB production potential. The sample was selected so that one half of the plots had a riparian physiographic class (n = 30) and the other half of the plots had an upland physiographic class (n = 30). An equal number of plots (n = 10) from riparian and upland physiographic classes represented one of three stand-size classes: large-diameter, medium-diameter, or small-diameter.

For each plot, the diameter of each live ash tree greater than 4 cm d.b.h. was used to calculate phloem area (y) using the following second order polynomial equation that was developed to fit the relationship between phloem area (surface area) and d.b.h. ($r^2 = 0.9501$) (McCullough and Siegert, 2007):

$$y = 0.0216x^2 - 0.0922x \quad (1)$$

where y = phloem area and x = d.b.h. (cm). The phloem area of each ash tree (n = 216 on riparian plots and n = 136 on upland plots) was multiplied by the mean estimate of the potential number of adult EAB that can be produced per m^2 of phloem, defined as EAB production potential (McCullough and Siegert, 2007) (table 1). EAB production potential is an average, calculated for 5 different diameter classes, of the total number of observed EAB adults found per m^2 of phloem and is based on results from 148 ash trees. The production potential for each ash was summed with all other ash trees on the same plot to yield the total EAB production potential per plot.

To test the difference in total EAB production potential per plot by physiographic class and stand-size, a two-way analysis of variance was performed using SAS software (SAS Institute, Inc. 2004).

Table 1.—*Mean emerald ash borer production potential per m^2 of phloem area by diameter class (McCullough and Siegert, 2007).*

Diameter class (cm)	Mean (number of adult insects)
2.3–13	68.8 (5.86)
14–25	108.3 (10.08)
26–42	106.2 (10.15)
43–60	105.2 (11.18)
> 60	94.7 (8.28)

Note: Standard error in parentheses.

[5] Diameter at breast height is measured at 1.3 m (USDA Forest Service 2001).

Results

Riparian and upland ash forests differ in their potential to produce EAB. Riparian ash forests had a significantly higher total EAB production potential than upland ash forests ($\alpha = 0.05$, $P < 0.01$). Stand-size class was also found to impact production potential. Total EAB production potential was significantly higher in the medium- and large-diameter classes than in the small-diameter stand-size class ($\alpha = 0.05$; $P < 0.01$ for the medium-diameter class, $P = 0.01$ for the large-diameter class). The medium- and large-diameter stand-size classes were not significantly different ($\alpha = 0.05$, $P = 0.74$). The interaction effect between physiographic class and stand-size class was not significant ($\alpha = 0.05$, $F = 0.25$, df = 2, $P = 0.78$).

Discussion

Riparian ash forests were found to have a significantly greater potential to produce adult EAB than upland ash forests. The difference in estimates of EAB production potential is due to a difference in ash density. A higher degree of species diversity (i.e., non-ash hardwoods) in upland forests is likely to contribute to the lower total production potential found among upland ash forests. In general, riparian forests contain more ash than upland forests and have more phloem area. Greater phloem area increases the potential area available to support the development of EAB. Higher total production potential in riparian forests indicates that EAB has a greater reproductive ability in riparian forests. Damage within riparian forests may be more significant, since populations of EAB may build more rapidly and spread more quickly. As such, the rate at which EAB can spread may be greater in riparian forests.

In addition to physiographic class, total EAB production potential was also found to be influenced by stand-size class. Medium- and large-diameter stands, where the majority of stocking is in trees larger than 12.5 cm, have more available phloem area. As a result, these stands can produce more EAB compared to small-diameter stands. EAB presents a significant risk to medium- and large-diameter stands, as these stands are areas of potentially high EAB density.

Eradication and containment are primary responses to EAB. Current control strategies include the removal of all ash within a 0.5 mile radius of an infested tree, voluntary reduction of ash in un-infested areas, and the application of insecticides (Poland and McCullough 2006). Results suggest that such efforts to mitigate the spread of EAB should be focused in riparian forests and in medium- to large-diameter stands where the potential to support high densities of EAB is the greatest. Knowledge of potential epicenters of EAB activity will allow managers to (1) reduce the rate of EAB spread, (2) lessen damage caused by EAB infestations, and (3) effectively allocate limited resources and management efforts.

Acknowledgments

The authors thank Cassandra Olson, Therese Poland, and Robert Venette of the Forest Service, Northern Research Station for reviewing this manuscript.

Literature Cited

Barnes, B.V.; Wagner, W.H., Jr. 2004. Michigan trees, a guide to the trees of the Great Lakes Region. Ann Arbor: MI: University of Michigan Press. 447 p.

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: 152-165.

Haack, R.A.; Jendek, E.; Liu, H. *et al.* 2002. The emerald ash borer: a new exotic pest in North America. Newsletter of the Michigan Entomological Society. 47(3, 4): 1-5.

Illinois Department of Agriculture. 2006. Emerald ash borer confirmed in Illinois. http://www.agr.state.il.us/newsrels/r0613061.html. (15 June 2006).

Maryland Department of Agriculture. 2006. Emerald ash borer in Maryland. http://www.mda.state.md.us/plants-pests/eab/. (5 October).

McCullough, D.G.; Katovich, S.A. 2004. Emerald ash borer. NA-PR-02-04. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Area State and Private Forestry. 2 p.

McCullough, D.G.; Siegert, N.W. 2007. Potential production of emerald ash borer adults: tree, site and landscape-level applications. In: Gottschalk, K.W., ed. Proceedings, 17th U.S. Department of Agriculture interagency research forum on gypsy moth and other invasive species 2006. Gen. Tech. Rep. NRS-P-10. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station: 74-75.

Michigan Department of Agriculture. 2005. Emerald ash borer confirmed at Brimley State Park. Available at: http://www.michigan.gov/mda/0,1607,7-125-1572_3628-126024--,00.html. (10 January 2009).

Miles, P.D.; Brand, G.J.; Alerich, C.L. *et al.* 2001. The Forest Inventory and Analysis database description and user's 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.

Poland, T.M.; McCullough, D.G. 2006. Emerald ash borer: invasion of the urban forest and the threat to North America's ash resource. Journal of Forestry. 104(3): 118-124.

SAS Institute, Inc. 2004. SAS for Windows: SAS OnlineDoc®. Version 9.1.3. Cary, NC: SAS Institute, Inc.

Tepley, A.J.; Cohen, J.G.; Huberty, L. 2004. Natural community abstract for southern floodplain forest. Lansing, MI: Natural Features Inventory. 14 p.

U.S. Department of Agriculture (USDA) Forest Service. 2001. National core field guide. Vol. 1. Version 1.5. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station, Forest Inventory and Analysis. 247 p. http://www.ncrs.fs.fed.us/4801/regional-programs/Inventory/fieldcrews/manuals/. (15 June 2006).

The Distribution of Mercury in a Forest Floor Transect Across the Central United States

Charles H. (Hobie) Perry[1], Michael C. Amacher[2], William Cannon[3], Randall K. Kolka[4], and Laurel Woodruff[5]

Abstract.—Mercury (Hg) stored in soil organic matter may be released when the forest floor is consumed by fire. Our objective is to document the spatial distribution of forest floor Hg for a transect crossing the central United States. Samples collected by the Forest Service, U.S. Department of Agriculture's Forest Inventory and Analysis Soil Quality Indicator were tested for Hg using cold-vapor atomic absorption. We found patterns of Hg concentration that differ from earlier studies; the patterns of Hg concentration had a bimodal distribution with high values in both the northern Rocky Mountains and the Great Lake States. Future work will include an evaluation of the role of elevation and forest type on Hg storage.

Introduction

Soils can be sinks for atmospherically deposited mercury (Hg) (Grigal *et al.* 2000, Kolka *et al.* 2001). Soils act as sinks because the Hg that binds to organic matter in soils is not volatilized back to the atmosphere. Hg retention in soils increases with organic carbon content and climatic factors that influence microbial activity in soils (Fleck *et al.* 1999, Grigal *et al.* 2000, Kolka *et al.* 2001). Forest fires may release into the ecosystem Hg stored in organic matter (Kelly *et al.* 2006). As mercury emissions come under increasing scrutiny and regulation, the contribution of Hg from forest fires relative to other anthropogenic sources is unknown (Friedli *et al.* 2003, Turetsky *et al.* 2006).

The Forest Service, U.S. Department of Agriculture's Forest Inventory and Analysis (FIA) program produces data, information, and knowledge about the extent, condition, status, and trends of the Nation's forest resources across all land ownership categories (Smith 2002). Traditionally, FIA concentrated on the Nation's timber resources, but a change in focus was codified by the passage of the Agricultural Research, Extension and Education Reform Act of 1998, integrating FIA with the ground sampling components of the Forest Health Monitoring program. FIA collects samples from forested areas across the United States as part of its phase 3 sampling, and annual soils inventories are underway or completed in 44 of the 50 States (Alaska, Hawaii, Mississippi, Montana, New Mexico, and Oklahoma have yet to be sampled). The data is being used to facilitate State, national, and international assessments (Coulston *et al.* 2005, Smith 2002, Stolte *et al.* 2002).

Our objective is to document initial investigations into the spatial distribution of forest floor Hg for a transect crossing the central coterminous United States (fig. 1). We selected this region because atmospheric wet deposition rates and concentrations vary from west to east across the country (National Atmospheric Deposition Program 2007, Sweet and Prestbo 1999), providing gradients of Hg deposition and retention across the study area.

Methods

The collection of forest floor samples was accomplished as part of the standard FIA Phase 3 Soil Quality Indicator program. The forest floor is defined as "the entire thickness of organic material overlying the mineral soil, consisting of the litter and the duff (humus)," and field protocols include the measurement

[1] Research Soil Scientist, U.S. Department of Agriculture (USDA) Forest Service, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108. E-mail: charleshperry@fs.fed.us.

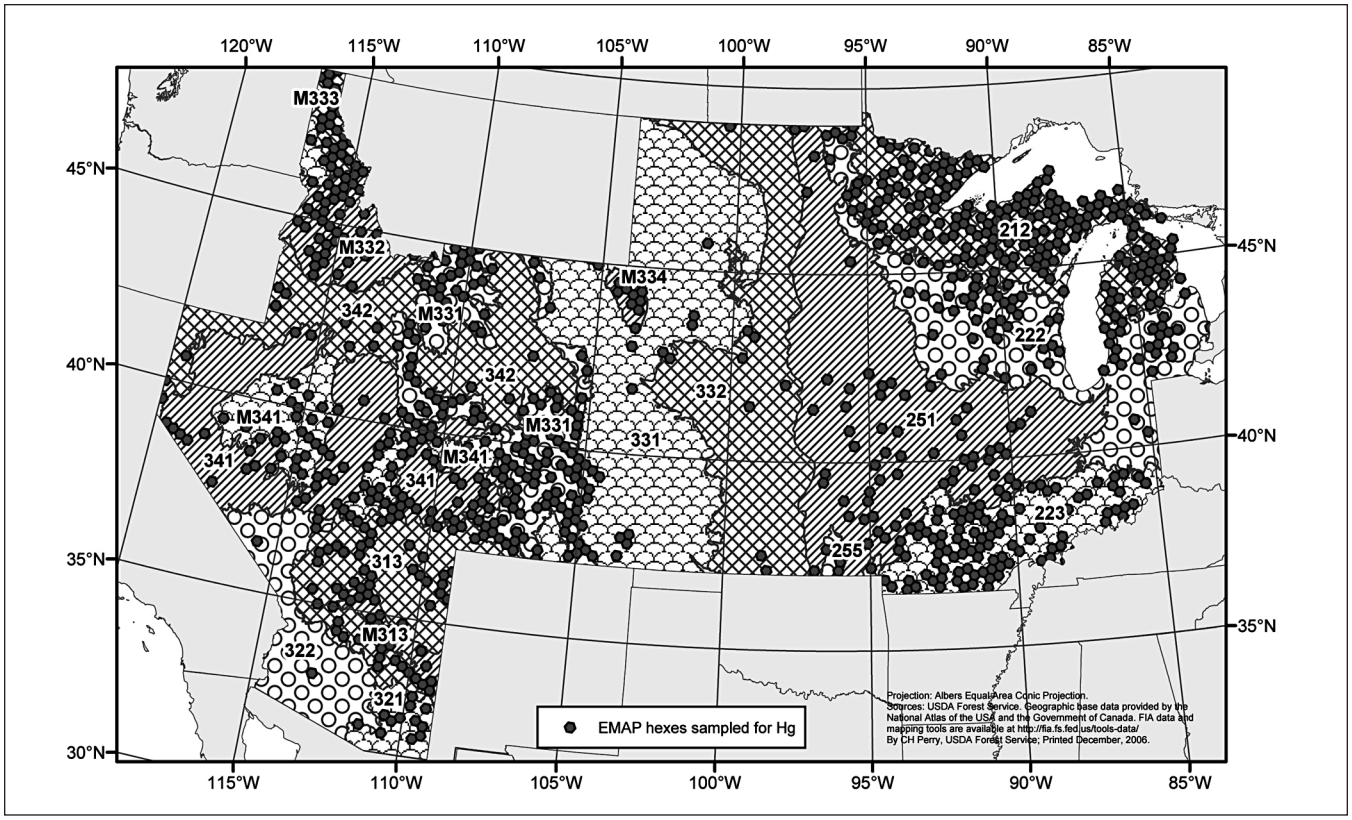
[2] Research Soil Scientist, USDA Forest Service, Rocky Mountain Research Station, 860 North 12th East, Logan, UT 84321. E-mail: mamacher@fs.fed.us.

[3] Geologist, U.S. Geological Survey (USGS), Eastern Mineral Resources, 954 National Center, Reston, VA 20192. E-mail: wcannon@usgs.gov.

[4] Project Leader, USDA Forest Service, Northern Research Station, 1831 Hwy. 169 East, Grand Rapids, MN 55744. E-mail: rkolka@fs.fed.us.

[5] Geologist, USGS, Eastern Mineral Resources, 2280 Woodale Drive, Mounds View, MN 55112. E-mail: woodruff@usgs.gov.

Figure 1.—*The distribution of soil samples tested for mercury. Patterned polygons and numbers refer to ecoprovinces of the region.*



of the thickness of the forest floor and the collection of the entire forest floor found within a sampling frame with a diameter of 30 cm (USDA Forest Service 2006). The samples were tested for a number of chemical and physical properties, not including Hg (Amacher *et al.* 2003).

We removed approximately 0.1 g of the sample for plots in our region of interest, and these were sent to two different laboratories for Hg analysis: the Forest Service Forestry Sciences Lab in Logan, UT, and XRAL Laboratories in Toronto, ON. Both laboratories used cold-vapor atomic absorption to measure the amount of Hg and calibrated their instruments against common Hg standards. We found good agreement between samples analyzed at both laboratories.

Observations of mercury concentrations were joined with the FIA Database (Alerich *et al.* 2006) to assign plot locations. These plots were assigned to ecoprovinces (table 1) and Environmental Monitoring and Assessment Program hexagons (Spence and White 1992, White *et al.* 1992) using a geographic

Table 1.—*Ecoprovinces sampled for forest floor mercury.*

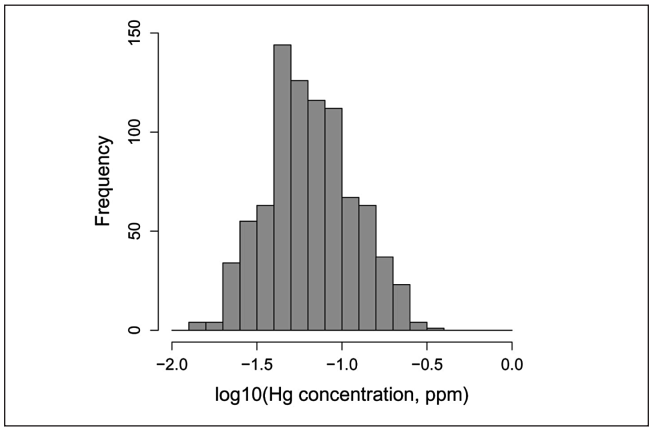
Label	Ecoprovince name	Number of samples
212	West Laurentian mixed forest	210
222	Midwest broadleaf forest	69
223	Central interior broadleaf forest	101
251	Prairie parkland (temperate)	54
313	Colorado Plateau semidesert	56
321	Chihuahuan semidesert	11
331	Great Plains—Palouse dry steppe	20
332	Great Plains steppe	11
341	Intermountain semidesert and desert	36
342	Intermountain semidesert	14
M313	Arizona-New Mexico mountains semidesert—open woodland— coniferous forest—alpine meadow	17
M334	Black Hills coniferous forest	11
M332	Middle Rocky Mountains steppe—coniferous forest—alpine meadow	43
M341	Nevada-Utah Mountains semidesert—coniferous forest—alpine meadow	55
M333	Northern Rocky Mountains steppe—coniferous forest—alpine meadow	25
M331	Southern Rocky Mountain steppe—open woodland—coniferous forest—alpine meadow	120

information system (fig. 1). Mean plot-level values of Hg concentration were tested against ecoprovince, latitude, and longitude using analysis of variance in R, a free software environment for statistical computing and graphics that is similar to the S language (Maindonald and Braun 2003).

Results and Discussion

We expected a strong regional component, with Hg concentrations increasing to the north and east, mimicking observations of mercury deposition (Nater and Grigal 1992, National Atmospheric Deposition Program 2007, Sweet and Prestbo 1999). The concentration of Hg in our samples was log-normally distributed (fig. 2) and ranged between 0.013 and 0.36 parts per million (ppm); the mean and median were 0.077 and 0.063 ppm, respectively. All statistical analyses were completed using the log-transformed concentration data.

Figure 2.—*Histogram of mercury concentrations (parts per million) for the study region.*



Mercury concentrations were highest in the northeastern part of our study region (ecoprovince 212), but high values were observed in the Rocky Mountains (ecoprovinces M332 and M333) as well (fig. 3). Ecoprovince was a significant predictor of Hg concentration in our analysis of variance model ($p < 0.001$) (fig. 4). As might be surmised from figure 3, latitude was a significant predictor ($p = 0.006$) but longitude was not ($p = 0.60$).

Figure 3.—*Spatial distribution of mean mercury concentrations (parts per million) for the study region. Numbers refer to ecoprovinces in the study region*

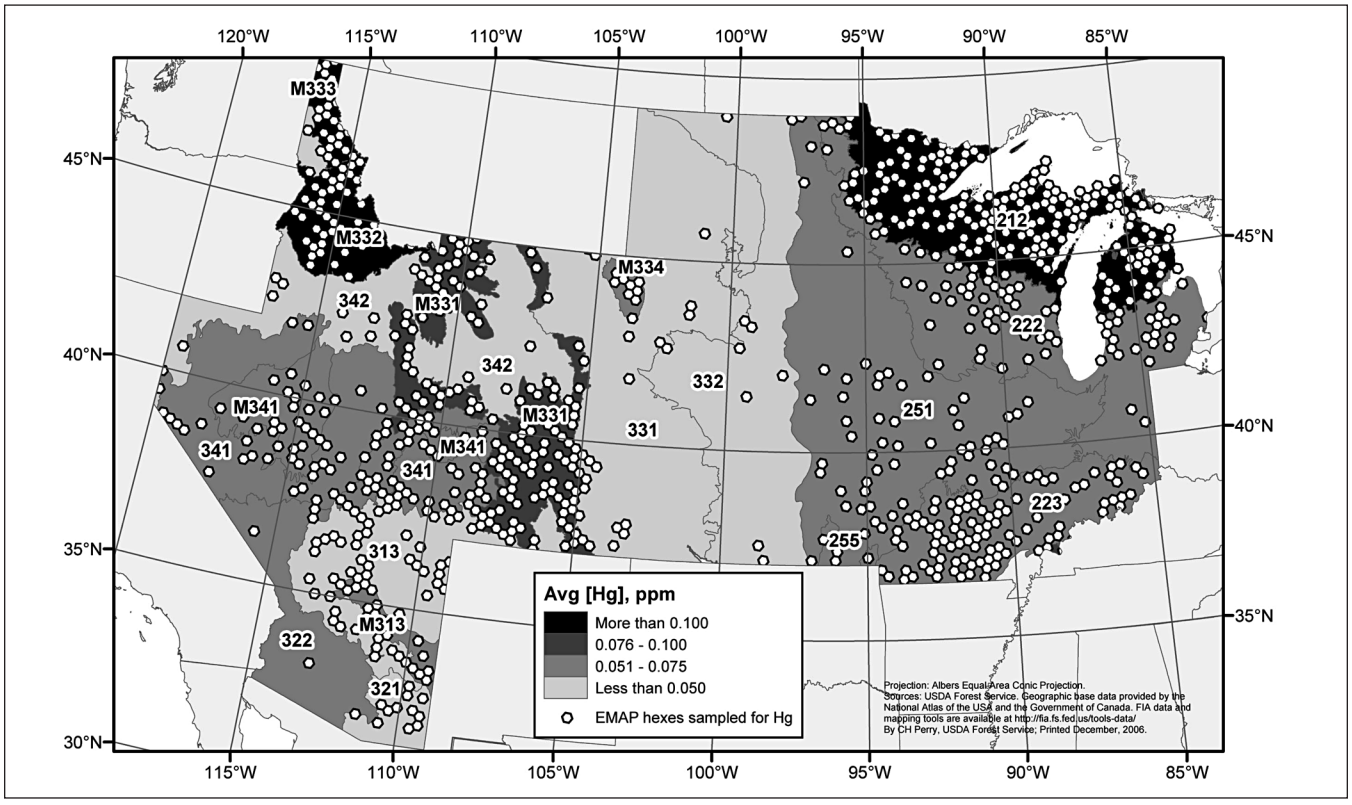
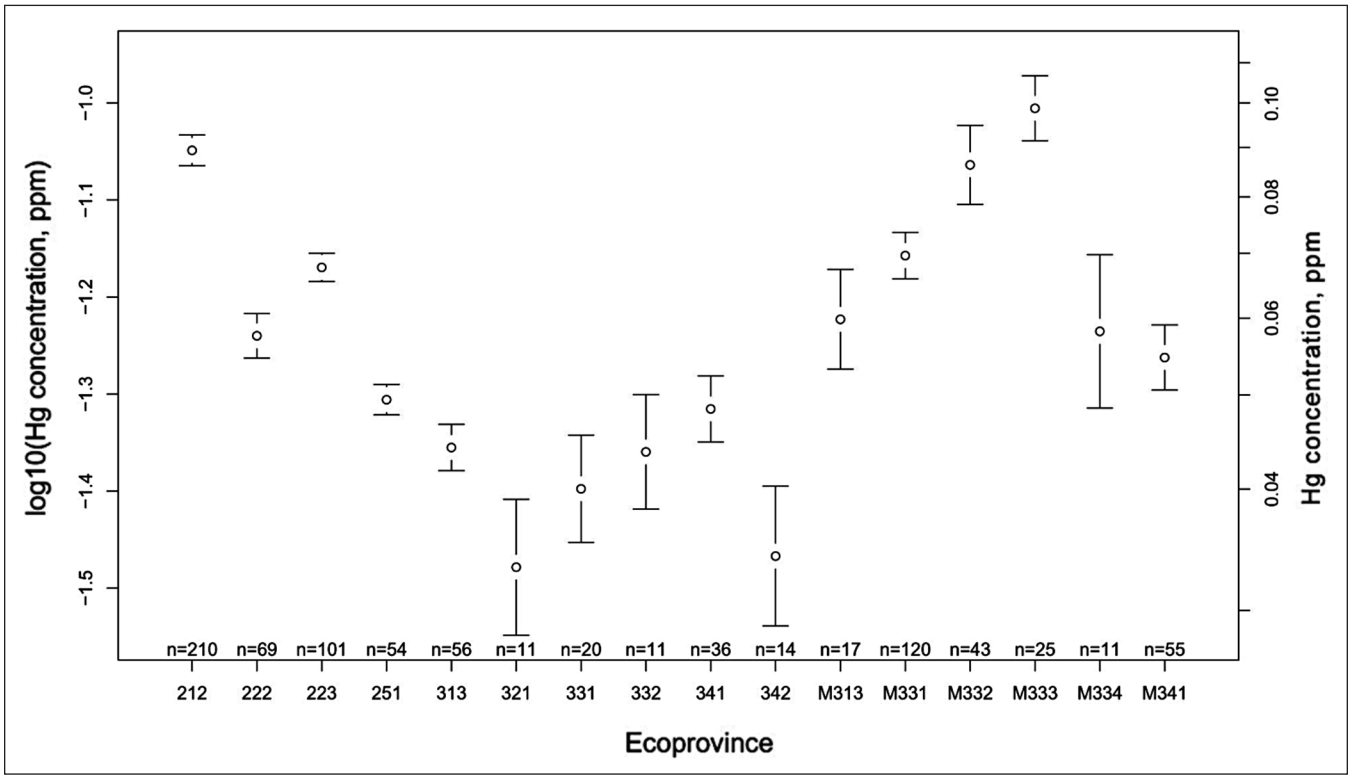


Figure 4.—*The mean concentration of mercury (parts per million) by ecoprovince. Ecoprovinces are mapped in figures 1 and 3. Error bars represent one standard error.*



The spatial trends in Hg concentration were not consistent across all ecoprovinces; significant interactions were found between ecoprovince and latitude ($p = 0.02$) as well as ecoprovince and longitude ($p = 0.003$). For example, the general trend was increasing Hg concentration with latitude, but concentrations increased more slowly in ecoprovince 251 and more quickly in ecoprovince M334. Also, the concentration of Hg generally increased from west to east, but concentrations actually declined moving east in ecoprovince M341 and M334. Because these interactions are largely observed in mountainous ecoprovinces, this evidence would suggest that elevation should be investigated further.

It is important to remind the reader that we are reporting concentration data; the total mass of the forest floor on each plot is required to determine the total mass of Hg stored in the forest floor. Carbon storage in the forest floor is comparable between the northern Rocky Mountains and the Lake States (Perry and Amacher, in press) indicating similar amounts of total forest floor mass (kg ha^{-1}). This suggests that the stored mass of Hg will follow the patterns of Hg concentrations in forest floor material across the study region.

Summary

This article provides initial results of a spatial model of forest floor Hg for a transect crossing the central United States, and we are very curious about the patterns emerging from our data. Mercury concentration in the forest floor ranged from 0.013 to 0.36 ppm across our study area. The elevated concentrations (and anticipated large mass) of Hg in the northern Rocky Mountains require further analysis as the spatial patterns do not completely mimic current patterns of Hg deposition (National Atmospheric Deposition Program 2007). Additionally, preliminary tests on subsets of the data suggest the important predictive role of forest-type groups. Each of these issues will be addressed in greater detail as the study progresses.

Acknowledgments

The authors acknowledge funding support from the Forest Service, U.S. Department of Agriculture State and Private Forestry and the Forest Health Monitoring program, Evaluation Monitoring Program. Additional in-kind support was provided by R.K. Kolka's research unit, RWU-NC-4351: Ecology & Management of Riparian/Aquatic Ecosystems.

Literature Cited

Alerich, C.L.; Klevgard, L.; Liff, C.; Miles, P.D.; Knight, B.; Conkling, B.L. 2006. The forest inventory and analysis database: database description and users guide. Version 2.1. http://www.ncrs2.fs.fed.us/4801/FIADB/fiadb_documentation/FIADB_v2-1_07_05_06.pdf. (1 December).

Amacher, M.C.; O'Neill, K.P.; Dresbach, R.; Palmer, C. 2003. Forest inventory and analysis manual of soil analysis methods. Unpublished report. On file with: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Forestry Sciences Lab, 860 North 12th East, Logan, UT 84321. 62 p.

Coulston, J.W.; Ambrose, M.J.; Riitters, K.H.; Conkling, B.L. 2005. Forest health monitoring: 2004 national technical report. Gen. Tech. Rep. SRS-90. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 81 p.

Fleck, J.A.; Grigal, D.F.; Nater, E.A. 1999. Mercury uptake by trees: an observational experiment. Water, Air, and Soil Pollution. 115(1-4): 513-523.

Friedli, H.R.; Radke, L.F.; Prescott, R.; Hobbs, P.V.; Sinha, P. 2003. Mercury emissions from the August 2001 wildfires in Washington State and an agricultural waste fire in Oregon and atmospheric mercury budget estimates. Global Biogeochemical Cycles. 17(2): 1039.

Grigal, D.F.; Kolka, R.K.; Fleck, J.A.; Nater, E.A. 2000. Mercury budget of an upland-peatland watershed. Biogeochemistry. 50(1): 95-109.

Kelly, E.N.; Schindler, D.W.; St. Louis, V.L.; Donald, D.B.; Vladicka, K.E. 2006. Forest fire increases mercury accumulation by fishes via food web restructuring and increased mercury inputs. Proceedings of the National Academy of Sciences. 103(51): 19380-19385.

Kolka, R.K.; Grigal, D.F.; Nater, E.A.; Verry, E.S. 2001. Hydrologic cycling of mercury and organic carbon in a forested upland-bog watershed. Soil Science Society of America Journal. 65(3): 897-905.

Maindonald, J.; Braun, J. 2003. Data analysis and graphics using R: an example-based approach. New York: Cambridge University Press. 362 p.

Nater, E.A.; Grigal, D.F. 1992. Regional trends in mercury distribution across the Great Lakes states, north central USA. Nature. 358: 139-141.

National Atmospheric Deposition Program. 2007. Mercury deposition network: a NADP network. http://nadp.sws.uiuc.edu/mdn/. (5 January).

Perry, C.H.; Amacher, M.C. [In press]. Soil carbon. In: Ambrose, M.J.; Conkling, B.L., eds. Forest Health Monitoring 2005 national technical report. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station.

Smith, W.B. 2002. Forest inventory and analysis: a national inventory and monitoring program. Environmental Pollution. 116: S233-S242.

Spence, M.; White, D. 1992. EMAP sampling grid technical report. Corvallis, OR: ManTech Environmental Technology, Inc.; U.S. Environmental Protection Agency, Environmental Research Laboratory. 64 p.

Stolte, K.; Conkling, B.L.; Campbell, S.; Gillespie, A. 2002. Forest health indicators: Forest Inventory and Analysis program. Washington, DC: U.S. Department of Agriculture, Forest Service. 24 p.

Sweet, C.W.; Prestbo, E. 1999. Wet deposition of mercury in the U.S. and Canada. In: Proceeedings, international specialty conference on mercury in the environment. Pittsburgh, PA: Air and Waste Management Association: 306-315.

Turetsky, M.R.; Harden, J.W.; Friedli, H.R. *et al.* 2006. Wildfires threaten mercury stocks in northern soils. Geophysical Research Letters. 33: L16403.

U.S. Department of Agriculture (USDA) Forest Service. 2006. Phase 3 field guide: soil measurements and sampling. http://www.fia.fs.fed.us/library/field-guides-methods-proc/. (1 December).

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(1): 5-21.

Dimensionality and the Sample Unit

Francis A. Roesch[1]

Abstract.—The sample unit and its implications for the Forest Service, U.S. Department of Agriculture's Forest Inventory and Analysis program are discussed in light of a generalized three-dimensional concept of continuous forest inventories. The concept views the sampled population as a spatial-temporal cube and the sample as a finite partitioning of the cube. The sample serves to cut the volume of the cube into a finite number of pieces like a three-dimensional jigsaw puzzle. Each puzzle piece is defined by the spatial-temporal selection volumes of observation sets on the individual trees existing in the forest during the period of interest. The concept is developed as a temporal extension of the alternative view of forest sampling offered in Roesch *et al. (*1993).

Introduction

Roesch *et al.* (1993) gave an alternative view of forest sampling that could be applied to all forest sampling schemes that select trees based on the location of a random point. They explain the idea as a jigsaw puzzle view of forest sampling. In this view, the sample units are the mutually exclusive sections of ground resulting from the overlapping selection areas of the individual trees in the forest. The population is the puzzle picture and it is partitioned into puzzle pieces that are mutually exclusive, exhaustive sample units that together define the sample frame. Each ground segment had a defined probability of selection proportional to its size, and the total of these probabilities over all segments is 1. In the case of point sampling, the size of each segment is determined by the basal areas and spatial distribution of the trees and the sampling angle chosen. In the case of plot sampling, the size of each segment is determined by the plot radius and the spatial distribution of the trees. Thus, all schemes can be thought of in a common fashion, differing only in the method used to cut the puzzle into pieces. After the puzzle is segmented, the segments are selected with probability proportional to their size, and each segment contains a set of attributes. Roesch *et al.* (1993) also showed the theory applicable to remeasured samples for two specific points in time. Since the paper's publication, the Forest Service Forest Inventory and Analysis program has adopted new sampling methods that require the addition of a temporal dimension to the jigsaw puzzle view. A formal description of that extension is the focus of current research by the author. That description involves the extension of the visualization to include time as a third dimension of the puzzle. The addition of time (to the puzzle) is required when the set of observation times is randomly determined. This addition results in a sampled population and sampling frame that are three dimensional.

Theory of the Three-Dimensional Sample

Assume a three-part process: first, a random point is located in two-dimensional space; second, a set of observation times is determined; and third, at each selected time, a cluster of trees near the point is selected for measurement by some rule. The two most common temporally specific rules for selecting the clusters of trees (i.e., two dimensional) are known as (circular, fixed-area) plot sampling and (horizontal) point sampling. In this article, we will concentrate on point sampling, in which a tree i, with radius of the cross-section r_i, is selected with probability proportional to tree basal area, πr_i^2. In considering continuous forest inventories, if a tree is selected at sample time t_1 from a permanent point and it lives until sample time t_1+k, most forest sampling rules would result in it also being measured at time t_1+k. Therefore, the probability of selection of the previously measured tree at time t_1+k is actually equal to 1 and not independent of the time t_1 sample. Successively applying the two-dimensional view is erroneous because new *independent* samples are not taken each time. One could argue that the population is an infinite set of points, with a point

[1] Mathematical Statistician, U.S. Department of Agriculture, Forest Service, Southern Research Station, 200 W.T. Weaver Blvd., Asheville, NC 28804–3454. E-mail: FRoesch@fs.fed.us.

being the sample unit, and therefore the same sample is merely reobserved. Alternatively, one could argue that because the set of trees occurring over the area of interest throughout the period of interest is the biological population of interest, our sampled population should be directly associated with it. This position requires knowledge of the probability of inclusion for the realized set of observations on each tree in the sample over the course of the period of interest. Because, when we also randomly sample time, potentially many sets of observations are realizable for each tree in the population, we must reconsider the sampled population and the sample unit. For example, table 1 shows the nine potential sets (excluding the null set) of observations on a single live tree over two cycles of a three-panel rotating panel design.

The sample unit is a three-dimensional puzzle piece created by partitioning the population volume with a solid of revolution for each tree, created by integrating πr_i^2 over time. The individual tree spatial-temporal volumes may be truncated on the sides when adjacent to an areal edge of the population and on the tops and bottoms by time limits of the population. We would divide the volume of a sample unit (in area x time units) by the volume of the population to determine the probability of selection for each unit. Although time is a continuous variable, we will treat it as discrete.

The three-dimensional view removes potential confusion related to inclusion probabilities. Note that inclusion probabilities can only differ between samples, not within samples. The population in this case is three-dimensional, with time being one of the dimensions; ergo, it is obvious that the inclusion probability cannot change over time unless a new sample is taken. Any change that would appear to change an inclusion probability in the two-dimensional view can be seen to actually define a separate subpopulation or sample unit by the three-dimensional definition. Subpopulations are defined by land area and/or by time.

If time is treated as discrete with a partition length of 1 year, the sample unit appears as a set of puzzle pieces created by partitioning the volume by overlapping sets of discs, one set for each tree in each panel. The disc size varies over time with tree size, enclosing an area known as the K-circle (Grosenbaugh and Stover 1957). The K-circle of tree *i* at time *j*, $K_{i,j}$, is an

Table 1.—*The marginal and conditional probability expressions for the nine possible sets of observations of a single live tree over two cycles (6 years) of an annual rotating panel design consisting of three consecutive panels. A nondecreasing probability occurs in live round trees that do not shrink.*

Year	1	2	3	4	5	6	Case	
$p(Y_i)$	p_1	p_2	p_3	p_4	p_5	p_6		
Panel	1	2	3	1	2	3		
$p(P_j)$	1/3	1/3	1/3	1/3	1/3	1/3		
Set							Unrestricted probability	Nondecreasing Probability
1	X						$p_1(1-p(Y_4\|Y_1))/3$	0
2	X			X			$p_1(p(Y_4\|Y_1))/3$	$(p_1)/3$
3				X			$((1-p_1)(p(Y_4\|Y_1)))/3$	$(p_4-p_1)/3$
4		X					$p_2(1-p(Y_5\|Y_2))/3$	0
5		X			X		$p_2(p(Y_5\|Y_2))/3$	$(p_2)/3$
6					X		$((1-p_2)(p(Y_5\|Y_2)))/3$	$(p_5-p_2)/3$
7			X				$p_3(1-p(Y_6\|Y_3))/3$	0
8			X			X	$p_3(p(Y_6\|Y_3))/3$	$(p_6)/3$
9						X	$((1-p_3)(p(Y_6\|Y_3)))/3$	$(p_6-p_3)/3$
	$(p_1)/3$	$(p_2)/3$	$(p_3)/3$	$(p_4)/3$	$(p_5)/3$	$(p_6)/3$		
		$(p_1 + p_2 + p_3) / 3$			$(p_4 + p_5 + p_6) / 3$			

imaginary circle, centered at tree center, with radius $\alpha r_{i,j}$. Assume that N trees are present, with labels 1,2,...,N, associated with the population of temporal length P starting in year 1. The selection area for tree *i* at time *j*, of size $a_{i,j}$ (in acres), is the portion of tree *i*'s K-circle that is within the population at time *j* and is the area from within which a random point will select the associated observations of tree attributes for the sample. Index $a_{i,j}$ annually, assuming $a_{i,j}$ remains constant for an entire year. Collect the $a_{i,j}$ into the matrix:

$$\mathbf{A} = \begin{bmatrix} a_{1,P} & a_{1,P-1} & \cdots & a_{1,1} \\ a_{2,P} & a_{2,P-1} & & \\ \vdots & & \ddots & \vdots \\ & & & \\ a_{N,P} & & \vdots & a_{N,1} \end{bmatrix}$$

We wish to estimate change over a defined area (**A**) and temporal period, and we'll assume a continuous forest inventory using a rotating panel design and leave the simplification to a single panel (or nonpaneled design) to the reader. An example of a rotating panel design is in use by the Forest Service's Forest Inventory and Analysis units (e.g., see Roesch and Reams 1999). Designs of this type consist of *g* mutually exclusive temporal panels. One panel per year is measured for *g* consecutive years, after which the panel measurement sequence reinitiates. Assume that the continuous inventory consists of n_c cycles. A cycle is a complete set of measurements on all panels, so $P = n_c g$ years. That is, if panel 1 is measured in year *t*, it will also be measured in years $t+g$, $t+2g$, etc., until year $t + (n_c - 1)g$. Panel 2 would then be measured in years $t+1$, $t+1+g$, $t+1+2g$, etc. Set *t* to 1 and reindex **A** as follows:

$$\mathbf{A} = \begin{bmatrix} a_{1,n_c g} & a_{1,(g-1)+(n_c-1)g} & \cdots & a_{1,1} \\ a_{2,n_c g} & a_{2,(g-1)+(n_c-1)g} & & \\ \vdots & & \ddots & \vdots \\ & & & \\ a_{N,n_c g} & & \vdots & a_{N,1} \end{bmatrix}$$

Without loss of generality, assume that a tree's assignment to temporal panel *p* is random, with probability equal to $1/g$.

Under this assumption, the unconditional probabilities of observation owing to a random point in three-dimensional spatial-temporal volume can be represented by the following matrix:

$$\mathbf{V} = g^{-1}A^{-1}\mathbf{A} = \begin{bmatrix} v_{1,n_c g} & v_{1,(g-1)+(n_c-1)g} & \cdots & v_{1,1} \\ v_{2,n_c g} & v_{2,(g-1)+(n_c-1)g} & & \\ \vdots & & \ddots & \vdots \\ & & & \\ v_{N,n_c g} & & \vdots & v_{N,1} \end{bmatrix}$$

Potentially n_c sets of observations are present on each tree for each panel. The probability of selection, by a random point in three-dimensional space, for a specific set, *s*, of observations (indexed by time of observation *o*) on tree *i* (π_{is}) is equal to the intersection of the selection volumes (for each time in the set) divided by the population (areal-temporal) volume, or

$$\pi_{is} = \frac{\bigcap_{o=1}^{n_s} v_{i,o\in s}}{V^T}$$, where n_s is the number of observations in the set. The joint probability of selection, by a random point in three-dimensional space, for a set, s_1, of observations (indexed by time of observation *o*) on tree *i*, and a set s_2 on tree *j* in panel *p* is

$$\pi_{is_1 j s_2} = \frac{\bigcap_{i,j}\left(\bigcap_{o=1}^{n_{s_1}} v_{i,o\in s_1}, \bigcap_{o=1}^{n_{s_2}} v_{j,o\in s_2}\right)}{V^T}.$$

When we consider all the possible intersections of all observation sets for all trees associated with the population, we have fully defined all sample units and the sample frame. That is, we have carved the spatial-temporal volume (or statue) into chunks that are selected with probability proportional to their size. Each chunk is associated with a unique set of observations on trees in the forest over time.

Assuming that tree centers do not move over time, then the intersection is equal to the smallest volume in the set,

$$\pi_{is} = \frac{\min(v_{i,o\in S})}{V^T}.$$

If we further assume the selection areas do not shrink, then the intersection is equal to the first volume in the set,

$$\pi_{is} = \frac{v_{i,\min(o)\in S}}{V^T}$$

The probability of having made any observation from panel p on tree i is $\pi_{ip} = \frac{\bigcup_{S=1}^{n_S} v_{ips}}{V^T} = \max_{S=1}^{n_s}(\pi_{ips})$, where v_{ips} is the selection volume for observation set s, of panel p for tree i, and n_S is the number of unique sets of observations on tree i in panel p.

Suppose that we randomly drop a point on the surface of a forest and use any rule to observe tree attributes over time. If the function is temporally dependent, then one must integrate area over time to determine a probability of inclusion. The probability of observing the attributes associated with tree i in year o is

$\pi_{o(i)} = \sum_{j\in S} \pi_{o(i)s} Z_{o(i),s}$, where

$$Z_{o(i)s} = \begin{cases} 1 & \text{if segment } s \text{ partitions the selection volume intersecting} \\ & \quad \text{observation year } o \text{ for tree } i \text{ and} \\ 0 & \text{otherwise} \end{cases}$$

and

$\pi_{0(i)s}$ is the probability that observation in year o for sample tree i was selected from the particular population segment s, corresponding to observation set S.

Therefore, $\pi_{o(i)s} = \left(\frac{\pi_{is}}{\pi_{o(i)}}\right)$. The sum over s of $\pi_{0(i)s}$ is 1.

Let y_{io} be the value of an attribute of interest for tree i at time o and $Y_o = \sum_{i=1}^{N} y_{io}$ be the total value of interest at time o across all trees.

We can now write a temporally specific observation for each segment as a sum of weighted tree values:

$$\bar{y}_{o,s} = \sum_{i=1}^{N} \pi_{o(i)s} y_{io} \qquad (1)$$

Now suppose that we randomly drop m points into the population volume with the same assumptions as those mentioned previously in the text and we are sampling with replacement. An unbiased estimator of the temporally specific total value of interest for a sample selected with probability proportional to size, V_s, is

$$\hat{Y}_o = \frac{V^T}{m} \sum_{s=1}^{m} \frac{\bar{y}_{o,s}}{V_s} = \frac{V^T}{m} \sum_{s=1}^{M} \frac{\bar{y}_{o,s}}{V_s} W_s \qquad (2)$$

where

$V^T = \sum_{s=1}^{M} V_s$, the total areal-temporal volume of the population,

m = the number of sample points,

M = the number of volume segments, and

W_s = the number of times the unit s appears in the sample.

Note that W_s is an integer between 0 and m, inclusive. V_s and $\bar{y}_{o,s}$ are fixed and W_s is random. From the development above, recall that o ranges from 1 to $n_c g$ years, so it is of interest to estimate the vector $\mathbf{y} = Y_1, Y_2, ..., Y_{n_c g-1}, Y_{n_c g}$ with a vector. $\hat{\mathbf{y}} = \hat{Y}_1, \hat{Y}_2, ..., \hat{Y}_{n_c g-1}, \hat{Y}_{n_c g}$ Focusing on the elements of the vectors, we define

$Y_o^* = \sum_{s=1}^{M} \bar{y}_{o,s}$ as the total time o value of interest across all segments.

Similar to Roesch *et al.* (1993) $\hat{Y}_o$ is easily shown to be unbiased for Y_0. By definition, the variance of $\hat{Y}_o$ is

$$Var(\hat{Y}_o) = \left(\frac{1}{m V^T}\right) \sum_{s=1}^{M} V_j \left(\frac{V^T \bar{y}_{o,s}}{V_s} - Y_o\right)^2. \qquad (3)$$

The sample estimate of the variance is then represented by the following equation (Cochran 1977):

$$\text{var}(\hat{Y}_o) = \frac{1}{m(m-1)} \sum_{s=1}^{m} \left(\frac{V^T \bar{y}_{o,s}}{V_s} - \hat{Y}_o\right)^2. \qquad (4)$$

Expanding equation 2 to include the definition of $\overline{y}_{o,s}$ and subsequent rearrangement gives

$$\begin{aligned}
\hat{Y}_o &= \frac{V^T}{m} \sum_{s=1}^{M} \frac{\overline{y}_{o,s}}{V_s} W_s \\
&= \frac{V^T}{m} \sum_{s=1}^{M} \frac{\sum_{i=1}^{N} \pi_{o(i)s} y_{io}}{V_s} W_s \\
&= \frac{V^T}{m} \sum_{s=1}^{M} \sum_{i=1}^{N} \frac{y_{io} Z_{o(i)s} W_s}{v_{o(i)}} \\
&= \frac{V^T}{m} \sum_{i=1}^{N} \frac{y_{io}}{v_{o(i)}} \sum_{s=1}^{M} Z_{o(i)s} W_s \\
&= \frac{V^T}{m} \sum_{i=1}^{N} \frac{y_{io}}{v_{o(i)}} w_{o(i)}
\end{aligned} \tag{5}$$

where $w_{0(i)}$ equals the number of times tree i is selected for observation at time o. The final expression in equation 5 is the three-dimensional point sample estimator.

Conclusion

This article discusses a unified discrete, three-dimensional population sampling theory for continuous forest inventories. It shows that, regardless of the method used to determine the observation points in space or time (e.g., remeasured plot sampling or point sampling, with or without temporal panels), all schemes can be thought of as carving the population volume into pieces and selecting the pieces with probability proportional to their size. The general development in equations 1 through 5 can be used for any specific type of areal-temporal forest sampling. This alternative view of continuous forest inventories may be useful in many ways. Most notably, it can contribute to database design by highlighting the minimally necessary information unit.

Literature Cited

Cochran, W.G. 1977. Sampling techniques. 3rd ed. New York: John Wiley. 444 p.

Grosenbaugh, L.R.; Stover, W.S. 1957. Point-sampling compared with plot-sampling in southeast Texas. Forest Science. 3: 2-14.

Roesch, F.A.; Reams, G.A. 1999. Analytical alternatives for an annual inventory system. Journal of Forestry. 97(12): 33-37.

Roesch, F.A., Jr.; Green, E.J.; Scott, C.T. 1993. An alternative view of forest sampling. Survey Methodology. 19(2): 199-204.

Assessment of Black Ash *(Fraxinus nigra)* Decline in Minnesota

Kathleen Ward[1], Michael Ostry[2], Robert Venette[3], Brian Palik[4], Mark Hansen[5], and Mark Hatfield[6]

Abstract.—Black ash *(Fraxinus nigra)* is an important component of wetland forests throughout the Upper Midwest and northeastern United States and is highly valued for paneling, furniture, and basketry. Decline of black ash has been noted with increasing frequency, although no detailed studies of the pattern of decline across the region have been done. From analyses of Forest Health Monitoring aerial sketchmapping data, an association was found between dieback and decline of black ash and proximity to city, county, and State roads. In addition, relationships between growth and mortality levels of black ash and climatic, edaphic, and physiographic factors were found through analyses of U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis (FIA) field plot data collected in Minnesota between 1977 and 2005. FIA data were limited, however, in revealing factors that could have caused the decline, such as damage from biotic and abiotic agents.

Introduction

Black ash *(Fraxinus nigra)* is present throughout the Upper Midwest and northeastern United States and is often found in lowland or swamp hardwoods forest types (Erdmann *et al.* 1987). Black ash also grows on more well drained and mesic sites. Black ash seed is an important food for game birds, song birds, and small mammals, and the twigs and foliage are used by white-tailed deer and moose. Black ash wood has limited commercial uses; it is highly valued for paneling, furniture, and specialty products. In addition, black ash wood is ideal for Native American basketry because it is strongly ring porous and the wood can be easily separated into basket splints (Benedict 2001). In recent years, the availability of quality basket trees has diminished because of black ash decline (Benedict 2001).

Black ash decline and relatively high levels of tree mortality have been observed throughout the range of black ash in recent years and at times throughout past decades (Croxton 1966, Livingston *et al.* 1995, USDA Forest Service 2004). For example, in 2004, more than 27,000 acres of black ash trees were reported to be affected by dieback and decline in Minnesota (Northeastern Area State and Private Forestry 2005). Declining trees typically exhibit sparse crowns, twig dieback, epicormic sprouting, and slow growth. The cause of black ash decline is unknown but was thought to be related to past drought conditions (Livingston *et al.* 1995), subfreezing winter temperatures with little snow cover, or late spring frosts (USDA Forest Service 2004). Black ash is a shallow-rooted species and, as such, is susceptible to the effects of varying water table levels and winter freeze-thaw injury. We identified several other hypotheses for black ash decline—advanced stand age, damage from biotic agents, hydrologic changes from road development, and road salt runoff.

The objectives of this study were to use Forest Service forest inventory and analysis (FIA) data and Forest Health Monitoring (FHM) data to assess the pattern and extent of black ash decline in Minnesota and to relate decline occurrence and variation to mapped landscape-scale climatic, physiographic, and edaphic data.

[1] Formerly Forester, currently Plant Health Safeguarding Specialist, U.S. Department of Agriculture (USDA), Northern Research Station, St. Paul, MN 55108. E-mail: kathleen.ward@aphis.usda.gov.

[2] Research Plant Pathologist, USDA, Northern Research Station, St. Paul, MN 55108. E-mail: mostry@fs.fed.us.

[3] Research Biologist, USDA, Plant Protection and Quarantine, Olney, TX 76374. E-mail: rvenette@fs.fed.us.

[4] Research Ecologist, USDA, Northern Research Station, Grand Rapids, MN 55744. E-mail: bpalik@fs.fed.us

[5] Research Forester, USDA Forest Service, Northern Research Station, St. Paul, MN 55108. E-mail: mhansen01@fs.fed.us.

[6] Forester, USDA Forest Service, Northern Research Station, St. Paul, MN 55108. E-mail: mahatfield@fs.fed.us.

Methods

Aerial Survey Data

Aerial survey data collected in Minnesota in 2004 were obtained from the FHM Aerial Survey Results Viewer (Northeastern Area State and Private Forestry 2005). Dieback and decline polygons in the black ash cover type were joined to three Minnesota Department of Transportation roads layers—major interstates and trunk highways, county and State roads, and city streets—using ArcGIS. A field containing distance values from the dieback/decline polygons to roads was added to the data tables during the joins. Random points equal in numbers to the dieback/decline polygons were generated within the black ash cover type previously determined by the Minnesota Gap Analysis Program. In a manner similar to the dieback/decline polygons, the random points were spatially joined to each of the three roads layers and distance from roads was calculated for each. Numbers of dieback/decline polygons and numbers of random points at distances from roads at 150-ft intervals between 0 and 9,990 ft were analyzed using contingency tables and the likelihood ratio statistic.

Public Inventory Data

TREE, PLOT, and CONDITION data tables were obtained for four Minnesota FIA inventory cycles initiated between 1977 and 2005—cycles 4, 5, 12, and 13 (FIA 2006). For cycles 12 and 13, black ash tree data from the tables were filtered to include only plots that were coded as black ash forest type. Black ash tree data from cycles 4 and 5 were filtered to include only plots coded as the elm/ash/cottonwood group because the codes were not specific to black ash forest type. Mean mortality, diameter at breast height (d.b.h.), and dead tree-to-live-tree ratios for each of the four FIA inventory cycles were derived from the TREE tables, averaged on a plot basis, and used as response variables. FIA variables, ASPECT, OWNCD, OWNGRPCD, PHYSCLCD, RDDISTCD (2003 and 2005 cycles only), SITECLCD, SLOPE, TRTOPCD, and WATERCD (2003 and 2005 cycles only), were used as potential predictors of black ash growth and mortality (see table 1 for variable definitions). The data were analyzed with a generalized linear model and Fisher's least significance difference. In addition, numbers of

Table 1.—*Statistically significant associations of public forest inventory and analysis data in four inventory cycles and black ash (a) dead-to-live tree ratio and (b) diameter at breast height (d.b.h.).*

FIA Variable	Cycle							
	1977 era	1990 era	2003	2005	1977 era	1990 era	2003	2005
	(a) Dead-to-live tree ratio				(b) d.b.h.			
OWNCD[a]				X[b]	X			
OWNGRPCD						X	X	
PHYSCLCD			X	X			X	
SITECLCD		X			X		X	
TRTOPCD		X	X	X		X	X	X
SLOPE						X		
ASPECT								
RDDISTCD	NA[c]	NA			NA	NA		
WATERCD	NA	NA			NA	NA		

[a] *OWNCD = owner (National Forest System; U.S. Fish & Wildlife Service; other Federal, State, county/municipal; private).*
OWNGRPCD = owner group (Forest Service; other Federal, State, and county/municipal; private).
PHYSCLCD = physiographic class (1977- and 1990-era cycles [xeric, mesic]; 2003 and 2005 cycles [dry ridge tops, deep sands, flatwoods, rolling uplands, moist slopes, narrow flood plains, broad flood plains, other mesic, swamps/bogs, small drains, wet bays, beaver ponds, boggy bays, other hydric]).
SITECLCD = site productivity class (cubic feet/acre/year) (165 to 224, 120 to 164, 85 to 119, 50 to 84, 20 to 49, 0 to 19).
TRTOPCD = physical opportunity to improve stand conditions by applying management practices (regeneration without site preparation, regeneration with site preparation, stand conversion [e.g., undesirable species], thin seedlings and saplings, thin poletimber, other stocking control [e.g., remove undesirable material], other intermediate treatments [e.g., fertilize, prune], clearcut harvest, partial cut harvest, salvage harvest, no treatment).
SLOPE = slope angle (percent).
ASPECT = slope direction (degree).
RDDISTCD = horizontal distance to improved road (feet) (2003 and 2005 cycles [less than or equal to 100, 101 to 300, 301 to 500, 501 to 1000, 1,001 to 2,640, 2,641 to 5,280, 5,281 to 15,840, 15,841 to 26,400, greater than 26,400]).
WATERCD = water on plot (2003 and 2005 cycles [none, small permanent streams or ponds; deep swamps, bogs, marshes; temporary streams; flood zones; other temporary water]).
[b] *X = a statistically significant association ($P<0.05$).*
[c] *NA = not available.*

trees in damaging agent categories (AGENTCD) were summarized for all four cycles, and counts of tree damage types (DAMTYP1) were summed for the 2003 and 2005 cycles. Damage types (DAMTYP1) were not recorded in the 1977 and 1990 cycles. Plot coordinates for all public inventory data are adjusted to ensure that the FIA plot data cannot be linked to individual landowners, so the true coordinates for the public data were not available.

True Coordinate Inventory Data

Data collected in Minnesota from 1,605 black ash trees measured in the 1990-era cycle and remeasured in the 2003 cycle were accessed from the FIA Spatial Data Services center in St. Paul, MN. The data included the true plot coordinates which were spatially joined with several ancillary datasets—county boundaries (Minnesota DNR 2003a), ecological subsections (Minnesota DNR 1999), temperature and precipitation (PRISM Group 2006), STATSGO soils data (NRCS 2006), the National Wetlands Inventory (NWI) (Minnesota DNR 2003b), and the National Hydrography Dataset (NHD) (USGS 2005). Spatial relationships of black ash growth and mortality among State climate divisions, ecological subsections, and counties were analyzed using contingency tables, the likelihood ratio statistic, and StatXact software. SAS software and linear regression were used to determine relationships between growth and mortality variables of black ash and mean temperature, mean precipitation, and STATSGO soil characteristics.

Results

Aerial Survey Data

Black ash dieback/decline polygons were significantly closer to city streets ($P = 0.030$) and to county and State roads ($P < 0.001$) than were random black ash points. Distances to highways were similar between dieback/decline polygons and random points ($P = 0.341$).

Public Inventory Data, 1977 Era

Mean d.b.h. differed by SITECLCD (table 1) and was smallest in the lowest productivity class (20 to 49 cf/ac/yr); most declining and nondeclining trees (85 percent) were on sites in the lowest productivity class. Mean d.b.h. differed by OWNCD (table 1) and was significantly larger on private land than on State land ($P = 0.050$) (table 1). Neither mean d.b.h. nor dead-to-live tree ratio was associated with ASPECT, OWNGRPCD, PHYSCLCD, SITECLCD, SLOPE, or TRTOPCD. Damaging agents were found to affect 20 percent of the trees, and the most common damaging agents were disease (13 percent) and fire (3 percent).

Public Inventory Data, 1990 Era

Dead-to-live tree ratio differed by SITECLCD ($P < 0.001$) (table 1), and the largest mean ratio was on medium productivity sites (85 to 119 cf/ac/yr). Most declining and nondeclining trees (88 percent) were on the poorest sites (20 to 49 cf/ac/yr). Dead-to-live tree ratio ($P < 0.001$) and d.b.h. differed by TRTOPCD ($P < 0.001$) (table 1), and the smallest dead-to-live tree ratios and the smallest mean d.b.h. were on sites requiring thinning treatments. Mean d.b.h. was positively related to increasing SLOPE values ($P < 0.001$) (table 1). Mean d.b.h. also differed by OWNGRPCD ($P = 0.053$) (table 1) and was largest on private land and smallest on National Forest land. Neither mean d.b.h. nor dead-to-live tree ratio was associated with ASPECT, OWNCD, or PHYSCLCD. Damaging agents were recorded on 14 percent of the trees, and, similar to the 1977-era cycle, the most common damaging agents were disease (5 percent) and fire (3 percent).

Public Inventory Data, 2003

Dead-to-live tree ratio ($P < 0.001$) and mean d.b.h. ($P < 0.001$) differed by TRTTOPCD (table 1). The largest mean dead-to-live tree ratio was on sites requiring the greatest degrees of treatment (e.g., stand conversion, regeneration with site preparation), and the smallest mean d.b.h. was on sites requiring stand conversion or thinning. Dead-to-live tree ratio ($P < 0.001$) and mean d.b.h. ($P = 0.006$) also differed by PHYSCLCD (table 1). The largest mean dead-to-live tree ratios were in hydric classes, such as beaver ponds and swamps and bogs, and the smallest mean ratios were in mesic and xeric classes—moist slopes and dry ridge tops. Similarly, black ash on sites with beaver ponds had a smaller mean d.b.h. than black ash on dry ridge tops or moist slopes. Mean d.b.h. also differed by SITECLCD ($P = 0.005$) (table 1); the smallest mean d.b.h. and the least mean stand

ages were in the largest productivity class (120 to 164 cf/ac/yr). Most declining and nondeclining trees (84 percent) were on sites in the poorest productivity class (20 to 49 cf/ac/yr). Mean dead-to-live tree ratio differed by OWNGRPCD (table 1) and was significantly larger on Forest Service land than on State or local government or private land (P = 0.050). Neither mean d.b.h. nor dead-to-live tree ratio was associated with ASPECT, OWNCD, RDDISTCD, SLOPE, or WATERCD. The most common damaging agent on dead trees was unknown (47 percent), and the most frequent damage types were dead terminals (3 percent) and conks or decay (2 percent).

Public Inventory Data, 2005

Dead-to-live tree ratio differed by OWNCD (P = 0.016) (table 1); the largest mean dead-to-live tree ratio was on county- or municipal-owned land, and the smallest was on National Forest land. Dead-to-live tree ratio differed by PHYSCLCD (table 1) and was significantly greater on narrow flood plains (P = 0.050) than on any other physiographic class. In addition, mean d.b.h. differed by TRTOPCD (P < 0.001) (table 1), and black ash on sites requiring stand conversion or regeneration with site preparation had smaller mean d.b.h. than black ash on sites with other TRTOPCD codes. Dead-to-live tree ratio also differed by TRTOPCD (table 1) and was significantly greater (P = 0.050) in stands requiring site preparation for regeneration than in stands with adequate stocking levels. Most declining and nondeclining trees (78 percent) were on sites in one of the poorest productivity classes (20 to 49 cf/ac/yr). Neither mean d.b.h. nor dead-to-live tree ratio was associated with ASPECT, OWNGRPCD, SITECLCD, RDDISTCD, SLOPE, or WATERCD. Unknown (21 percent) and logging or human (21 percent) were the most common damaging agents on dead trees, and, similar to the 2003 cycle, the most frequent damage types were dead terminals (5 percent) and conks or decay (3 percent).

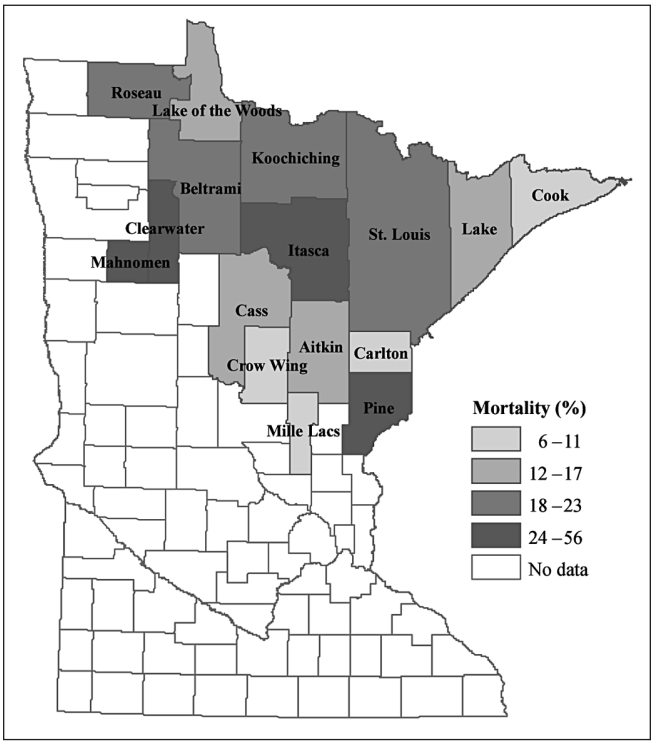
True Coordinate Inventory Data, 1990 Era and 2003

Black ash mortality increased by 18 percent between 1990 and 2003, and levels of mortality were spatially concentrated. Mortality differed among 16 counties (P < 0.001) (fig. 1) and was greatest in Mahnomen County (56 percent) and was least in Crow Wing and Mille Lacs Counties (6 percent and 7 percent, respectively). Mortality also differed among five Minnesota climate divisions in 1990 (P ≤ 0.001) but was similar among divisions in 2003 (P = 0.176). Mortality between 1990 and 2003 was largest in the Central (24 percent), Northwest (23 percent), and North Central (15 percent) Divisions. d.b.h. also differed among climate divisions in 1990 (P = 0.052) and in 2003 (P = 0.006), and the mean d.b.h. was largest in the Central Division in both time periods. In 1990, differences in black ash mortality (P = 0.02) existed among 20 ecological subsections, and the greatest mortality was in the Mille Lacs Uplands subsection. Mortality was similar among ecological subsections (P = 0.540) in 2003.

Between 1990 and 2003, change in d.b.h. was significantly greater in Upland than in palustrine/lacustrine NWI classes (P = 0.004). Similarly, between 1990 and 2003, change in d.b.h. increased as distance from a water-saturated NHD class (swamp) increased (P = 0.029).

Although significant relationships between black ash mortality and STATSGO soils variables, such as Pnddep (depth of surface water ponding on the soil), Clay (clay content), and Kffact (susceptibility of soil particles to detachment and movement by

Figure 1.—*Differences in levels of black ash tree mortality between 1990 and 2003 among 16 Minnesota counties.*



water), were observed, little of the variation was explained for tree mortality in the 1990 cycle ($R^2 = 0.10$) or the 2003 cycle ($R^2 = 0.02$). Between 1990 and 2003, change in black ash d.b.h. was significantly related to the STATSGO variable Bd (moist bulk soil density) and to mean temperature, but again little of the variation was explained ($R^2 = 0.03$).

Discussion

Black ash decline was associated with multiple, interacting factors. For example, black ash growth and survival were affected by site characteristics, such as physiographic features, including topography, soils, and wetness; weather conditions; land ownership; and proximity to city, State, and county roads. Generally, black ash growth and survival were poorer on flat, water-saturated sites of low quality. A positive relationship also existed between black ash decline and proximity to city, county, and State roads; several factors could contribute to this relationship. Road construction can alter the natural hydrologic flow through black ash stands and result in stagnant, standing water which can adversely impact tree growth and survival. Other factors include high levels of road deicing salt spray and runoff on land adjacent to roads in the winter. Road salt spray causes bud death and twig dieback in deciduous trees, and high levels of soil salt can damage leaves and reduce tree growth and vigor (Johnson and Sucoff 1999). In addition, road salt can decrease the cold hardiness of plants (Sucoff and Hong 1976). Vegetation near roadways can also be exposed to damaging pollutants from car and truck emissions.

Increases in black ash d.b.h. over time were negatively related to water-saturated classes in the NWI and NHD datasets, but no association occurred with black ash d.b.h. increases and the FIA plot water class (WATERCD). The difference may be due to disparity of scale because the WATERCD is on a point basis and both the NWI and NHD data are polygon based. In addition, the FIA water class definitions are different from definitions used by NWI and NHD. WATERCD is defined as a water body less than 1 acre in size or a stream less than 30 feet wide, whereas, definitions for water class in the NWI and NHD datasets do not include size limitations. Similarly, significant positive relationships existed between numbers of FHM dieback/decline polygons and proximity to city streets and county and State roads, but no associations existed between growth and mortality variables and the FIA distance to roads class. Again, a disparity of scale exists between the datasets. In addition, FHM dieback/decline variables and FIA growth and mortality variables are not interchangeable. Differences in mean d.b.h. among land ownership classes also could be due to varying geographic locations, management types, stand ages, legacies of land acquisition, or other unidentified causes.

The most common damage types recorded on black ash trees were dead terminals and decay. Disease and unknown causes were the most frequent damaging agents associated with black ash. It is difficult to meaningfully compare damaging agent codes and damage types from different FIA inventory cycles because methods have changed over time. For example, before 2000, damaging agents were coded for all trees. Beginning in 2000, however, the variable was collected for only dead and removed trees. In addition, data are collected in all seasons, and the ability to discriminate among types of tree damage and between damaged and undamaged trees differ depending on the season. Wetland species, such as black ash in Minnesota, are much easier to survey in the winter when the ground and surrounding water features are frozen and easier to traverse. It is very difficult, however, to differentiate damaged from undamaged trees in the winter because of the lack of foliage. In addition, decline results from complex interactions among multiple factors, so no one code or combination of codes could be expected to be sensitive enough to reliably distinguish declining from nondeclining trees.

In summary, FIA growth and mortality data proved valuable for discriminating among several factors that could be associated with black ash decline, such as physiographic and edaphic classes, but no FIA variables were found to separate declining trees from nondeclining trees. Many of the tree and plot characteristics used in this study could be easily derived from publicly available FIA data. The advantage of the true-coordinate FIA data is that it can be joined to various geospatial layers, and previously undetermined relationships may be revealed.

Acknowledgments

The authors thank Manfred Mielke and the Forest Health Monitoring program for partial funding of the study and Ebrahim Abdela for field observations (unpublished data).

Literature Cited

Benedict, M. 2001. Black ash: its use by Native Americans, and site factors affecting seedling abundance and ring growth in northern Minnesota. St. Paul, MN: University of Minnesota. 116 p. M.S. thesis.

Croxton, R.J. 1966. Detection and classification of ash dieback on large-scale color aerial photographs. U.S. Department of Agriculture, Forest Service. Res. Pap. PSW-35. Berkeley, CA: U.S. Department of Agriculture, Forest Service, Forest and Range Experiment Station. 13 p.

Erdmann, G.G.; Crow, T.R.; Peterson, R.M., Jr.; Wilson, C.D. 1987. Managing black ash in the Lake States. Gen. Tech. Rep. NC-115. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station. 10 p.

Forest Inventory and Analysis (FIA). 2006. FIA data mart: download files. http://www.ncrs2.fs.fed.us/FIADatamart/fiadatamart.aspx/. (14 December 2006).

Johnson, G.R.; Sucoff, E. 1999. Minimizing de-icing salt injury to trees. Extension Pub. FO-1413. Minneapolis/St. Paul, MN: University of Minnesota. 7 p.

Livingston, W.H.; Hager, A.; White, A.S. 1995. Drought associated with brown ash dieback in Maine. Phytopathology. 85: 1558.

Minnesota Department of Natural Resources (DNR). 1999. Ecological subsections of Minnesota. http://deli.dnr.state.mn.us/data_catalog.html/. (14 December 2006).

Minnesota DNR. 2003a. Minnesota county boundaries. http://deli.dnr.state.mn.us/data_catalog.html/. (14 December 2006).

Minnesota DNR. 2003b. National wetlands inventory polygons. http://deli.dnr.state.mn.us/data_catalog.html/. (14 December 2006).

Natural Resources Conservation Service (NRCS). 2006. U.S. general soil map (STATSGO). http://www.ncgc.nrcs.usda.gov/products/datasets/statsgo/. (14 December).

Northeastern Area State and Private Forestry. 2005. Aerial survey results viewer. http://www.na.fs.fed.us/ims/aerial/viewer.htm/. (14 December 2006).

PRISM Group. 2006. PRISM group. http://www.ocs.orst.edu/prism/index.phtml. (14 December 2006).

Sucoff, E.; Hong, S.G. 1976. Effect of NaCl on cold hardiness of *Malus* spp. and *Syringa vulgaris*. Canadian Journal of Botany. 54(24): 2816-2819.

U.S. Department of Agriculture (USDA) Forest Service. 2004. Lake States forest health watch. http://na.fs.fed.us/fhp/fhw/lsfhw/2004/dec_04/12_04.pdf. (14 December 2006).

U.S. Geological Survey (USGS). 2005. National hydrography dataset. http://nhd.usgs.gov. (14 December 2006).

Use of Forest Inventory and Analysis Grid-Based Animal Population Data To Develop an Index of Ecological Diversity

Patricia N. Manley[1], Kristian K. McIntyre[2], Matthew D. Schlesinger[2], Lori A. Campbell[2], Susan Merideth[3], and Dennis D. Murphy[3]

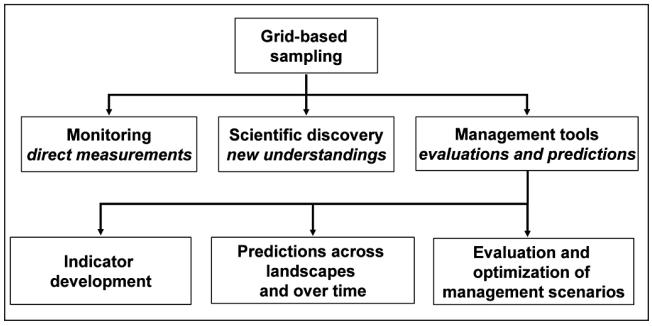
Abstract.—Data collected in association with the forest inventory and analysis (FIA) systematic grid has the potential to make multiple contributions to meeting land management information needs, including the development of indicators for application in management and monitoring programs. We derived bird, small mammal, and mammalian carnivore indexes of ecological diversity for the Lake Tahoe Basin using population data collected near FIA grid points, along with two additional data sets. The resulting index provided a robust measure of ecological diversity relative to existing landscape patterns of human development. We used the index to compare reference and existing conditions, and demonstrate its potential to meet National Forest System needs to evaluate progress toward desired ecological diversity conditions across landscapes and over time.

Introduction

Emphasis on large-scale environmental monitoring on public lands has increased dramatically over the past 10 years. Within the Forest Service, U.S. Department of Agriculture, the development of national protocols has been promoted and supported in an attempt to foster monitoring efforts that are implemented consistently and are compatible with national databases, both intended to improve the utility of environmental data (e.g., Holthausen *et al.* 2006; Manley, Van Horne, Roth *et al.* 2006, Woodbridge and Hargis 2006). The Forest Inventory and Analysis (FIA) program is designed to gather critical information on the status of forests in the United States (Roesch and Reams 1999). The systematic grid upon which FIA program data collection is based has the potential to serve as a valuable foundation for other complementary data collection efforts with grid-based data collection making multiple contributions to meeting public land management agency information needs, including the development of indicators, predictive models, and evaluation and assessment tools (fig. 1).

The National Forest Management Act (1976) and its implementation target the maintenance of species and ecological diversity in pursuit of sustainability. Indicators of species and ecological diversity are needed as measures of success in sustaining ecosystems on National Forest System (NFS) lands. To that end, ecological indicators are commonly employed in monitoring and assessing trends in environmental conditions, signaling rapid environmental changes, and diagnosing ecosystem health (Cairns *et al.* 1993). The national forests face many potential challenges, with loss of open space from urbanization being one

Figure 1.—*Potential information yields from grid-based population and habitat inventory and monitoring, including indicator development.*



[1] Research Wildlife Biologist, U.S. Department of Agriculture (USDA) Forest Service, Pacific Southwest Research Station, Sierra Nevada Research Center, 2121 2nd Street, Suite A-101, Davis, CA 95618. E-mail: pmanley@fs.fed.us.

[2] Research Wildlife Biologist, USDA Forest Service, Pacific Southwest Research Station, Sierra Nevada Research Center, 2121 2nd Street, Suite A-101, Davis, CA 95618.

[3] Doctoral Candidate and Graduate Program Director, Department of Ecology, Evolution, and Conservation Biology, University of Nevada, 1664 N. Virginia Street, Reno, NV 89557.

of the four primary threats, as recently identified by the Chief of the Forest Service (www.fs.fed.us/projects/four-threats). To meet management challenges on the nation's forest lands, measures that indicate the status of this threat are needed.

One tool for representing ecosystem diversity or integrity is to use indices that reflect multiple aspects of the biological community. A multimetric index of biotic integrity (IBI) was first developed by Karr (1981), using responses of fishes to indicate the biological condition or health of streams. Since its inception, the concept of the IBI has evolved and undergone modifications using various taxonomic indicators to assess diverse environments, from streams and wetlands, to terrestrial environments (e.g., Angermeier and Karr 1986, Bradford *et al.* 1998, Bryce *et al.* 2002, Karr and Kimberling 2003, Moyle and Marchetti 1999). The premise of indicators is based on the assumption that characteristics of the indicator (e.g., presence/absence, changes in the abundances, occurrences, survival) reflect changes or perturbations in the environment and associated species, processes, or functions (Noon *et al.* 1999). Many authors warn against the selection of only one or a few species as surrogate indicators of ecological condition (Feinsinger 2001, Landres *et al.* 1988, Morrison 1986, Noss 1990). IBI-type indices avoid this potential pitfall because they are typically based on a diversity of multispecies metrics, such as a diversity of trophic levels, species-group nesting guilds or foraging strategies, species richness and diversity measures, and tolerances of human disturbance.

Our objective was to demonstrate how indicators can be derived using basic field survey data collected, in part, in association with the FIA systematic grid, and then use those indicators to develop an index of ecological diversity that reflects changes in vertebrate biota in response to land development. Here we develop an index of terrestrial ecological diversity for the Lake Tahoe Basin patterned after the IBI approach. We then demonstrate how such an index of ecological diversity (IED) can be used to evaluate the status of terrestrial ecological diversity on a landscape subject to development and a diversity of other land uses.

Methods

Study Area

The Lake Tahoe Basin is a distinctive montane ecosystem located in the central Sierra Nevada Mountains. Lake Tahoe occupies nearly 10 percent of the 88,000-ha basin, with more than 80 percent of the remaining area consisting of national forest lands. The basin spans elevations from approximately 1900 m at Lake Tahoe to 3300 m at the surrounding peaks and encompasses four life zones (Holdridge 1967, Lugo *et al.* 1999): lower montane, upper montane, subalpine, and alpine (Manley *et al.* 2000). The basin's diversity has been altered over the past 150 years through a variety of human activities, most recently in the form of land development, including urban centers and more dispersed land conversion activities (Lindstrom *et al.* 2000). Currently human development and associated recreation impacts are greatest in the lower elevation zones (Manley *et al.* [2009]).

Sample Sites and Biotic Sampling

The Lake Tahoe Basin implemented a monitoring program based on the FIA grid (Roesch and Reams 1999) in 2002. The program entailed establishing monitoring points as per the national Multiple Species Inventory and Monitoring (MSM) protocol (Manley, Van Horne, Roth *et al.* 2006), with the addition of three more sample points randomly located within each FIA hexagon for a density of four points per 2400 ha. An increased density of points was needed to reach a sufficient sample size for monitoring in the relatively small Lake Tahoe Basin. In 2002–2004, 80 monitoring points distributed across all the hexagons were sampled for birds and small mammals (Roth *et al.* 2004), and in 2002–2005, a total of 61 monitoring points distributed across all the hexagons were sampled for mammalian carnivores. We combined these FIA-based MSM data with additional data sets for birds, small mammals, and mammalian carnivores from studies conducted in the Tahoe Basin. One data set (URB) included a survey conducted in 2003–2004 across land ownerships at lower elevations (less than 2200 m) in the basin (Manley, Murphy, Campbell *et al.* 2006), where 71 to 77 sites were sampled for birds, small mammals, and mammalian carnivores. Sites were randomly selected, stratified by degree of surrounding land development (in an area of 300 m radius),

which ranged from 0 to 80 percent. The final data set (RIP) was from a survey of streamside riparian ecosystems that was conducted in 1995–1996 across multiple land ownerships (Manley 2000). The data set consisted of 80 sites that were sampled for birds and small mammals, which included four stream reaches randomly selected from the headwaters to the mouth of 20 randomly selected watersheds. The three data sets were combined for a total of 305 sites. The three data sets overlapped spatially, and together they provided a broad representation of the diversity of ecological conditions in the basin.

We used point counts to characterize the bird community at 235 sites, and we detected 94 terrestrial bird species. Field methods were consistent among surveys as per standardized methods (Blondel *et al.* 1981, Ralph *et al.* 1993): distance between stations (200 m), timing (from mid-May to mid-August between 0530 and 1000 hours), duration (10 min), and number of visits ($n = 2$). The number and spatial arrangement of count stations varied minimally among data sets: MSM used seven count stations in a hexagonal configuration; URB used five stations in a cross-like configuration; and RIP used eight stations in a crosshatch configuration. We used frequency of detection among count stations (i.e., proportion of count stations per site with detections) as a representative measure of abundance to avoid any potential bias associated with the year of sampling or number of count stations.

We used live trapping to characterize the small mammal community at 231 sites, and we detected 22 small mammal species. Field methods were consistent among surveys: timing (June to September), duration of sampling (3 days and nights), bait (oats and seeds), trap check frequency (twice per day), distance between adjacent traps (15 m), primary trap type (extra-long Sherman traps; 7.5 x 9 x 22.5 cm), and marking (all captured individuals were marked). The number, configuration, and size of traps did vary: MSM used 104 extra-long traps in an open hexagonal configuration occupying approximately 10 ha; URB used 50 to 64 traps arrayed in a square grid, with every third trap being an extra large Sherman trap (10 x 11 x 37.5 cm); and RIP used 108 extra-long traps arrayed in a rectangular grid of 6 x 18 traps centered on the stream reach. We used the number of first captures per trap opportunity as a measure of relative abundance to avoid potential biases associated with recapture rates based on grid configuration. Trap opportunity was corrected (TO_c) for traps that were stolen, sprung, or otherwise rendered unable to capture animals,

$$TOc = (trapsset*trapchecks*trapdays) – (unavailabletraps / 2)$$

Mammalian carnivores were sampled at a total of 128 sites, and we detected 10 carnivore species. Sampling consisted of four or five enclosed sooted aluminum track plates (Barrett 1983, Fowler 1994, Zielinski and Kucera 1995) and two remote cameras (Zielinski and Kucera 1995). One track plate was positioned at the center of each sample unit and one camera was positioned 100 m from the center track plate at a random azimuth. For MSM, four track-plate stations were located 500 m from the center, one in each cardinal direction. The URB study had a similar configuration, with the exception of having only three track-plate stations placed 250 m from the center at 0, 120, and 240 degrees. In both efforts, one additional camera was paired and placed in the same fashion with one of the surrounding track plates. Track plates and cameras were baited with chicken and baby carrots, and a commercial scent was used as a lure. Track plates and cameras were checked every 2 days for a total of five visits. A species was determined to be present in a sample unit if any device recorded a detection event during the survey period. The response variable derived from these methods was species presence.

We grouped bird and small mammal species detected during sampling into species assemblages—21 and 8 groups, respectively—based on their ecological characteristics, in an effort to provide insights into the sensitivity of species that are known to be associated with specific forest resources (e.g., nesting and foraging substrates, primary diet during the breeding season). Assemblage or group membership was determined based on a variety of published sources and databases (as per Manley *et al.* 2004), and is documented (along with scientific names) in Manley and McIntyre (2006). Species detected incidentally or not directly associated with terrestrial ecosystems were not included in species groups. Species frequently belonged to more than one assemblage. For each group, we calculated richness and abundance. Abundance of species groups per site was calculated as the sum of detection frequencies across all species in the group.

Vegetation Community Classification

We first classified all 305 sites into vegetation community types using cluster analysis (CA). The CA was based on the proportion of the 50-ha (400-m radius) area occupied by each of five forest types (Jeffrey pine, mixed conifer, lodgepole pine, subalpine conifer, and aspen), shrubs, and meadow. The proportion of the area occupied by nonvegetative cover types (e.g., barren, water, development) and riparian vegetation was not included in the CA to ensure that the clusters reflected emergent properties of the primary environmental gradients in the Tahoe Basin. We chose a 50-ha area because it was the minimum area that encompassed most survey locations. The EVEG vegetation layer (USDA Forest Service 1991, 2001) derived from satellite imagery was the source for vegetation data. The California Wildlife Habitat Relationships vegetation classification scheme was used to represent vegetation types (CDFG 2002, Mayer and Laudenslayer 1988). We used ArcView Geographic Information System (GIS) (ESRI, Inc. 1995) to perform calculations. The CA was performed with flexible beta linkage (*beta* = -0.25) and Sørensen's (Bray-Curtis) distance in PC-ORD (Windows Version 4.17; McCune and Mefford 1999).

Once clusters were derived, we used principal components analysis (PCA) to examine the relationships of vegetation community types relative to principal environmental axes. For the PCA, we entered the original seven vegetation variables, plus four additional environmental variables: mean elevation, mean percent slope, mean precipitation, and percent riparian. Mean elevation and slope were derived from digital elevation models, while mean precipitation was derived from PRISM data (Daly 1995). The amount of riparian vegetation was represented by a map derived from infrared aerial photo interpretation of the entire Lake Tahoe Basin in the early 1990s. The cross-products matrix was created using Pearson correlation coefficients among all column variables to produce a "standardized" PCA, and a varimax rotation was used. Use of correlation coefficients serves to standardize column variables (McCune and Mefford 1999). Only factors with eigenvalues of more than 1.0 were retained and reported.

Land Development

We generated a spatially explicit GIS model of human development derived from land-use designations on county parcel maps, combined with roads and trails from Forest Service transportation maps (see Manley *et al.* [2009]). Development was defined as the removal of native vegetation. Average percent of parcels occupied by development was estimated based on land use designation (as identified on county parcel maps), roads, and trails. An average percent developed value was calculated for each of 90 land uses based on a sample of parcels in each land use. For each land use type, we estimated the amount of development within a 300 m radius using digital orthophoto quadrangles. The average percent developed was assigned to all parcels of each land use type. Roads and trails were assigned widths based on their respective average widths (Caltrans 2001), and added to the parcel development map to create the final development map. Sites were placed into one of three development classes: No/Low (less than or equal to 01.0 percent), Moderate (more than 1.0 to 10 percent), or High (more than 10 percent) development. For comparisons where the sample of sites was low at the upper end of the development gradient, only two classes were used: No/Low (less than or equal to 1.0 percent) or Mod/High (more than 1.0 percent) development.

Indicator Analysis

We used a combination of indicator species analysis (ISA) (PC-ORD for Windows Version 4.17; McCune and Mefford 1999), a technique originally developed by Dufrêne and Legendre (1997), and simple linear correlations to identify species and species groups that showed strong responses to development across all vegetation community types.

Indicator values (IV) produced from the ISA were based on a combination of species frequency of occurrence and abundance, and those values determine how strongly each species is associated with a particular group (vegetation or development). IV range from 0 (never present) to 100 (always present and exclusive to that group) (McCune and Mefford 1999). We tested the significance of each maximum IV using a Monte Carlo permutation procedure, and then compared the observed IV to randomly generated values (per Dufrêne and Legendre 1997). We used an alpha-level of $\alpha = 0.10$ to determine significance. Species and species groups that differed significantly between development classes were considered candidate indicators. We did not perform ISA for large

mammals, because only presence/absence data were available for these species.

In addition to ISA, linear relationships between species and development were examined. We identified all species and species groups that were significantly correlated ($P < 0.10$ based on Pearson's correlation coefficient) with development across all sites from all vegetation communities. We chose a minimum P value of < 0.10 for significance so that patterns of association would be more apparent. We then evaluated correlations of those species and species groups with development by vegetation community type. Despite apparent trends based on scatterplots, the limited occurrence of many species and the limited development in some vegetation communities (e.g., higher elevations) rendered most correlations insignificant; therefore, bivariate scatter plots also were consulted to evaluate relationships between species and development in each vegetation community.

Index Derivation and Ecological Diversity Assessment

Species and species groups were first evaluated for their potential as metrics in the IED. Species or species groups with strong and/or consistent relationships with human development among vegetation communities and analyses (i.e., ISA and correlations) were considered primary candidates. If their relationships were less consistent or weaker, then they were identified as secondary candidates. IED metrics selected were combinations of ecological groups and composites of individual species that exhibited strong shared relationships with development within and among vegetation communities. Individual species identified as having a strong positive or negative relationship with development were combined creating development-tolerant and development-intolerant species groups, respectively.

We wanted to have the same metrics for each vegetation community to enhance comparisons of conditions among them and reduce any potential for bias associated with sample size inequities among development classes. Based on the strength of the relationship with development, no more than 10 metrics were selected to represent each taxonomic group, consistent with the IBI approach (Karr and Chu 1995); however, the actual number was usually less than 10, as a function of the strength and number of candidate indicators.

Once metrics were selected, the values of each metric were rescaled to vary from 0 to 1. Generally, metric rescaling is done by dividing each metric value by the maximum value; however, in some cases, high values were considered outliers, effectively depressing the remaining range of values. The range and distribution of values for selected metrics were examined to select maximums to represent the primary range of values for each metric and to be used as the denominator in the rescaling procedure. For intolerant metrics, 1 indicates high diversity and 0 indicates low diversity. The values for tolerant metrics were subtracted from 1.0 to invert the values (1 indicates low diversity, 0 indicates high diversity). All rescaled metric values were then summed for each site, and the result was divided by the number of metrics used, and multiplied by 100 to obtain the final index with a potential maximum value of 100.

For birds and small mammals, we evaluated differences in richness and abundance relative to land development within and among the vegetation communities to identify potential metrics. Only species that were present at more than or equal to 10 percent of sites within two or more vegetation communities were evaluated for inclusion in tolerant or intolerant species groups. Species present at less than 10 percent of all sites were considered rare, and were included in the rare species group.

Only 10 mammalian carnivore species were detected across the three studies in the Lake Tahoe Basin. Large mammal data consisted of detection or nondetection events for each of those 10 species at each of the 128 sites, and ($n = 10$). As a result of the small species pool, the IED calculation was simplified relative to birds and small mammals. Species were classified as being either tolerant of land development or intolerant based on bivariate scatterplots. Given the small range of values associated with each metric, we assigned set values (0, 0.5, or 1) to each increment of increase in species richness. As for the birds and small mammals, the assigned values for the three metrics were added together and divided by 3, the theoretical maximum, to obtain a commensurate index value per site, which can range from 0 to 1.

It is likely that strong signals will be provided by the more abundant, lower trophic-level species (e.g., herbivorous small mammals, insectivorous birds), because of the greater range of values they may portray, but higher trophic-level species can be very sensitive to human disturbance because of their large area requirements. A diversity of trophic levels in the IED thus is likely to provide a stronger indication of ecological conditions than one alone. We used the subset of sites from which birds, small mammals, and large mammals all were sampled (n = 98 of the 305 sites) to generate a multi-taxonomic IED. We averaged the bird, small mammal, and large mammal IED values for each site to create the composite IED.

Finally, we used the composite IED to compare existing conditions in the basin to those of undeveloped (i.e., reference) conditions to provide a context for evaluating current terrestrial ecological diversity based on vertebrate biota. Reference sites were those with less than 5 percent development (n = 60 basinwide). Existing conditions were represented by a randomly selected subset of sites stratified in a basinwide grid to ensure equivalent representation. We then summarized the statistical distribution of IED values for existing conditions and compared them to reference conditions by vegetation community. We used t-tests to compare existing and reference conditions.

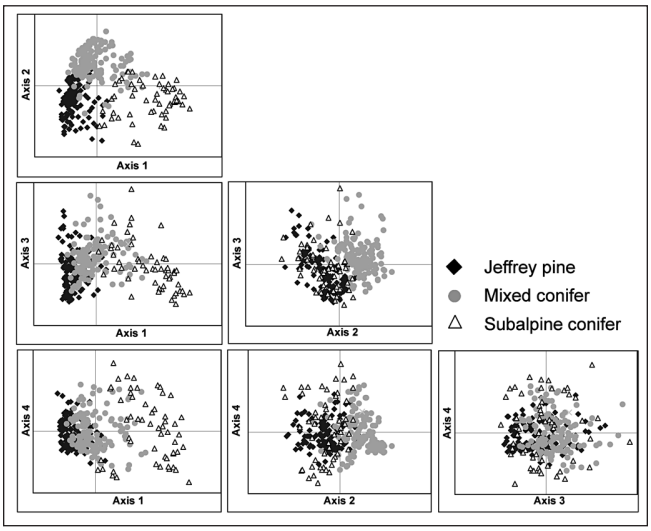
Results

Vegetation Community Types and Associated Land Development

The cluster analysis revealed three primary vegetation community types: Jeffrey pine dominated (n = 123 sites), mixed conifer (n = 127 sites), and subalpine conifer (n = 55 sites; fig. 2) communities. Four principal components had eigenvalues of more than 1.0 that together explained 70.0 percent of the variation in vegetation and environmental variables among sites. Factor 1 separated the Jeffrey pine community type (lowest elevation) from the subalpine conifer community type (highest elevation), and factor 2 separated the mixed conifer community type (midelevation) from the Jeffrey pine type.

The number of sites in each development class differed among the vegetation communities and among taxonomic groups. Based on the entire sample of 305 sites, the Jeffrey pine type had the most equitable distribution of sites along the development gradient, but only had eight sites with low development. Conversely, mixed conifer and subalpine conifer types had many sites with low development (42 and 44 sites, respectively), but only 11 sites in the mixed conifer type had high development, and only 11 sites in the subalpine conifer type had more than 1 percent development.

Figure 2.—*Factor score plots for axes 1–4 from principal components analysis of vegetation and environmental features of 305 sites illustrating the features separating the three primary vegetation community types, Lake Tahoe Basin, CA and NV.*



Individual and Cross-Taxonomic IEDs

Birds. Although 40 bird species were identified as significant indicators of a particular development level in one or more vegetation communities based on ISA results, only 22 bird species were consistently associated with low or high development conditions across vegetation communities and/or consistent with correlation results (appendix A). Similarly, 16 bird species groups, including total bird abundance, were significant indicators of development level based on ISA; but only five were consistently associated with low or high development conditions across vegetation communities. With the addition of species and species groups significantly correlated with development, 28 bird species and 16 species groups were identified as negatively associated with developed conditions, and 11 species and 8 species groups as positively associated with developed conditions (appendix A).

A final set of metrics was determined by evaluating the consistency and strength of the relationships between a species or species group and land development among the three vegetation communities. Eight metrics were selected: seven negatively associated and one positively associated with development (table 1). Richness value responses of foliage foragers, invertivores, shrub nesters, ground nesters, habitat specialists, and rare species either exceeded or paralleled corresponding abundance values; therefore, we chose to include those richness-based metrics. Consistently, granivore richness and ground forager richness were positively associated with development, but it is likely that those species groups would not increase in abundance at the highest land development levels, and their relationships with development were not steeply sloped, so these two groups were not included as metrics. The development-intolerant species group included all 28 species classified as either primary or secondary candidates. Intolerant species richness and abundance were highly correlated ($r = 0.804$); therefore, only richness was retained as a metric.

The resulting bird IED for each vegetation community, and the three types combined, displayed consistent declines in IED values with development. The strongest patterns were apparent in the Jeffrey pine and mixed conifer types, where the development gradient was strongest (fig. 3a).

Small Mammals. A total of 12 small mammal species, 8 species groups, and total small mammal abundance were significant indicators of a land development level in one or more vegetation communities, based on ISA results; however, only six species and eight species groups were consistently associated with low or high development conditions across vegetation communities and/or were consistent with correlation results (appendix A). With the addition of species significantly correlated with development, five small mammal species and nine species groups were identified as negatively associated with developed conditions and 5 species and 10 species groups were positively associated with developed conditions (appendix A).

Strong correlations between species and species groups eliminated many candidate groups from inclusion in the final set of metrics. Deer mouse abundance was a primary contributor to the abundance of three species groups, resulting in high correlations between deer mouse abundance and the abundance of associated groups: *Peromyscus* ($r = 0.99$), nocturnal species abundance ($r = 0.92$), and omnivore abundance ($r = 0.94$). Similarly, Douglas squirrel abundance was a primary contributor to canopy associate abundance ($r = 0.99$). Therefore, none of these groups was considered for inclusion as metrics in the final index.

A final set of five species metrics was selected, including one tolerant and four intolerant (table 1). Metrics of species intolerant of land development included the richness of intolerant species, rare species, ground associates, and herbivorous species. Six species comprised the intolerant species group, including the infrequently detected long-tailed weasel and ermine. These two mustelid species were included in the group of intolerant species, as opposed to being a separate metric (*Mustela* abundance), because they were infrequently occurring species and alone did not constitute a strong metric. Only the relative abundance of development-tolerant species (comprised of four species) was included as a tolerant metric. The small mammal IED for each vegetation community displayed consistent declines in the IED values with development (fig. 3b).

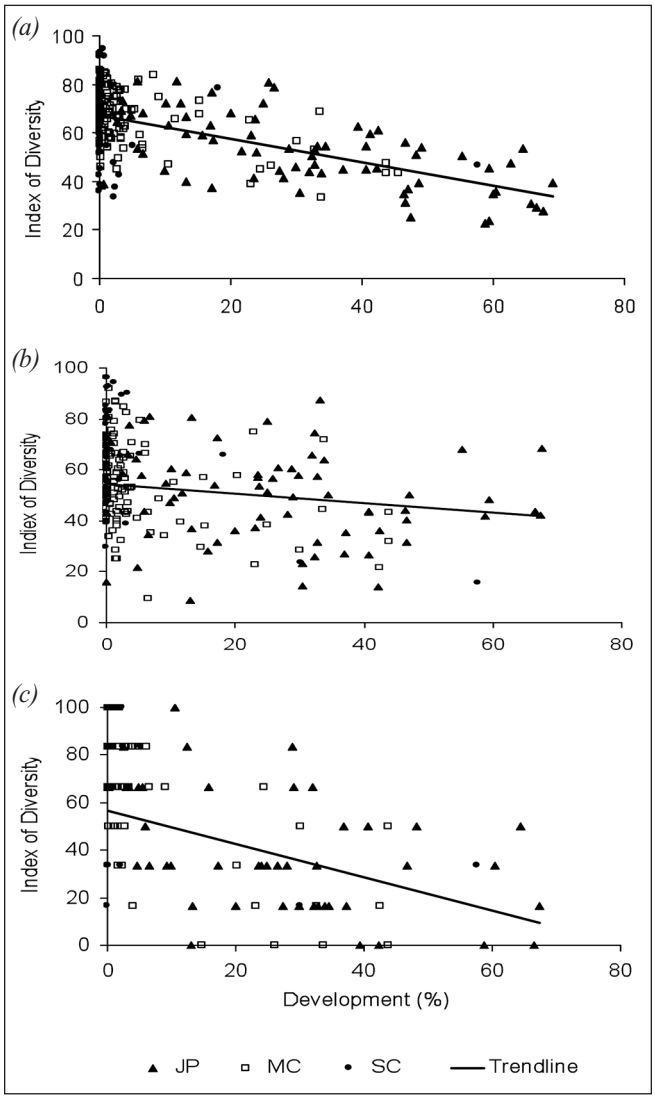
Mammalian Carnivores. Three metrics were used in the large mammal carnivore IED: richness of species intolerant of or impacted by land development (marten, bobcat, weasels, spotted skunk, bear); richness of species tolerant of development and able to exploit and benefit from it (coyote, ringtail, striped

Table 1.—*Descriptive statistics for composite multitaxonomic Index of Ecological Diversity values for reference and existing conditions in each of the three primary vegetation community types the Lake Tahoe basin based on conditions in 2006.*

Vegetation community type	Reference condition			Existing condition			t-test	
	n	Mean	s.d.	n	Mean	s.d.	T	P
Jeffrey pine	5	64.5	7.93	11	40.6	17.57	– 3.74	0.002
Mixed conifer	33	67.0	10.43	15	63.2	15.96	– 0.85	0.408
Subalpine conifer	22	67.93	13.27	14	68.0	13.65	0.01	0.992

skunk, gray fox, raccoon); and richness of domestic animals known to harass and negatively impact wildlife (domestic dogs and cats). The large mammal IED was significantly correlated with development at the basinwide scale, and within each of the three vegetation communities (fig. 3c). Although the variance in large mammal index values was high relative to development, an evident decline in maximum values was observed as development increased particularly within the Jeffrey pine and mixed conifer types. Maximum large mammal index values of 1.0 were observed for sites with 10.6 percent or less development, and index values greater than 0.5 were not observed at sites exceeding 32 percent development.

Figure 3.—*Index of Ecological Diversity values for birds (a), small mammals (b), and mammalian carnivores (c) in the Lake Tahoe Basin. Sites associated with each of the three primary vegetation community types are indicated (JP = Jeffrey pine; MC = mixed conifer; SC = subalpine conifer).*



Cross-Taxonomic Relationships. The composite IED showed significant negative associations with land development for each vegetation community (fig. 4). Correlations of the composite index and development were strongest in the mixed conifer type (r = -0.688, $P < 0.001$), followed by the Jeffrey pine type (r = -0.629, $P < 0.001$), and the subalpine conifer type (r = -0.539, $P = 0.007$). The basinwide composite index was also highly significant (r = -0.722, $P < 0.001$).

Ecological Diversity Assessment

The existing conditions of Jeffrey pine, mixed conifer, and subalpine conifer types were represented by 11, 15, and 14 sites, respectively. Reference sites in the Jeffrey pine type were limited (n = 5); however, comparisons between reference and existing conditions across vegetation communities showed clear patterns (fig. 5). Existing conditions in Jeffrey pine were significantly degraded compared to reference conditions based on IED values, and showed greater deviation from reference compared to mixed conifer and subalpine conifer. The majority of existing Jeffrey pine sites (more than 80 percent) had IED values less than 50, while reference sites had values more than 50 (fig. 5). Existing conditions in mixed conifer had a wider range of IED values, including sites with IED values less than 40, which were absent from reference sites, but as a group, IED values for existing conditions were not significantly

Figure 4.—*Index of Ecological Diversity values for the composite-taxonomic index in the Lake Tahoe Basin. Sites associated with each of the three primary vegetation community types are indicated (JP = Jeffrey pine; MC = mixed conifer; SC = subalpine conifer).*

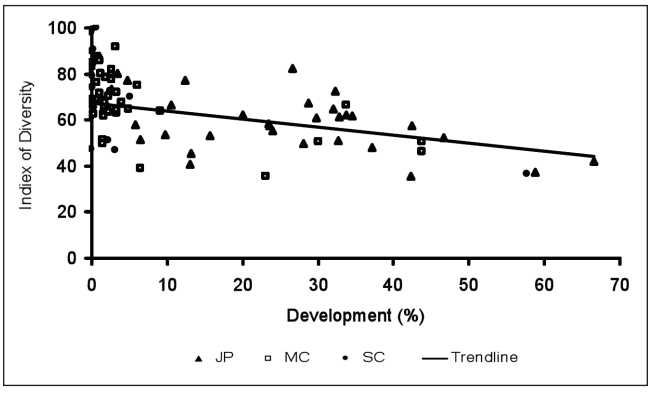
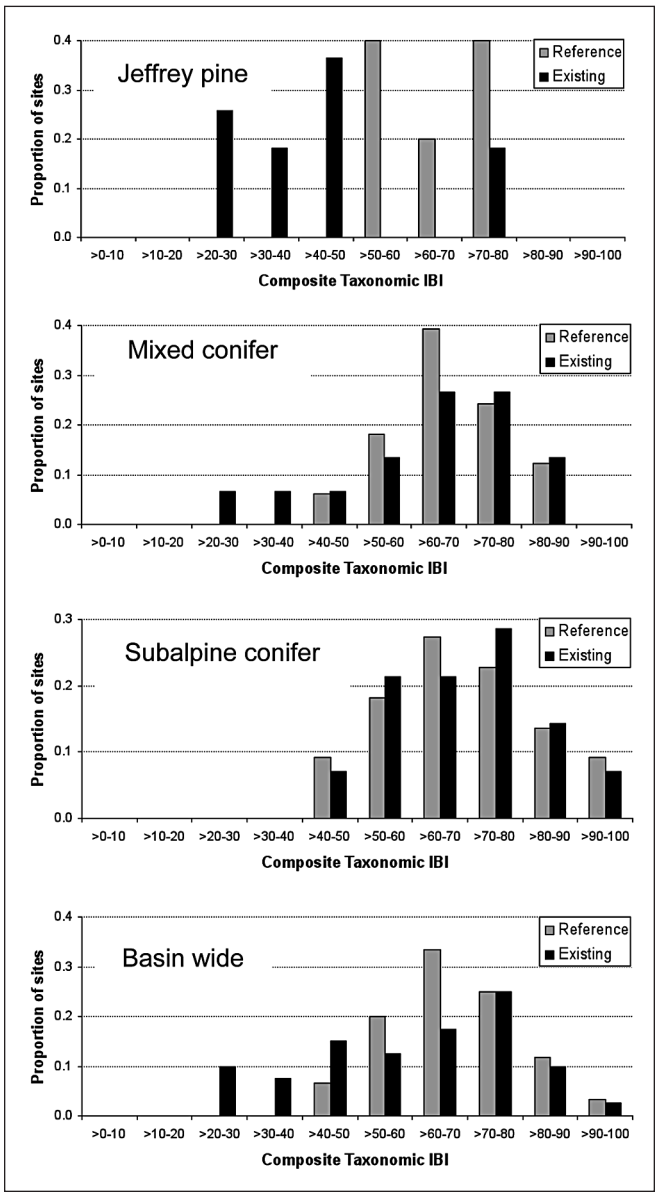


Figure 5.—*A comparison of the distribution of composite taxonomic Index of Ecological Diversity values for reference (less than 5 percent development) and existing sites in each of the three primary vegetation community types (Jeffrey pine, mixed conifer, and subalpine conifer) and basinwide in the Lake Tahoe Basin.*



different from those of reference sites. Existing conditions in the subalpine type did not deviate significantly from reference conditions. Basinwide roughly 20 percent of the sites fell outside the range of reference conditions, with most of those sites in the Jeffrey pine type.

Discussion

Efficiency of Multispecies Indicators

The search for indicators of ecological diversity is a challenging endeavor. The label indicator in the context of individual species has been used in multiple references, including individual species whose presence or abundance indicates a set of resource conditions associated with a number of other species (i.e., umbrella species), species dependent upon a particular resource for some aspect of their life history (e.g., secondary cavity nesters), and species that are particularly sensitive to environmental change (Caro and O'Doherty 1999). The alternative to individual species indicators is to use composite species metrics that represent a significant proportion of the ecosystem being monitored. In general, species assemblages have a greater probability of representing the condition of the resources upon which they are dependent, since individual species respond idiosyncratically to changing environmental conditions.

Monitoring species occurrences based on a systematic sampling design, specifically the FIA grid, provides a wealth of information that can serve to address agency information needs, including the development of indicators of ecological diversity. Despite the typical need to minimize the cost of environmental monitoring and assessment, it is unlikely that one or a few species, vertebrate or otherwise, will be able to provide a sensitive and informative representation of the condition of entire communities or ecosystems. Although only three survey methods were employed in this example, a large number (n = 126) and proportion (more than 90 percent of species expected to be detected with these methods and 84 percent of all terrestrial vertebrate species in the study area) were actually detected (Manley *et al.* 2005). This indicates that multiple-species and multi-taxonomic surveys in association with a systematic grid are likely to detect a portion of resident of species sufficient to derive species assemblage indicators.

ISA succeeded in allowing identification of species uniquely associated with one end of the land development gradient; however, it was important to consider multiple sources of information on patterned associations between species and species groups, and land development. ISA is a valuable tool for identifying species that are highly sensitive to development, but

it is insensitive to more subtle impacts on species (e.g., gradual declines in species abundance or frequency of occurrence). Furthermore, ISA cannot be used to evaluate species richness measures for potential indicator metric purposes. Conversely, correlations are sensitive to consistent declines in abundance or richness, but are insensitive to declines in maximum abundance and richness (i.e., depression of highest values as development increases or decreases). Notably both methods fail to identify impacts to species with low frequency of occurrence, a condition exhibited by a large proportion of species in most ecosystems (Gaston 1994), particularly upper trophic-level species.

The bird IED showed the most consistent and precise indication of biotic changes associated with land development. The strength of the bird IED is probably a function of both the sensitivity of songbirds to local conditions and the large number of species in the bird community conferring a greater range of richness and abundance values. The small mammal IED was intermediate in its precision relative to birds and mammalian carnivores. Although mammalian carnivores were the least precise, they did not decrease the precision of the composite IED, and they added a valuable ecological component to the index.

The systematic grid design made it possible to stratify the sample of sites into two different subpopulations of sites—reference and existing conditions—such that we could evaluate the current condition of ecological diversity for each vegetation community and throughout the study area. The use of reference conditions as a baseline against which existing conditions are evaluated has proved to be a valuable and reliable environmental assessment method (e.g., Wright *et al.* 2000). Although the sample size was limited, we were able to identify a significant deviation of existing conditions in the lowest elevation Jeffrey pine type, and thus identify an increase in the integrity of native vertebrate diversity in vegetation communities that were successively distant from the primary concentration of land development in the Lake Tahoe Basin.

The success of this approach in assessing ecological diversity suggests promise for applications of grid-based population monitoring using basic survey methods in meeting agency requirements to evaluate and monitor progress toward meeting desired ecological conditions. The approach provided sufficient information to derive reliable indicator metrics and apply them in developing a baseline of condition descriptors associated with reference conditions, and in evaluating existing conditions relative to reference or desired conditions. In this research effort, land development was used as the measure of human-generated land disturbance, but IEDs and deviations between existing and desired conditions could be developed for any number of anthropogenic stressors, such as pollution, human activity, and forest management activities. In the case of National Forest System lands management, it would be both prudent and feasible to develop IEDs that are informed by multiple stressors, using them individually and together to inform management on the status of ecological diversity and the environmental factors that put that diversity at risk.

Acknowledgments

Many agencies and individuals contributed to the collection of data used in this analysis. Funding and staff support was provided by the California Tahoe Conservancy, Tahoe Regional Planning Agency, Nevada Division of State Lands, and Forest Service Lake Tahoe Basin Management Unit. S. Parks, J. Roth, S. Romsos, S. Willits, and J. Karr supported core components of the project.

Literature Cited

Angermeier, P.L.; Karr, J.R. 1986. Applying an index of biotic integrity based on stream-fish communities: considerations in sampling and interpretation. North American Journal of Fisheries Management. 6(3): 418-429.

Barrett, R.H. 1983. Smoked aluminum track plots for determining furbearer distribution and relative abundance. California Fish and Game. 69: 188-190.

Blondel, J.; Ferry, C.; Frochot, B. 1981. Point counts with unlimited distance. In: Ralph, C.J.; Scott, J.M., eds. Estimating numbers of terrestrial birds. Studies in Avian Biology. 6: 414-420.

Bradford, D.F.; Franson, S.E.; Neale, A.C.; Heggam, D.T.; Miller, G.R.; Canterbury, G.E. 1998. Bird species assemblages as indicators of biological integrity in Great Basin rangeland. Environmental Monitoring and Assessment. 49: 1-22.

Bryce, S.A.; Hughes, R.M.; Kaufmann, P.R. 2002. Development of a bird integrity index: using bird assemblages as indicators of riparian condition. Environmental Management. 30: 294-310.

Cairns, J.; McCormick, P.V.; Neiderlehner, B.R. 1993. A proposed framework for developing indicators of ecosystem health. Hydrobiologica. 236: 1-44.

California Department of Fish and Game (CDFG). 2002. California Wildlife Habitat Relationships, version 8.0 personal computer program. Sacramento, CA: California Department of Fish and Game, California Interagency Wildlife Task Group.

California Department of Transportation (Caltrans). 2001. Highway design manual. Sacramento, CA: California Department of Transportation, Project Delivery, Division of Design.

Caro, T.M.; O'Doherty, G. 1999. On the use of surrogate species in conservation biology. Conservation Biology. 13: 805-814.

Daly, C. 1995. U.S. monthly precipitation, 1961 to 1990 base period (digital map layer). Corvalis, OR: Oregon State University.

Dufrêne, M.; Legendre, P. 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecological Monographs. 67: 345-366.

ESRI, Inc. 1995. ArcView ver. 3.2. Redlands, CA: Environmental Systems Research Institute.

Feinsinger, P. 2001. Designing field studies for biodiversity conservation. Washington, DC: Island Press.

Fowler, C.H. 1994. Marten and fisher survey techniques. Arcata, CA: Humboldt State University. 120 p. M.S. thesis.

Gaston, K.J. 1994. Rarity. New York: Chapman and Hall.

Holdridge, L.R. 1967. Life zone ecology. San Jose, Costa Rica: Tropical Science Center.

Holthausen, R.; Czaplewski, R.; DeLorenzo, D.; Hayward, G.; Kessler, W.; Manley, P.N.; McKelvey, K.; Powell, D.S.; Ruggiero, L.; Schwartz, M.; Van Horne, B.; Vojta, C. 2005. Strategies for monitoring terrestrial animals and habitats. Gen. Tech. Rep. RMRS-GTR-161. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Karr, J.R. 1981. Assessment of biotic integrity using fish communities. Fisheries. 6(6): 21-27.

Karr, J.R.; Chu, E.W. 1995. Ecological integrity: reclaiming lost connections. In: Westra, L.; Lemons, J., eds. Perspectives on ecological integrity. Dordecht, Netherlands: Kluwer Academic: 34-48.

Karr, J.R.; Kimberling, D.N. 2003. A terrestrial arthropod index of biological integrity for shrub-steppe landscapes. Northwest Science. 77: 202-213.

Landres, P.B.; Verner, J.; Thomas, J.W. 1988. Ecological uses of vertebrate indicator species: a critique. Conservation Biology. 2: 316-329.

Lindstrom, S.; Rucks, P.; Wigand, P. 2000. A contextual overview of human land use and environmental conditions. In: Murphy, D.D.; Knopp, C.M., eds. Lake Tahoe watershed assessment. Gen. Tech. Rep. PSW-GTR-175. Albany, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station: 23-127. Vol. I.

Lugo, A.E.; Brown, S.L.; Dodson, R.; Smith, T.S.; Shugart, H.H. 1999. The Holdridge life zones of the conterminous United States in relation to ecosystem mapping. Journal of Biogeography. 26(5): 1025-1038.

Manley, P.N. 2000. Biological diversity and its measure: an assessment of lotic riparian ecosystems in the Lake Tahoe Basin. Berkeley, CA: University of California. 416 p. Ph.D. dissertation.

Manley, P.N.; Fites-Kaufman, J.A.; Barbour, M.G.; Schlesinger, M.D.; Rizzo, D.M. 2000. Biological integrity. In: Murphy, D.D.; Knopp, C.M., eds. Lake Tahoe watershed assessment. Gen. Tech. Rep. PSW-GTR-175. Albany, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station: 403-598. Vol. I.

Manley, P.N.; McIntyre, K.K. 2006. Indicators of ecological integrity for terrestrial ecosystems in the Lake Tahoe basin. Final report submitted to the Forest Service, Lake Tahoe Basin Management Unit, South Lake Tahoe, CA: 104 p.

Manley, P.N.; Murphy, D.D.; Campbell, L.A.; Heckmann, K.E.; Merideth, S.; Parks, S.A.; Sanford, M.P.; Schlesinger, M.D. 2006. Biotic diversity interfaces with urbanization in the Lake Tahoe Basin. California Agriculture. 60(2): 59-64.

Manley, P.N.; Parks, S.A.; Campbell, L.A.; Schlesinger, M.D. 2009. Modeling urban land development as a continuum to address fine-grained habitat heterogeneity. Landscape and Urban Planning: 28-36.

Manley, P.N.; Schlesinger, M.D.; Roth, J.K.; Van Horne, B. 2005. A field-based evaluation of a presence-absence protocol for monitoring ecoregional-scale biodiversity. Journal of Wildlife Management. 69(3): 950-966.

Manley, P.N.; Van Horne, B.; Roth, J.K.; Zielinski, W.J.; McKenzie, M.M.; Weller, T.J.; Weckerly, F.W.; Hargis, C. 2006. Multiple species inventory and monitoring technical guide. Gen. Tech. Rept. WO-GTR-73. Washington, DC: U.S. Department of Agriculture, Forest Service, Washington Office.

Manley, P.N.; Zielinski, W.J.; Schlesinger, M.D.; Mori, S.R. 2004. Evaluation of a multiple-species approach to monitoring species and ecosystem conditions at the ecoregional scale. Ecological Applications. 14(1): 296-310.

Mayer, K.E.; Laudenslayer, W.F., eds. 1988. A guide to wildlife habitats of California. Sacramento, CA: California Department of Fish and Game.

McCune, B.; Mefford, M.J. 1999. PC-ORD. Multivariate analysis of ecological data. Version 4.17. Gleneden Beach, OR: MjM Software Design.

Morrison, M.L. 1986. Bird populations as indicators of environmental change. Current Ornithology. 3: 429-451.

Moyle, P.B.; Marchetti, M.P. 1999. Applications of indices of biotic integrity to California streams and watersheds. In: Simon, T.P., ed. Assessing the sustainability and biological integrity of water resources using fish communities. Boca Raton, FL: CRC Press.

Noon, B.R.; Spies, T.A.; Raphael, M.G. 1999. Conceptual basis for designing an effectiveness monitoring program. In: Mulder, B.S.; Noon, B.R.; Spies, T.A.; Raphael, M.G.; Palmer, C.J.; Olsen, A.R.; Reeves, G.H.; Welsh, H.H., Jr., eds. The strategy and design of the effectiveness program for the Northwest forest plan. Gen. Tech. Rep. PNW-GTR-437. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Station: 21-48.

Noss, R.F. 1990. Indicators for monitoring biodiversity: a hierarchical approach. Conservation Biology. 4: 355-364.

Ralph, C.J.; Geupel, G.R.; Pyle, P.; Martin, T.E.; DeSante, D.F. 1993. Handbook of field methods for monitoring landbirds. Gen. Tech. Rept. GTR-PSW-144. Albany, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station. 41 p.

Roesch, F.A.; Reams, G.A. 1999. Analytical alternatives for an annual inventory system. Journal of Forestry. 97: 33-37.

Roth, J.K.; Manley, P.N.; McKenzie, M.M.; Schlesinger, M.D. 2004. Multiple-species inventory and monitoring 2002 monitoring report. Interim evaluation and report. South Lake Tahoe, CA: U.S. Department of Agriculture, Forest Service, Lake Tahoe Basin Management Unit. 198 p.

U.S. Department of Agriculture (USDA) Forest Service. 1991. Existing vegetation. Vallejo, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Region.

USDA Forest Service. 2001. Existing vegetation (updated). Vallejo, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Region.

Woodbridge, B.; Hargis, C.D. 2006. Northern Goshawk inventory and monitoring technical guide. Gen. Tech. Rep. WO-71. Washington, DC: U.S. Department of Agriculture, Forest Service. 80 p.

Wright, J.F.; Sutcliffe, D.W.; Furse, M.T., eds. 2000. Assessing the biological quality of fresh waters; RIVPACS and other techniques. Ambleside, United Kingdom: Freshwater Biological Association.

Zielinski, W.J.; Kucera, T.E. 1995. American marten, fisher, lynx, and wolverine: survey methods for their detection. Gen. Tech. Rep. PSW-GTR-157. Albany, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station.

Appendix A (1 of 3)

Bird and small mammal species and species groups identified as candidate metrics in Indexes of Biological Diversity (scientific names available in Manley and McIntyr 2006). Selected metrics indicated in bold. Species relationships with development were evaluated based on indicator species analysis (ISA), Pearson's correlation coefficients across all three vegetation community types (= P < 0.05, ** P < 0.01, *** P < 0.001) and per type (significant correlations per type represented by X), and bivariate scatterplots (not shown). Dashes indicate not applicable*

Potential metrics	Basinwide	Vegetation community type					
		Jeffrey pine		Mixed conifer		Subalpine conifer	
	Correlation	ISA	Correlation	ISA	Correlation	ISA	Correlation
Birds—intolerant metrics							
Primary species:							
Black-backed woodpecker	**– 0.149***	**X**	**X**				
Chipping sparrow	**– 0.163***		**X**	**X**			
Dark-eyed junco	**– 0.493****	**X**	**X**	**X**	**X**		**X**
Dusky flycatcher	**– 0.430****	**X**	**X**	**X**	**X**		
Green-tailed towhee	**– 0.183****	**X**	**X**	**X**			
Hairy woodpecker	**– 0.143****	**X**	**X**			**X**	
Hermit thrush	**– 0.384*****	**X**	**X**	**X**	**X**	**X**	
MacGillivray's warbler	**– 0.220*****		**X**	**X**	**X**		
Mountain quail	**– 0.245*****	**X**	**X**		**X**		
Nashville warbler	**– 0.286*****	**X**	**X**				
Olive-sided flycatcher	**– 0.292*****	**X**	**X**	**X**	**X**		
Townsend's solitaire	**– 0.375*****	**X**	**X**		**X**		
Warbling vireo	**– 0.256*****	**X**	**X**	**X**	**X**		
Western tanager	**– 0.163****	**X**	**X**	**X**			
Secondary species:							
Brown creeper	**– 0.206****				**X**		
Cassin's vireo	**– 0.204****		**X**				
Golden-crowned kinglet	**– 0.304*****		**X**		**X**		
Hermit warbler	**– 0.193****				**X**		
Lincoln sparrow	**– 0.178****		**X**				
Pine grosbeak	**– 0.237*****					**X**	
Pine siskin	**– 0.154****		**X**				
White-breasted nuthatch	**– 0.159****		**X**				
Wilson's warbler	**– 0.267*****				**X**		
Yellow-rumped warbler	**– 0.334*****		**X**		**X**		
Blue grouse	**– 0.179****						
Calliope hummingbird	**– 0.169****						
Cassin's finch	**– 0.200****						
Pileated woodpecker	**– 0.140***						
Primary groups:							
Foliage forager richness	**– 0.427*****	**—**	**X**	**—**	**X**	**—**	
Invertivore richness	**– 0.417*****	**—**	**X**	**—**	**X**	**—**	
Foliage forager abundance	– 0.434***		X	X	X		
Invertivore abundance	– 0.266***	X	X	X	X		
Secondary groups:							
Air forager richness	– 0.114*	—	X	—	X	—	
Ground nester richness	**– 0.493*****	**—**	**X**	**—**	**X**	**—**	
Habitat specialist richness	**– 0.335*****	**—**	**X**	**—**		**—**	
Rare species richness	**– 0.225*****	**—**		**—**	**X**	**—**	
Moderate specialist richness	– 0.136*	—	X	—		—	
Shrub nester richness	**– 0.321*****	**—**	**X**	**—**	**X**	**—**	
Air forager abundance	– 0.315***		X				
Ground forager abundance	– 0.334***		X				

ISA = indicator species analysis.

Appendix A (2 of 3)

Bird and small mammal species and species groups identified as candidate metrics in Indexes of Biological Diversity (scientific names available in Manley and McIntyr 2006). Selected metrics indicated in bold. Species relationships with development were evaluated based on indicator species analysis (ISA), Pearson's correlation coefficients across all three vegetation community types (= P < 0.05, ** P < 0.01, *** P < 0.001) and per type (significant correlations per type represented by X), and bivariate scatterplots (not shown). Dashes indicate not applicable*

Potential metrics	Basinwide	Vegetation community type					
		Jeffrey pine		Mixed conifer		Subalpine conifer	
	Correlation	ISA	Correlation	ISA	Correlation	ISA	Correlation
Ground nester abundance	– 0.535***	X	X	X	X		
Habitat specialist abundance	– 0.247***	X	X		X		
Habitat generalist abundance	– 0.258***		X				
Shrub nester abundance	– 0.345***		X	X	X		
Birds—tolerant metrics							
Primary species:							
Band-tailed pigeon	**0.421*** **			**X**	**X**		
Brewer's blackbird	**0.671*** **	**X**		**X**	**X**		**X**
Brown-headed cowbird	**0.635*** **	**X**		**X**	**X**	**X**	**X**
Mountain chickadee	**0.236*** **	**X**		**X**			
Mourning dove	**0.468*** **		**X**	**X**	**X**	**X**	**X**
Pygmy nuthatch	**0.643*** **	**X**	**X**	**X**	**X**	**X**	**X**
Steller's jay	**0.485*** **	**X**	**X**	**X**	**X**		
Secondary species:							
Common raven	**0.304*** **		**X**		**X**		**X**
Downy woodpecker	0.128*				X		
Northern flicker	**0.342*** **				**X**		
White-headed woodpecker	**0.317*** **				**X**		
Primary groups:							
Corvid abundance	0.292***	X		X	X		X
Omnivore abundance	0.303***	X			X		X
Malentity abundance	0.605***	X	X	X	X	X	X
Malentity richness	0.338***				X		X
Moderate specialist abundance	0.122*	X					
Tolerant species—relative abundance	0.859***	X	X	X	X	X	X
Secondary groups:							
Granivore richness	0.168*	—		—		—	X
Ground forager abundance	0.333***		n				X
Small Mammals— Intolerant Metrics							
Primary species:							
Lodgepole pine chipmunk	**– 0.253*** **			**X**	**X**		**X**
Secondary species:							
Deer mouse	**– 0.332*** **				**X**		**X**
Golden-mantled ground squirrel	– 0.122						X
Shadow chipmunk	**– 0.133***					**X**	
Western jumping mouse	**– 0.117**						

ISA = indicator species analysis.

Appendix A (3 of 3)

Bird and small mammal species and species groups identified as candidate metrics in Indexes of Biological Diversity (scientific names available in Manley and McIntyr 2006). Selected metrics indicated in bold. Species relationships with development were evaluated based on indicator species analysis (ISA), Pearson's correlation coefficients across all three vegetation community types (= P < 0.05, ** P < 0.01, *** P < 0.001) and per type (significant correlations per type represented by X), and bivariate scatterplots (not shown). Dashes indicate not applicable*

Potential metrics	Basinwide	Vegetation community type					
		Jeffrey pine		Mixed conifer		Subalpine conifer	
	Correlation	ISA	Correlation	ISA	Correlation	ISA	Correlation
Primary groups:							
Ground associate richness	**– 0.240*** **	**—**		**—**		**—**	
Herbivore richness	**– 0.329*** **	**—**	**X**	**—**	**X**	**—**	
Peromyscus species abundance	– 0.332***			X			
Nocturnal species abundance	– 0.281***			X			
Omnivore species abundance	– 0.130*			X			
Secondary groups:							
Habitat generalist richness	– 0.127	—		—		—	
Rare species richness	**– 0.149***	**—**		**—**		**—**	
Intolerant species richness	**– 0.139***						
***Mustela* species abundance**	**– 0.126***				**X**		**X**
Small mammals— tolerant metrics							
Primary species:							
Douglas squirrel	**0.496*** **	**X**	**X**	**X**	**X**		**X**
California ground squirrel	**0.487*** **	**X**	**X**	**X**	**X**	**X**	
Long-eared chipmunk	**0.325*** **			**X**	**X**		
Secondary species:							
Long-tailed vole	0.193**		X				
Yellow-pine chipmunk	**0.197****					**X**	**X**
Primary groups:							
Nocturnal species richness	0.066**	—		—		—	
Omnivore richness	0.340***	—		—		—	
Canopy species richness	0.480***	—		—		—	
Microtus species abundance	0.121*	X					
Diurnal species abundance	0.205**	X			X	X	X
Habitat specialist abundance	0.133*	X		X			
Spermophilus species abundance	0.189**			X			
Canopy species abundance	0.323***						
Tamias species abundance	0.121*					X	
Tolerant species relative abundance	**0.559*** **			**X**	**X**	**X**	**X**

ISA = indicator species analysis.

Wildlife Monitoring Across Multiple Spatial Scales Using Grid-Based Sampling

Kevin S. McKelvey, Samuel A. Cushman, Michael K. Schwartz, and Leonard F. Ruggiero

Abstract.—Recently, noninvasive genetic sampling has become the most effective way to reliably sample occurrence of many species. In addition, genetic data provide a rich data source enabling the monitoring of population status. The combination of genetically based animal data collected at known spatial coordinates with vegetation, topography, and other available covariates enables development of habitat relationships and evaluation of population attributes, such as connectivity. Colocating animal occurrence sampling on an extensive vegetation plot grid, such as the forest inventory and analysis grid, provides opportunities to develop statistical models and monitor animal populations.

Introduction

The forest inventory and analysis (FIA) plot network provides a strong framework for assessing changes in forest resources over time and, with some constraints due to plot spacing and resource rarity, across space (Frayer and Furnival 1999). Linking FIA data to other co-occurring resources at FIA plot locations provides both direct monitoring of the state of these resources and the development of relationships between these resources and FIA data (Zielinski *et al.* 2006). Here we discuss opportunities associated with collecting animal occurrence data on grids where vegetation data are simultaneously collected. In many cases, noninvasive genetic sampling (e.g., DNA from hair or feces) provides an efficient and information-rich method to survey species.

Historically, direct monitoring of wildlife populations has been considered difficult and expensive, and wildlife population dynamics have largely been inferred through opinion-based habitat relationships. These include systems such as the California Habitat Relationships Database (CDFG 2000) and specific Habitat Suitability Index (USFWS 1981) models that have been developed for many species. The accuracy of these models is unknown and probably low (Block *et al.* 1994); habitat relationships models generally evaluate habitat at a single spatial scale, involve relatively few variables, and weighting factors are based on guesses rather than on estimates that are statistically derived. In many cases, model quality is not formally testable because habitat rankings (e.g., high, medium, or low) have no precise definitions.

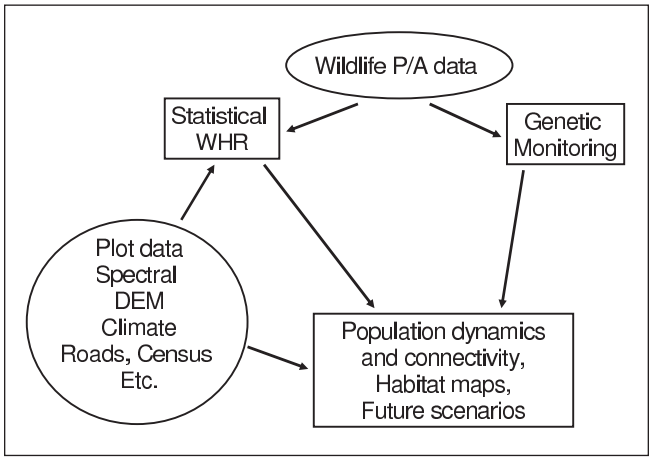
Recently, however, determining animal occurrence or abundance has become much easier due to the rapid evolution of genetic technologies. Rather than needing to trap or photograph an organism to determine presence, many organisms can be identified using evidence such as hair, scats, or feathers. These "noninvasive" genetic samples contain diagnostic DNA for identifying individuals, populations, and species (Morin *et al.* 2001, Taberlet *et al.* 1996). Snowshoe hares (*Lepus americanus*), for example, produce 400 to 500 pellets per day (Hodges 1999). Thus, if snowshoe hares are on a site, pellets (and through DNA confirmation that the pellets were produced by snowshoe hares) can be collected with high reliability and little effort. Because noninvasive sampling facilitates collection of representative samples, rare and elusive organisms such as Canada lynx (*Lynx canadensis*) can now be monitored (McDaniel *et al.* 2000).

An organism's DNA contains a rich source of data. Identifying an organism to species using DNA is inexpensive, reliable, and unambiguous (McKelvey *et al.* 2006, Mills *et al.* 2000). Species identification, however, represents only a small portion of the information that can be gleaned through genetic analyses. The size of the organism's population, genealogical relationships to other organisms identified in the area, movement between habitat islands, and many other population attributes can be inferred through analysis of genetic patterns (Luikart and England 1999, Schwartz *et al.* 2007). Many of these genetic analyses require relatively small samples on individuals (e.g., 20 to 40), obviating the need to sample a large proportion of the population to develop strong statistical inferences.

In addition to the direct monitoring of wildlife populations based on genetic measures, new statistical approaches (Cushman and McGarigal 2002, 2004) enable development of habitat associations between animal occurrence or abundance with vegetation data from multiple scales, thereby increasing both the utility of habitat understandings and statistical power. Together, these new developments enable cost-effective wildlife monitoring and the generation of robust statistical wildlife habitat relationships (WHR) models.

Importantly, the monitoring of wildlife populations and the development of statistical WHR models can be analyzed using the same data, and we view these activities as complementary parts of a generalized approach to evaluating multiple wildlife species across extensive landscapes (fig 1). Genetic samples can be used to directly track population status, but, lacking links to vegetation, genetic monitoring contains little information about the effects of management on populations. Multiscale WHR, being population surrogates, are intrinsically weaker than direct monitoring for the purpose of tracking populations but enable quantification of direct relationships between organisms and the landscapes they inhabit. The combination of these two approaches enables higher level analyses such as estimation of the habitat quality and population status in future landscapes and evaluation of landscape connectivity. Therefore, we begin with discussions of genetic monitoring and multiscale WHR and follow with an evaluation of landscape connectivity using vegetation, topography, and genetic relationships derived from grid-based black bear (*Ursus americanus*) data.

Figure 1.—*A diagram showing the relationships between animal occurrence data, vegetation, topography, and anthropogenic data, and monitoring products. The same data set serves multiple, mutually reinforcing purposes.*



Genetic Monitoring

Genetic monitoring can evaluate important population characteristics such as population size or connectivity. In many cases, information on the historical status of a population can be gleaned by recovering DNA from archived material (e.g., museum skins, fish scales, or trophy collections) or by inference from patterns of genetic variation in a single contemporary population sample (e.g., deficit of rare alleles [Luikart *et al.* 1998]). This process enables "retrospective monitoring" to assess historical conditions (Poulsen *et al.* 2006).

One type of genetic monitoring (called Category I genetic monitoring in Schwartz *et al.* [2007]) uses noninvasive genetic sampling to count individuals. Multiple samples can be analyzed in a capture-mark-recapture (CMR) framework (Otis *et al*. 1978) to provide population estimates. A second type evaluates changes in population status through changes in genetic diversity using metrics such as expected heterozygosity (H_e) or allelic diversity (A). Both of these metrics, which can be readily measured, theoretically decline over time inversely with population size. H_e, for example, declines at a rate inversely proportional to two times the effective population size (N_e; a genetic measure of population size) and thus can be used to track population status. In a retrospective study on brown trout (*Salmo trutta*), H_e and A were examined in Denmark between 1944 and 1997; older samples were from scale collections found in museums (Østergaard *et al.* 2003). The analysis concluded that the population was stable over this time period, but maintenance of genetic diversity was dependent on gene flow between small local populations (Østergaard *et al.* 2003).

Several methods exist that convert changes in H_e or temporal changes in allele frequencies into N_e, thus enabling direct tests for changes in effective population size. In another retrospective study, N_e was estimated for brown bears in Yellowstone National Park using samples from the 1910s, 1960s, and 1990s

(Miller and Waits 2003). N_e estimates were ~85 for both time periods (1910–1960s and 1960s–1990s), again suggesting a stable, effective population size but a relatively small effective size.

Population estimates based on CMR yield immediate results whereas detecting a reduction in N_e requires a lag time (e.g., one generation) to allow changes in H_e or allele frequencies to occur (Schwartz *et al.* 1999). Thus, where an immediate estimate of population size is needed, calculating N_e has historically been viewed as a less desirable approach. However, new methods requiring only one sample for each N_e estimate are emerging (England *et al.* 2006, Waples 2006). Although their precision and reliability have not yet been thoroughly quantified, the ability to estimate N_e based on a single sample would greatly expand the utility of N_e as a monitoring tool (Schwartz *et al.* 2007).

Multiscale Wildlife Habitat Relationships

Multiscale gradient modeling is a multivariate, nonlinear regression approach that enables the researcher to predict habitat quality and abundance of organisms based on ecological factors from a range of spatial scales (Cushman and McGarigal 2002, 2003, 2004). Hierarchical gradient modeling maximizes predictive power by optimally incorporating the influence of driving variables across multiple spatial scales, enabling researchers to quantify the independent and interactive effects of multiple factors and scales simultaneously. This includes presence and abundance of species, which are generally related to habitat elements associated with multiple spatial scales as well. For instance, Zabel *et al.* (1995) found that the northern spotted owl (*Strix occidentalis*) habitat was linked to the presence of nest and roost trees (fine scale) and proximity to dense young forest, which contained abundant prey (mid scale). In addition, northern spotted owl range is limited to wet, coastal climates (broad scale). By optimizing the relationship between organisms and environmental structure across scales gradient, models can be used to quantitatively describe species-environment relationships, assess predictive accuracy, and predict and map current species distributions across geographical regions under alternative scenarios.

Gradient models are particularly powerful for WHR because of three important attributes. First, empirical statistical models are developed by optimizing the measured relationship between patterns of organism distribution and measured environmental variation. Importantly, subjective opinion and assumptions are avoided because inserting expert opinion into empirically derived species distribution models may reduce model performance and predictive success (Seoane *et al.* 2005). Second, the distribution and abundance of each modeled species is predicted directly on the basis of measured environmental variables. Thus, the result is not a subjective proxy relationship like most past wildlife habitat models. Third, such models can be directly validated, and their precision, predictive power and error rates are quantified through grid-based sampling.

Landscape Connectivity

Predicting population-level impacts of landscape change depends on identifying factors that influence population connectivity. Most putative movement corridors and barriers have not been based on empirical data, however, largely due to the fact that dispersals are relatively rare and difficult to document. Historically, animal movements were tracked using radiotelemetry. Transmitters were high frequency and had relatively short ranges, so if animals left an area, their signals would be lost. For example, most long-range movements for Canada lynx are associated with lynx that could not be relocated by their transmitters but which were, by chance, turned in by trappers (Aubry *et al.* 2000, Mech 1977). Beyond the difficulties in reliably recording dispersals, a further difficulty lay in separating consequential movements. To be meaningful to population structure, a migrant must affect the population into which it immigrates. Generally, this action would mean reproducing. Because a collared migrant, almost by definition, moved from a studied population to an unstudied one, however, its activities and interactions at the new location were entirely unknown.

Because consequential movements involve breeding, migrants that breed reveal their movement patterns by the genes they leave behind. While genetic similarity is expected to decline with distance, areas with greater similarity than expected (as

a result of more genes shared) represent movement corridors. Conversely, areas that are more divergent are associated with barriers. In theory, reliable inferences about population connectivity can be obtained by correlating genetic similarity of individuals across large landscapes with hypothetical movement-cost models.

As an example, Cushman *et al*. (2006) compared the patterns of genetic similarity among 146 individual black bears sampled across a 3,000-square-km study area in northern Idaho. Genetic similarity was correlated with 110 movement-cost hypotheses describing a range of potential relationships between movement cost and land cover, slope, elevation, roads, Euclidean distance, and a putative movement barrier (a large agricultural valley that bisected the study area). Movement resistance hypotheses were divided into seven organizational models enabling the influence of barriers, distance, and landscape features to be statistically separated. It was found that gene flow patterns, and therefore consequential dispersal movements, were facilitated by contiguous forest cover at middle elevations. Cushman *et al.* (2006) were then able to map these understandings and present a data-based corridor map for the study area.

Discussion

Animal occurrence data, when colocated on a vegetation plot grid can provide statistical links between vegetative condition and population status. When simple occurrence data are expanded through genetic analyses, measures of population size, fragmentation, and connectivity can also be linked to vegetation and other landscape features. The examples presented above serve to demonstrate the potential of these approaches. Although many of the current examples of genetic monitoring have occurred in fisheries, this is not because of any intrinsic sampling advantages associated with aquatic systems. Fish were easy to catch (and therefore easier to collect DNA samples) but hard to count; DNA methods were therefore rapidly implemented to monitor the status and nature of fish populations. With new noninvasive DNA methods, however, DNA from many terrestrial organisms can also be easily collected.

Genetic technologies are rapidly changing; in the future, genetic analyses will be more powerful and less costly. Given the rapid rate of change in genetic technologies, coupled with promising analytical developments such as those illustrated here, the next decade will likely see an explosion of both new analysis methods and on-the-ground applications. Those who are planning animal monitoring should incorporate these understandings to ensure that implementations are maximally cost effective.

Literature Cited

Aubry, K.B.; Koehler, G.M.; Squires, J.R. 2000. Ecology of Canada lynx in southern boreal forests. In: Ruggiero, L.F. *et al*., eds. Ecology and conservation of lynx in the United States. Boulder, CO: University Press of Colorado: 373-396.

Block, W.M.; Morrison, M.L.; Verner, J.; Manley, P.N. 1994. Assessing wildlife-habitat-relationships models: a case study with California oak woodlands. Wildlife Society Bulletin. 22: 549-561.

California Department of Fish and Game (CDFG). 2000. Standards and guidelines for CWHR species models. Tech. Rep. No. 31. Sacramento, CA: California Department of Fish and Game, California Wildlife Habitat Relationships System, California Interagency Wildlife Task Group.

Cushman, S.A.; McGarigal, K. 2002. Hierarchical, multi-scale decomposition of species-environment relationships. Landscape Ecology. 17: 637-646.

Cushman, S.A.; McGarigal, K. 2003. Landscape-level patterns of avian diversity in the Oregon Coast Range. Ecological Monographs. 73(2): 59-281.

Cushman, S.A.; McGarigal, K. 2004. Hierarchical analysis of forest bird species-environment relationships in the Oregon Coast Range. Ecological Applications. 14: 1090-1105.

Cushman, S.A.; McKelvey, K.S.; Hayden, J.; Schwartz, M.K. 2006. Gene flow in complex landscapes: testing multiple hypotheses with causal modeling. American Naturalist. 168: 486-499.

England, P.R.; Cornuet, J.M.; Berthier, P.; Tallmon, D.A.; Luikart, G. 2006. Estimating effective population size from linkage disequilibrium: severe bias in small sizes. Conservation Genetics. 7: 303-308.

Frayer, W.E.; Furnival, G.M. 1999. Forest survey sampling designs: a history. Journal of Forestry. 97(12): 4-10.

Hodges, K.E. 1999. Proximate factors affecting snowshoe hare movements during a cyclic population low phase. Ecoscience. 6: 487-496.

Luikart, G.; Cornuet, J.M.; Allendorf, F.W.; Sherwin, W.B. 1998. Distortion of allele frequency distributions provides a test for recent population bottlenecks. Journal of Heredity. 89: 238-247.

Luikart, G.; England, P.R. 1999. Statistical analysis of microsatellite DNA data. Trends in Ecology and Evolution. 14: 253-256.

McDaniel, G.W.; McKelvey, K.S.; Squires, J.R.; Ruggiero, L.F. 2000. Efficacy of lures and hair snares to detect lynx. Wildlife Society Bulletin. 28: 119-123.

McKelvey, K.S.; von Kienast, J.; Aubry, K.B. *et al.* 2006. DNA analysis of hair and scat collected along snow tracks to document the presence of Canada lynx (*Lynx canadensis*). Wildlife Society Bulletin. 34: 451-455

Mech, L.D. 1977. Record movement of a Canadian lynx. Journal of Mammalogy. 58: 676-677.

Miller, C.R.; Waits, L.P. 2003. The history of effective population size and genetic diversity in the Yellowstone grizzly (*Ursus arctos*): implications for conservation. Proceedings of the National Academy of Sciences. 100: 4334-4339.

Mills, L.S.; Pilgrim, K.; Schwartz, M.; McKelvey, K.S. 2000. Identifying lynx and other North American cat species based on mtDNA analysis. Conservation Genetics. 1: 285-288.

Morin, P.A.; Chambers, K.E.; Boesch, C.; Vigilant, L. 2001. Quantitative polymerase chain reaction analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes verus*). Molecular Ecology. 10: 1835–1844.

Østergaard, S.; Hansen, M.M.; Loeschcke, V.; Nielson, E.E. 2003. Long-term temporal changes of genetic composition in brown trout (*Salmo trutta L.*) populations inhabiting an unstable environment. Molecular Ecology. 12: 3123–3135.

Otis, D.L.; Burnham, K.P.; White, G.C.; Anderson, D.R. 1978. Statistical inference from capture data on closed animal populations. Wildlife Monographs. 62: 1-135.

Poulsen, N.A.; Nielsen, E.E.; Schierup, M.H.; Loeschcke, V.; Grønkjær, P. 2006. Long-term stability and effective population size in North Sea and Baltic Sea cod (*Gadus morhua*). Molecular Ecology. 15: 321-331.

Schwartz, M.K.; Luikart, G.; Waples, R.S. 2007. Genetic monitoring as a promising tool for conservation and management. Trends in Ecology and Evolution. 22: 25-33.

Schwartz, M.K.; Tallmon, D.A.; Luikart, G. 1999. Using genetics to estimate the size of wild populations: many methods, much potential, uncertain utility. Animal Conservation. 2: 320-322.

Seoane, J.; Bustamante, J.; Dias-Delgado, R. 2005. Effect of expert opinion on the predictive ability of environmental models of bird distribution. Conservation Biology. 19: 512-522.

Taberlet, S.; Griffin, B.; Goossens, S. *et al.* 1996. Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic Acids Research. 26: 3189-3194.

U.S. Fish & Wildlife Service (USFWS). 1981. Standards for the development of habitat suitability index models for use in the Habitat Evaluation Procedures. ESM 103. Washington, DC: U.S. Department of the Interior, Fish and Wildlife Service, Division of Ecological Services.

Waples, R.S. 2006. A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. Conservation Genetics. 7: 167-184.

Zabel, C.J.; McKelvey, K.S.; Ward, J.P., Jr. 1995. Influence of primary prey on home-range size and habitat-use patterns of northern spotted owls (*Strix occidentalis caurina*). Canadian Journal of Zoology. 73: 433-439.

Zielinski, W.J.; Truex, R.L.; Dunk, J.R.; Gaman, T. 2006. Using forest inventory data to assess fisher resting habitat suitability in California. Ecological Applications. 16: 1010-1025.

Quantifying Forest Fragmentation Using Geographic Information Systems and Forest Inventory and Analysis Plot Data

Dacia M. Meneguzzo[1] and Mark H. Hansen[2]

Abstract.—Fragmentation metrics provide a means of quantifying and describing forest fragmentation. The most common method of calculating these metrics is through the use of Geographic Information System software to analyze raster data, such as a satellite or aerial image of the study area; however, the spatial resolution of the imagery has a significant impact on the results. Forest Inventory and Analysis (FIA) plot data also provide a way of quantifying fragmentation using measurements collected on the ground. In this study, the relationship between fragmentation metrics (total edge length, edge density, and forest proportion) calculated using FIA plot data and satellite imagery at two different spatial resolutions, 30 m and 250 m, is compared. Results for total edge length and edge density showed that estimates derived from the 30-m data were consistently larger than those from the FIA data, while estimates from the 250-m data were consistently smaller than those from the FIA data. For forest proportion, the percent forest values found using FIA plot data were very similar to those calculated using satellite imagery.

Introduction

Forest fragmentation is the breaking up of large, contiguous tracts of forest into smaller, more isolated patches. Fragmentation has many negative impacts on vegetation and wildlife; therefore, it is important that fragmentation be accurately quantified for management and monitoring purposes. This task can be accomplished through the use of fragmentation metrics, which are measurements that quantify and describe landscape pattern. The most common method of calculating fragmentation metrics is through the use of Geographic Information System (GIS) software to analyze raster data, such as a satellite or aerial image of the study area; however, the spatial resolution of the imagery has a significant impact on the results of the fragmentation metric calculations, and it is not known which spatial resolution produces the most accurate results. We believe that Forest Inventory and Analysis (FIA) plot data can offer some insight into this problem. In this study, we compare the relationship between fragmentation metrics calculated using FIA plot data and satellite imagery at two different spatial resolutions (30 m and 250 m). The fragmentation metrics include total edge length, edge density, and forest proportion. Fragmentation is indicated by longer total edge lengths, higher numbers of edge density, and lower amounts of forest proportion.

Data

Study Areas

Three study areas (fig. 1), ranging from heavily to sparsely forested, were selected in Michigan. These particular areas were chosen because they each are adjacent to one of the Great Lakes, do not contain any large tracts of Federal land, and, finally, do not contain any large urban or metropolitan areas. Study area 1 is Marquette County, which is located in the Upper Peninsula and is bordered by Lake Superior to the north; study area 2 is a group of three neighboring counties (Alpena, Montmorency, and Presque Isle) in the northeast portion of the Lower Peninsula and is bordered by Lake Huron to the east; and study area 3 is also a group of three neighboring counties (Berrien, Cass, and Van Buren) in the southwest corner of the Lower Peninsula and is bordered by Lake Michigan to the west.

[1] Forester, U.S. Department of Agriculture (USDA) Forest Service, Forest Inventory and Analysis Unit, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108. E-mail: dmeneguzzo@fs.fed.us.

[2] Research Forester, USDA Forest Service, Forest Inventory and Analysis Unit, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108. E-mail: mhansen01@fs.fed.us.

FIA Plots

The national plot configuration (fig. 2) consists of four circular subplots, each with a 24-ft radius. The centers of subplots 2, 3, and 4 are located 120 ft from the center of subplot 1, and the azimuths to subplots 2, 3, and 4 are 360, 120, and 240 degrees, respectively. The national plot configuration also requires mapping the different conditions (forest land, nonforest land, noncensus water, or census water) that occur on any of the four subplots if the area of the condition is at least 1 acre in size.

Plots from the first annual cycle of Michigan (2000 to 2004) were used in this study. More specifically, only plots where both forest and nonforest conditions occurred on the same subplot(s) were selected from the FIA database for the three study areas. Since both forest and nonforest conditions were present, this condition indicated that there was at least one forest/nonforest edge on each selected plot. Additionally, edges between forest and water and edges that occurred outside the subplot boundary were not included. Out of a possible 1,313 field plots, 150 were selected for this study based on the criteria described previously.

Imagery

The first set of satellite imagery was from the 1992 National Land Cover Dataset (Vogelmann *et al*. 2001) and has a spatial resolution of 30 m. This data set was previously corrected using a roads layer from the TIGER 2000 data (U.S. Census Bureau 2000) to update the urban class because a number of urban areas were initially classified as forest. The imagery was then reclassified into two categories, forest and nonforest, from the original 24 land cover classes.

The second set of satellite imagery was Moderate Resolution Imaging Spectroradiometer imagery with 250-m spatial resolution. Originally, the pixels in this imagery contained percent forest values ranging from 0 to 100, so a threshold of 36 percent was used to classify the pixels into forest and nonforest categories (Nelson *et al.* 2005). Pixels with values between 0 and 35 percent were classified as nonforest, while pixels with values greater than 35 percent were classified as forest.

Figure 1.—*Study area locations in Michigan (ESRI 2002).*

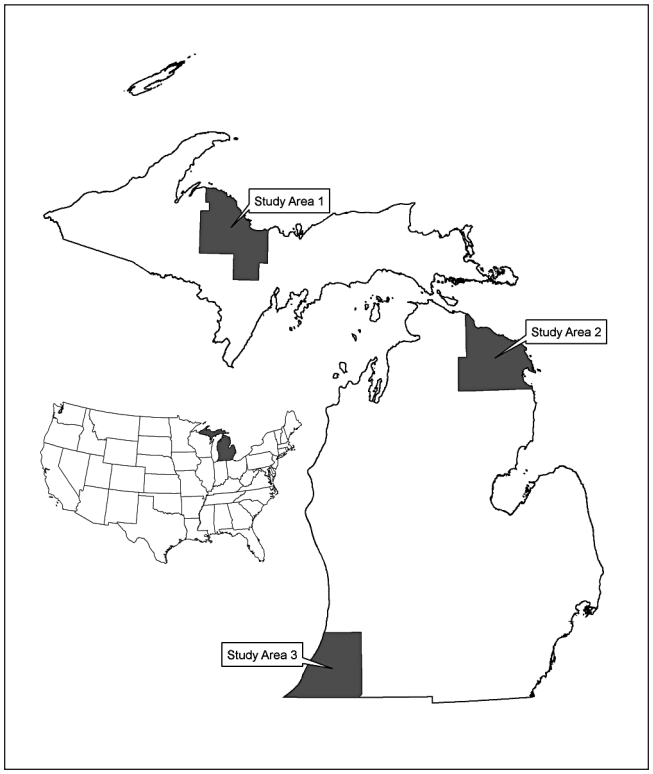
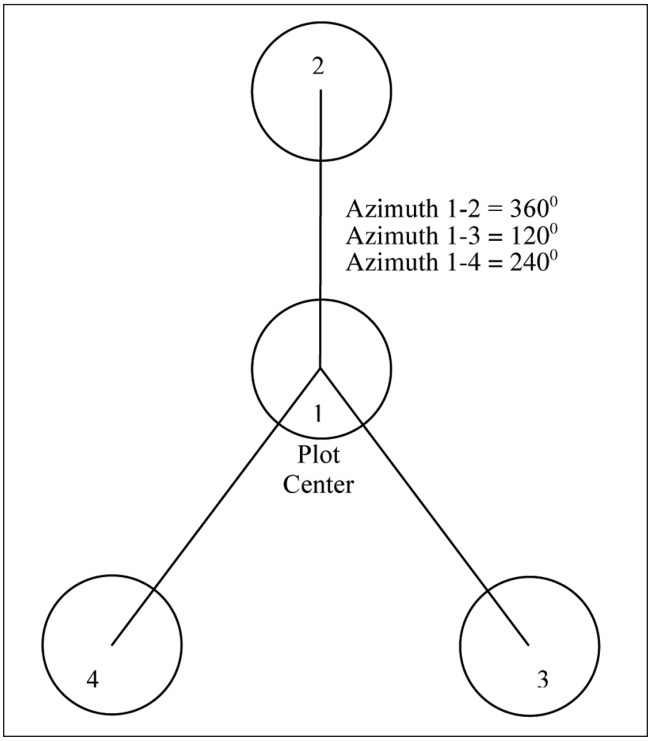


Figure 2.—*Forest Inventory and Analysis phase 2 plot configuration.*



Methods

Metric Calculations

Total Edge Length. The requirement to map different condition classes on subplots allows the boundary, or edge length, between the conditions to be obtained. For this study, however, the edge length was only obtained between forest and nonforest condition classes. This edge length may be measured as a straight line (fig. 3) or as two lines that meet at a corner (fig. 4). When two different conditions are encountered on a subplot and the edge is a straight line, the right and left azimuths are recorded, from subplot center, where the different conditions intersect the subplot circumference. Given that the radius of the subplot is fixed at 24 ft, the edge length has a maximum length of 48 ft. If the edge between the two conditions is not a straight line and contains a corner, however, a corner azimuth and a corner distance from the subplot center are also recorded. When this condition occurs, the total edge length may be longer than 48 ft, so, with the azimuth, radius, and/or corner distance information, the total edge length for both types of edges was calculated using the law of cosines:

$$a^2 = b^2 + c^2 - 2bc*\cos A$$

where a is the edge length and b and c are the length of the radii, or, when a corner is present, b and c are the lengths of the segments from the corner to the subplot circumference. The angle A is the angle in the triangle (opposite the straight-line edge) at the subplot center. When a corner was present, two angles were found and the law of cosines was used to find the two edge lengths, which were then summed to find the total edge length. After the edge lengths were obtained for each subplot, they were summed (if there was more than one subplot containing a forest/nonforest edge) to find the plot-level edge length total, which was then multiplied by the plot expansion factor. Finally, these expanded plot estimates were summed to find the total edge length for the study area.

To find the total edge length using the satellite imagery and GIS software, a short ARC Macro Language program was written to count the number of edges between the forest and nonforest pixels in the study area. Starting in the upper left corner of the image, all forest/nonforest pixel edges that occurred to the east and south of the subject center pixel were counted and the value was assigned to that pixel in an output grid. The numbers of edges in the output grid were summed and the total was multiplied by the spatial resolution to find the total edge length in meters. Meters were then converted to miles to reduce the large results and make it easier to compare the results.

Figure 3.—*How a straight boundary is measured on a subplot (USDA Forest Service 2005).*

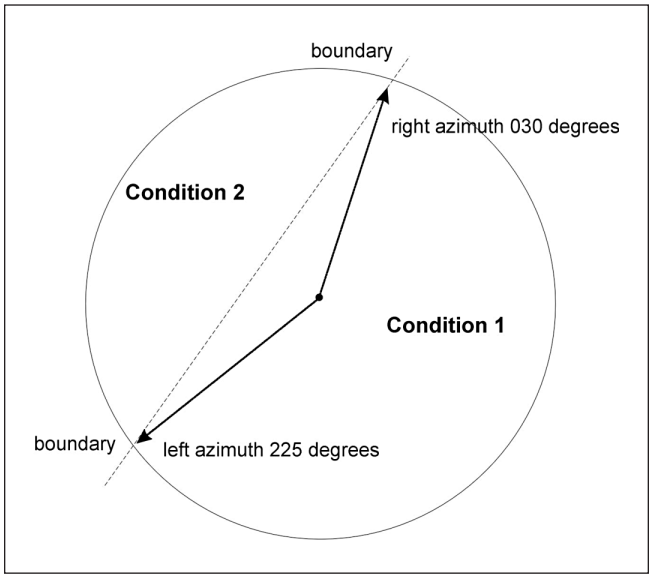
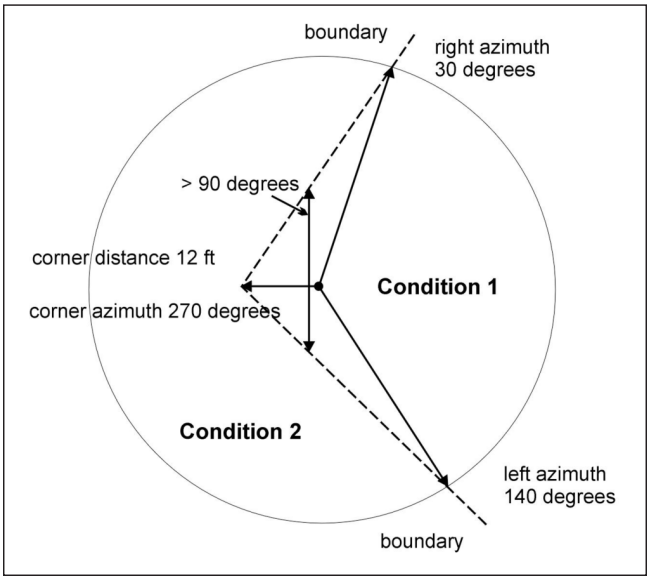


Figure 4.—*How a boundary with a corner is measured on a subplot (USDA Forest Service 2005).*



Edge Density. In an attempt to make the total edge length metric results more meaningful, they were converted to an edge density measurement: miles of edge per square mile of forest land. This procedure was accomplished by converting forest land area from acres to square miles. The total edge length in miles was then divided by the forest land area in square miles to obtain edge density.

Forest Proportion. The estimates of forest proportion found using FIA plot data were obtained by using two algorithms that expanded condition-level data (e.g., accessible forest, nonforest, noncensus water) from the FIA database to population estimates (Miles *et al.* 2001). Because each study area was considered a population, both algorithms were run for each area. The first algorithm calculated the area, in acres, of accessible forest land and the second calculated the total area of all land and noncensus water. The forest land area was then divided by the total area to find the proportion of forest. Estimates of forest proportion found using the satellite imagery were obtained by dividing the number of forest pixels by the total number of forest and nonforest pixels in each study area.

Results. The values obtained using FIA plot data and the satellite imagery approaches similarly separated the study areas into different levels of fragmentation through the use of the edge density and forest proportion metrics, but the approaches differed for total edge length (table 1). Study area 1, the most heavily forested, had the shortest total edge length. Study area 3, the most sparsely forested, had the longest total edge length according to the results found using the satellite imagery. The FIA plot data results showed that study area 2 had the longest total edge length (table 1), however. Overall, for all three study areas, the 30-m resolution imagery produced the longest edge lengths while the 250-m resolution imagery produced the shortest edge lengths.

In terms of edge density, in all cases, study area 1 had the lowest edge density and study area 3 had the highest edge density (table 2). Furthermore, the 30-m and 250-m images produced the highest and lowest results, respectively, as occurred with total edge length (table 2).

The results found for forest proportion (table 3) were similar for each study area using all three approaches. The proportion of forest in study area 1 differed by only 3 percent, but the FIA plot data and the 250-m imagery approaches produced the same result. The results for study area 2 were the most similar, and study area 3 had the widest range of results (table 3).

To further illustrate the degree of fragmentation in the three study areas, the percentages of field plots that were fragmented were also found by dividing the number of field plots that had forest and nonforest conditions by the total number of field plots in the study area. In study area 1, 6.4 percent of the field plots were fragmented. In study areas 2 and 3, 12.8 percent and 20.7 percent of the field plots, respectively, were fragmented.

Table 1.—*Total edge length metrics (miles) for three study areas in Michigan using three different approaches: (1) FIA plot data, (2) satellite imagery with 30-m spatial resolution, and (3) satellite imagery with 250-m resolution.*

Study area	FIA plot data	30-m resolution	250-m resolution
1	6908	14985	3578
2	11462	17912	6167
3	9135	23327	6710

FIA = Forest Inventory and Analysis.

Table 2.—*Edge density metrics (mile/mile2 of forest land) for three study areas in Michigan.*

Study area	FIA plot data	30-m resolution	250-m resolution
1	4.3	9.3	2.2
2	8.3	13.0	4.5
3	20.6	52.5	15.1

FIA = Forest Inventory and Analysis.

Table 3.—*Forest proportion metrics (percent) for three study areas in Michigan using three different approaches: (1) FIA plot data, (2) satellite imagery with 30-m spatial resolution, and (3) satellite imagery with 250-m resolution.*

Study area	FIA plot data	30-m resolution	250-m resolution
1	89	86	89
2	75	75	74
3	27	30	25

FIA = Forest Inventory and Analysis.

Summary

The objective of this study was to compare the fragmentation metric values calculated using FIA plot data with those calculated using satellite imagery obtained at two different spatial resolutions. Generally, the use of FIA plot data and 30-m and 250-m satellite imagery successfully separated the study areas into varying levels of fragmentation. In addition, forest proportion was similar among all approaches. The 250-m resolution imagery underestimated fragmentation, while the 30-m resolution imagery overestimated fragmentation, especially in study area 3, the most sparsely forested study area. This observation indicates that FIA plot data would produce other metric values similar to those found using satellite imagery with a spatial resolution somewhere between 30 and 250 m. Therefore, FIA data could potentially provide an answer to the question of which spatial resolution is the most appropriate for calculating fragmentation metric values. Accurate quantification of fragmentation would be valuable information that could be included in annual and 5-year State reporting.

Future work on this study will include calculating fragmentation metrics at a range of spatial resolutions, from less than 30 m. (e.g., 10 m) to resolutions between 30 and 250 m (e.g., 60 m, 90 m, 120 m), and will also include calculating additional metrics, such as average patch size. In addition, the relationship between metric values calculated using FIA plot data and those calculated at all of the various spatial resolutions will be further analyzed to find out exactly how spatial resolution impacts metric results. Finally, we hope this process will lead to useful conclusions about which spatial resolution is the most accurate and reliable for quantifying forest fragmentation.

Literature Cited

ESRI. 2002. ESRI data & maps 2002. ESRI white paper. Redlands, CA: ESRI. 34 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.

Nelson, M.D.; McRoberts, R.E.; Lessard, V.C. 2005. Comparison of U.S. forest land area estimates from forest inventory and analysis, national resources inventory, and four satellite image-derived land cover data sets. In: Proceedings, fifth annual forest inventory and analysis symposium. Gen. Tech. Rep. WO-69. Washington, DC: U.S. Department of Agriculture, Forest Service. 222 p.

U.S. Department of Agriculture (USDA) Forest Service. 2005. Forest inventory and analysis national core field guide. Version 3.0. Washington, DC: U.S. Department of Agriculture, Forest Service. 282 p. Vol. 1.

U.S. Department of Commerce, Census Bureau (U.S. Census Bureau). 2000. Census 2000 TIGER/Line files [machine-readable data files]. Washington, DC: U.S. Department of Commerce, Census Bureau. http://www.census.gov/geo/www/tiger/tiger2k/tgr2000.html. (17 September 2003).

Vogelman, J.E.; Howard, S.M.; Yang, L.; Laron, C.R.; Wylie, B.K.; Van Driel, N. 2001. Completion of the 1990s National Land Cover Dataset for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. Photogrammetric Engineering and Remote Sensing. 67: 650-662.

The Relationship Between Diversity and Productivity in Selected Forests of the Lake States Region (USA): Relative Impact of Species Versus Structural Diversity

W. Keith Moser[1] **and Mark Hansen**[2]

Abstract.—Ecological theory suggests that diversity and productivity (at some measure) are positively correlated, presumably because individuals engage in niche partitioning to occupy any unclaimed growing space. We examined this theory using inventory information from the U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis program. The study uses plot-level data from inventories of the Lake States region conducted between 1999 and 2005. Relationships between diversity and biomass productivity in two contrasting forest cover types—aspen and sugar maple-beech-birch—were examined using two measures of diversity. We expected aspen productivity to increase with increasing species diversity and sugar maple-beech-yellow birch productivity to increase with increasing height diversity, reflecting the niche occupancy of each type and reasonable strategies for differentiation. On a landscape level, matching diversity and productivity would better allow us to predict change as a result of management actions, ecological succession, or other factors.

Neither hypothesis was supported by the data. For aspen stands, the increase in competing, more shade-tolerant species apparently constrained aspen growth more than they added to total stand productivity. For sugar maple-beech-yellow birch stands, the freeing of growing space due to height differentiation did not result in a dramatic response by the individual trees capable of occupying that space. Forest managers faced with mandates to enhance biodiversity while maintaining productivity must be aware of what is and is not possible. Our results suggest that there are fundamental limitations in how managers can simultaneously manage for these two attributes.

Introduction

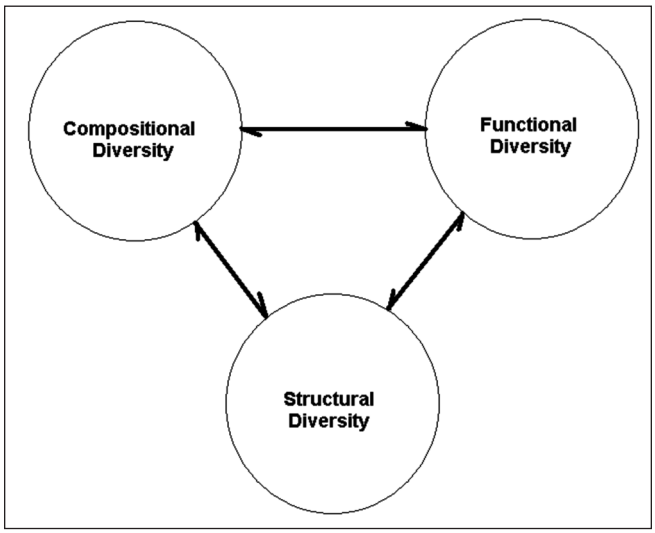
Forests may represent the most diverse ecosystems in the world. Forests provide a wide array of goods and services, both timber and nontimber resources, and play a critical role in carbon storage. Simultaneously, forests provide livelihoods and social and cultural benefits for millions of people throughout the world (Convention on Biological Diversity 2005). Biodiversity has long been considered an important concept in the analysis of ecosystems. Forests are critical areas for managing and protecting biodiversity (Probst and Crow 1991). These ecosystems are defined by the presence and composition of the tree species that inhabit them. A mental construct of such a system might start with the picture of the trees as the "skeleton" upon which one hangs the many attributes of the ecosystem, whether they are wildlife habitat, carbon storage, or nutrient cycling (Odum 1965). To paraphrase, a more diverse forest may offer many different pathways for each of these "skeletal ornament" systems, thus allowing a certain amount of redundancy in the face of disturbances, expected and unexpected.

Although diversity has frequently been described by measures of species or age, other accepted metrics, including ecosystem function, spatial arrangement, and height and diameter differentiation, exist (fig. 1). All have been shown to add value to the ecosystem in some way (Crow *et al.*, 1994). Spies (1998) also stated that particularly important components of forest structure include (1) tree size/age distribution, (2) vertical foliage distributions, (3) horizontal canopy distribution, and (4) dead wood.

[1] Corresponding Author, U.S. Department of Agriculture (USDA) Forest Service, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108. Phone: 651–649–5155. E-mail: wkmoser@fs.fed.us.

[2] Research Forester, USDA Forest Service, Forest Inventory and Analysis Unit, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108. E-mail: mhansen01@fs.fed.us.

Figure 1.—*Components of ecological diversity (from Crow* et al. *1994)*.



The public strongly supports maintaining and enhancing biodiversity, an important component of forest health (Patel *et al.* 1999, but see Rapport 1998). Forest managers are expected to consider the impacts of management actions upon biodiversity. Public land managers, in particular, are susceptible to pressure in this regard.

Forest management has traditionally focused on productivity as a measure of success (Assmann 1970). Although past generations may have defined productivity in terms of some economically useful product, or some other "output" such as water yield or wildlife populations, current thinking bases measures of productivity on some measurable quantity, such as volume or biomass. Regardless of how one measures it, productivity is still used as a yardstick for evaluating the progress in meeting management goals.

Given the public interest in and value for diversity, and the use of productivity as a scale against which we measure success, we ask, "What exactly are the tradeoffs, and how do they vary between different geographic, ecosystem, or management situations? How can the results from an investigation into these relationships be translated into reasonable, understandable management guidelines?"

To begin to answer these questions, we chose two cover types common in the Lake States region: aspen and sugar maple-beech-yellow birch. The aspen forest type is composed primarily of quaking aspen (*Populus tremuloides*) and/or bigtooth aspen (*Populus grandidentata*), both of which are shade-intolerant (Baker 1949), relatively short-lived species that historically existed on the forest-prairie ecotone, which was subjected to frequent fires (Laidly 1990, Perala 1990). Disturbance, old-field abandonment, and deliberate management action have resulted in the species' establishment throughout the region. The sugar maple-beech-yellow birch forest type (often referred to as "northern hardwoods") is a late successional forest type found in the cold-temperate forests of eastern North America (Eyre 1980). Primarily composed of sugar maple (*Acer saccharum*), American beech (*Fagus americana*), and yellow birch (*Betula allegheniensis*), this forest type is characterized by more shade-tolerant (Baker 1949), longer lived species than the aspen forest type and generally thrives on more mesic, richer sites that experience frequent but low-severity disturbances. In terms of forest type longevity and cohesion, and the physiological performance required by inhabitants of their relative ecological niches, these two forest types should present stark contrasts in their stand dynamics and response to variations in composition, structure, and disturbance.

Literature

Diversity and Forest Management

Many studies have examined the relationship between species mixture and productivity. Pretzsch (1997) looked at a series of stands ranging from pure Norway spruce to pure beech and different combinations of the two. He found that mixed stands are able to compensate for density reductions due to thinning by increased growth within the remaining stand. The greater elasticity of growth against a reduction in stand density can be attributed to the multilayer stand structure in mixed stands.

Any positive or negative deviations from average stand density, due to lack of treatment or unplanned disruption of stand canopy, could be more easily buffered in mixed stands. Kelty (1986) looked at productivity of mixed hardwood stands and

mixed hardwood-hemlock stands. He found that stands with hemlock were more productive without any material reduction in growth of the hardwoods. Edgar and Burk (2001) examined stands in northeastern Minnesota with a range of species mixtures. Although they expected pure aspen stands to be the most productive, they found that the two most productive stands were both vertically stratified aspen-balsam fir-paper birch mixtures.

Using forest inventory and analysis (FIA) data, Caspersen and Pacala (2001) found that successional composition has a significant relationship to forest ecosystem function: early successional stands were more productive than late successional stands. They concluded that successional diversity was positively correlated with productivity, as was species diversity. This result suggests that a forest managed in such a way as to maintain species diversity would maximize the function of interest.

Structural diversity in forested ecosystems has also been studied. MacArthur and MacArthur (1961) documented the relationship between height diversity and bird species diversity. Proponents of "new forestry" have pointed to diversity in tree size and canopy structure as one advantage of old growth forests (Franklin *et al.* 2002). For shade-intolerant forest types, such as longleaf pine, structural diversity can reduce overall productivity (Farrar 1996), but even in these situations, managed-forest structural diversity maintains the ecological benefits largely unchanged over time (Moser *et al.* 2002). In forests where species exhibit substantial niche overlap, structural diversity and temporal advantage can dictate both the species' proportion and the overall productivity (Oliver and Larson 1996). One study found a pattern of reciprocal replacement, suggesting the importance of structural diversity and species location in maintaining forest composition (Woods 1979).

Growing space is a holistic concept encompassing the availability of factors influencing growth, including sunlight, water, nutrients, oxygen, and carbon dioxide (Oliver and Larson 1996). Where one factor is limiting, its availability controls growth and the ability of a tree to use another resource (Assmann 1970). Tree species can exist on the same site if their minimum requirements and opportunities for capture of growing space differ (Oliver and Larson 1996).

Diversity Measurement Tools

In this article, we will examine species composition and aspects of the vertical and horizontal distributions within the forest. Metrics that quantify structural and species diversity in forests have existed for over a century. Some are distance independent and manipulate counts of species or sizes ("mixtures"). Some are distance dependent on an ordinal scale, where they compare a subject tree to some number of "nearest" trees ("differentiation"). Finally, there are quantitative distance-dependent methods ("positioning") that require knowledge of the actual positions of each tree in a group of trees.

Mixtures can be evaluated with counts, such as species richness; proportional measures, such as importance factors; or its opposite, evenness; or some combination of both, such as the Shannon or Simpson indices (Magurran 1988). In this study we focused on the Shannon index

$$H' = -\sum_{i=1}^{n} p_i * \ln(p_i)$$

where p_i is the proportion of the total number of individuals that belong to a particular category.

Mixtures can also be evaluated with some summary of categorical variables. One example of this type of tool is a measure of mingling (v. Gadow 1999). Mingling measures are based on the proportion of trees (usually totaling three) with characteristics different from the selected sample tree (Graz 2006). To evaluate the extent of horizontal, distance-independent diversity, we calculated mingling indices for species and height at the subplot level. The species mingling index M_i for a given sample tree, i, using n neighbors, is defined as

$$M_i = \frac{1}{n}\sum_{j=1}^{n} m_{ij}$$

where:

$$m_{ij} = \begin{cases} 1, & \text{if the tree is of another species} \\ 0, & \text{if the tree is of the same species} \end{cases}$$

If three neighbors are used to determine M_i, the index may be one of four possible values:

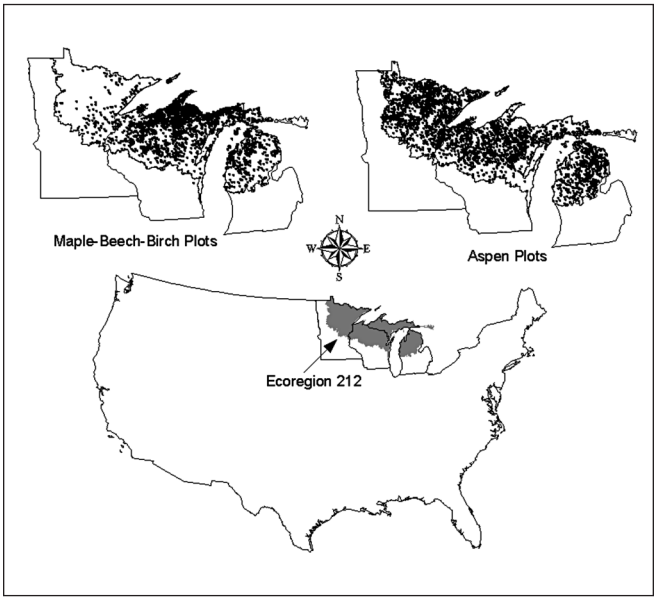
1. 0/3 if none of the neighbors is of a different species.
2. 1/3 if one of the neighbors is of a different species.
3. 2/3 if two of the neighbors are of a different species.
4. 3/3 if all of the neighbors are of a different species.

Mingling values close to 0 suggest that trees of the reference species or height category occur in groups, implying a high degree of aggregation. Higher values of mingling closer to 1 suggest that trees of the reference species or height category do not occur together (Graz 2006).

Objectives

This study examines the relationship between species and height diversity in two representative forest cover types: early successional aspen stands and late successional sugar maple-beech-yellow birch stands, in Ecoregion 212 (fig. 2). The forest types chosen represent endpoints on both single species: multiple species and shade-intolerant/shade-tolerant gradients for this ecoregion.

Figure 2.—*Location of Ecoregion 212 and sample plots for the forest types studied before the truncation of the data set; see the "Methodology" section of this article).*



Hypotheses

The following hypotheses were evaluated:

1. H_0: There is no difference in productivity due to species or height diversity.
2. H_1: High species diversity in aspen stands is associated with high productivity.
3. H_2: High height diversity in sugar maple-beech-yellow birch stands is associated with high productivity.

Methodology

FIA Data Set

The plot and tree measurements presented here come from paired observations from plots located in Ecoregion 212 in Michigan, Minnesota, and Wisconsin, with final measurements taken from 1999 to 2005. All plots consist of an initial measurement followed by a remeasurement of the plot 5 to 14 years later that accounts for all live trees measured initially and any new trees that have grown onto the plot. Growth and mortality for a plot are recorded based on changes observed over the remeasurement period. Plots are characterized for diversity, age, and current volume based on the second measurement. In cases where three measurements of a plot were available (e.g., a plot measured in 1990, 2000, and 2005), the first two measurements (e.g., 1990–2000) were treated as one plot and the second two measurements (e.g., 2000–2005) were treated as a second independent plot in our data set. These repeated plots at the same location account for 13 percent of the aspen and 23 percent of the sugar maple-beech-birch plots. We restricted measures of diversity to include only live trees on single-condition plots that were entirely on timberland and had not had any removals over the remeasurement period. The measure of productivity was the ratio of all live growth on timberland to current all-live volume on timberland. Complete documentation of the plot design and all measurements can be found at http://socrates.lv-hrc.nevada.edu/fia/dab/databandindex.html.

Tree measurements used in this study include species and height (total tree height measured to the nearest foot) and tree location (determined through distance and azimuth measure-

ments from plot center). Height was measured only on trees 5 in in diameter and larger.

Measures of relative growth rates and density are affected by the life-history strategy of the species and the relative stages of stand development. To eliminate the influence that different relative lifespans, high juvenile growth rates, or density declines with advanced age might have on our analysis, we truncated our original data set by removing the oldest 20 percent and youngest 20 percent for each forest type. This tactic reduced the number of aspen forest-type plots from 3,190 to 1,914 and the number of sugar maple-beech-yellow birch plots from 2,440 to 1,464.

We speak of "productivity," but it is actually a relative growth rate (Blackman 1919), or the annual growth rate ($ft^3\ ac^{-1}\ yr^{-1}$) divided by current-year volume ($ft^3\ ac^{-1}$). The advantage of this technique is that it matches growth to the volume rather than giving undue weight to stands with large volumes (which might have large growth). One disadvantage is that above-average growth rates occur in the early years of a stand, when they often have lower diversity values (particularly in aspen stands); so there is a confounding of age and diversity. Also, physiologically, we should be comparing active sapwood volume, not total volume, to growth. Still, there is a correlation (albeit lagged in time) between sapwood volume and total volume, and measuring volume is one real-world benchmark of stand performance.

Measures of Diversity

We used the Shannon index to calculate species diversity and diversity of categories of height. To convert a continuous variable-like height into a categorical one, we assigned heights to 3-ft classes. Each class was then treated in the same manner as species in the computation of the index. We calculated the mingling index for each subplot and then averaged the four values for each plot. We used analysis of variance of selected stand structure variables—basal area and stand density index (SDI), stand age, aspect and diversity indices—at the 95-percent level to test our hypotheses.

Results

Shannon Index

Our results suggested that measures of diversity—H' species and H' height—along with stand age and measures of stand density show a significant relation to productivity.[3] When we analyzed the data separately by forest type, we continued to observe a significant relationship between productivity and stand age, basal area, and H' species.[4] The sugar maple-beech-yellow birch data did not display a significant relationship with H' height where aspen did.[5]

Figure 3 displays growth percentage compared with the Shannon index for the two forest types. Growth percentage for aspen forest types was highest at the lower diversity classes, the opposite of H_1, and declined significantly as diversity increased. Sugar maple-beech-yellow birch productivity did increase slightly as diversity increased, with the two lowest classes being significantly less than the three highest diversity classes. The two forest types' productivity levels were not significantly different at the highest diversity level. Figure 4 displays a steady, significant decline in aspen productivity with increasing diversity. Sugar maple-beech-yellow birch productivity did not change significantly across height-diversity categories.

Mingling Index

As opposed to the Shannon index, which is a plot-level calculation, the mingling index reflects the association of each individual tree with its three nearest neighbors. We expected that fine-scale heterogeneity would be more easily detectable using this technique. Overall, the results of our analysis of variance found significant relationships between productivity and overstory density (both basal area and SDI), stand age, and both

[3] Basal area (F = 42.3, p(F) = 0), stand age (F = 204.3, p(F) = 0), SDI (F = 10.4, p(F) = 0), H' species (F = 8.7, p(F) = 0.03), H' height (F = 36.7, p(F) = 0).

[4] Aspen: basal area (F = 17.3, p(F) = 0), stand age (F = 49.3, p(F) = 0), SDI (F = 7.6, p(F) = 0.01), H' species (F = 7.9, p(F) = 0). Sugar maple-beech-yellow birch: basal area (F = 3.9, p(F) = 0.05), stand age (F = 4.8, p(F) = 0.03), SDI (F = 9.5, p(F) = 0), H' species (F = 6.6, p(F) = 0.01).

[5] Aspen: H' height (F = 20.7, p(F) = 0). Sugar maple-beech-yellow birch: H' height (F = 0.25, p(F) = 0.615).

mingling indices (species and height).[6] When only aspen stands were examined,[7] we found significance in the relationships between productivity and density, and age and the mingling index for species. For sugar maple-beech-yellow birch stands,[8] however, only density (overstory basal area and SDI) and stand age displayed significant relationships with productivity.

Productivity for aspen stands was significantly larger than sugar maple-beech-yellow birch stands at corresponding mingling index levels except for the highest category (fig. 5). The differential between the two forest types declined as the index increased. Across mingling indices for height, aspen productivity was significantly larger than sugar maple-beech-yellow birch productivity (fig. 6), although, as in species mingling,

Figure 3.—*Growth as a percentage of current live volume compared with H' species, by forest type.*

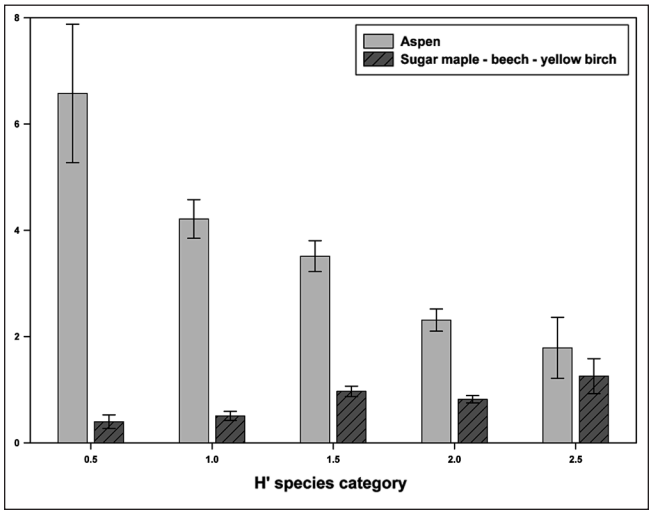


Figure 4.—*Growth as a percentage of current live volume compared with H' height, by forest type.*

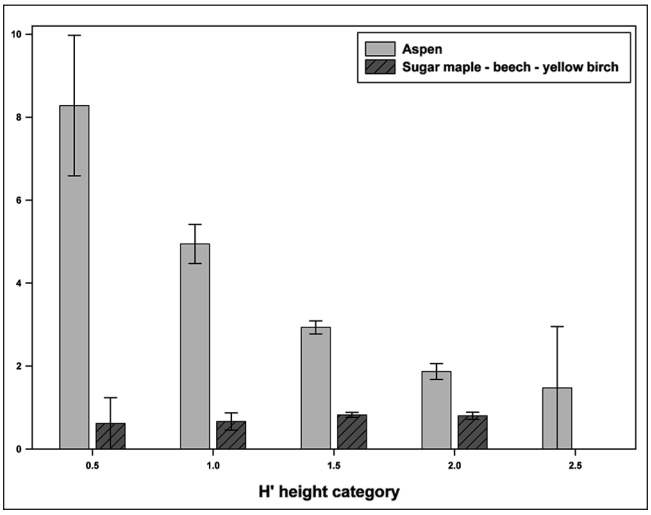


Figure 5.—*Productivity (growth/current volume) compared with mingling index for species, by forest type.*

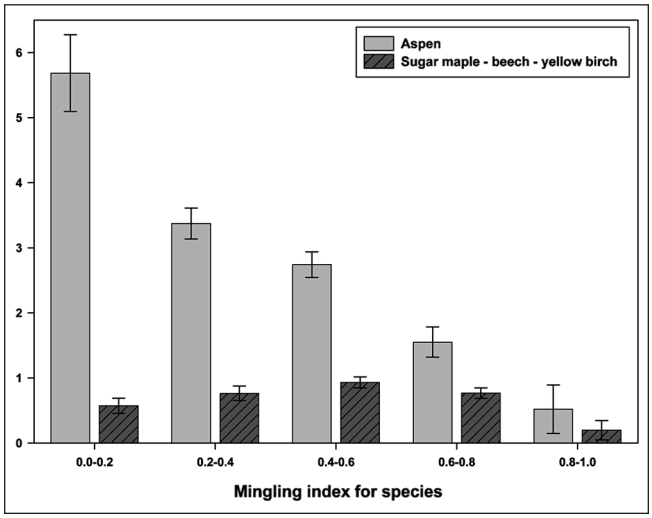
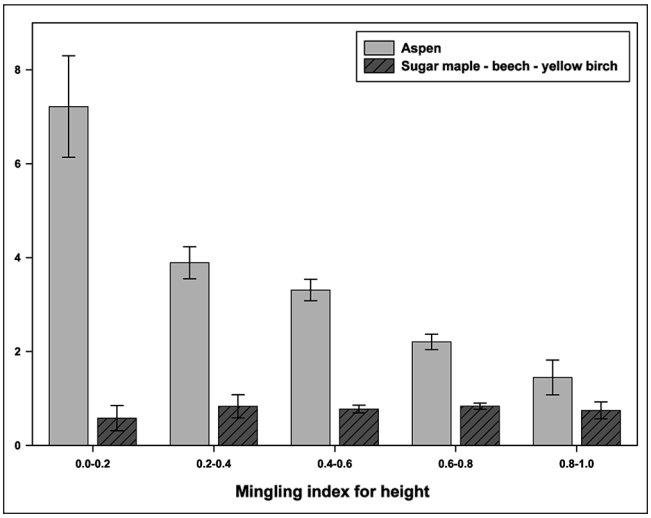


Figure 6.—*Productivity compared with mingling index for height, by forest type.*



[6] Overstory basal area (F = 42.0, p(F) = 0), stand age (F = 202.9, p(F) = 0), SDI (F = 10.4, p(F) = 0), Mi-species (F = 16.4, p(F) = 0), Mi-height (F = 6.3, p(F) = 0.01).

[7] Aspen: overstory basal area (F = 17.2, p(F) = (0)), stand age (F = 48.9, p(F) = 0), SDI (F = 7.5, p(F) = 0), Mi-species (F = 12.5, p(F) = 0).

[8] Sugar maple-beech-yellow birch: overstory basal area (F = 3.9, p(F) = 0.05), stand age (F = 4.8, p(F) = 0.03), SDI (F = 9.4, p(F) = 0).

the difference between the two forest types declined as the mingling index increased.

Discussion

Our original hypotheses were that (1) increased species diversity would increase productivity in aspen stands, and (2) increased height diversity would increase productivity in sugar maple-beech-yellow birch stands. Neither hypothesis was supported by our data. Aspen productivity declined as the Shannon index for species diversity increased. Although aspen productivity was usually higher than sugar maple-beech-yellow birch productivity across species diversity levels, the relative advantage declined as species diversity increased. Apparently, the available growing space in aspen stands is occupied by species that, in time, suppress more aspen growth than they contribute with their own growth. Similar relationships have been observed in mixed pine-hardwood stands in the southern United States, where the hardwood presence suppressed the growth of loblolly pine (Miller *et al.* 2003).

Our second hypothesis was that height diversity would increase sugar maple-beech-yellow birch productivity because the close-niche overlapping northern hardwoods would stagnate without some type of differentiation or release. Our data did not support this hypothesis. Apparently, the niche overlap was not as constricting as we first thought; the stand was not influenced by the increase in height diversity. Another explanation might be that the available growing space in a mixed-height stand was not captured by the remaining trees as dramatically as we had expected.

The mingling indices provided a smaller scale analysis of mixtures than did the plot-level Shannon indices. Mingling should be more sensitive than Shannon to clumping at scales smaller than subplot level. Productivity declined for aspen as M_i-species and M_i-height increased. The sugar maple-beech-yellow birch results for species were not significantly different with respect to the mingling index. Aspen productivity was generally higher than sugar maple-beech-yellow birch productivity. As the mingling indices are ordinal in their tree-to-tree categorization, we had difficulty discovering the influences of mingling compared with measures of density, although analyses of variance suggested density was a significant influence on productivity.

Other explanations of the trends include noise in the data set and unaccounted-for variables. For example, on highly productive sites, aspen may have gotten a good start; the species dominates the site and grows quickly. On other sites, where multiple species occurred during regeneration (possibly due to poor stocking), growth is slow. Also, we were measuring volume, not biomass, which may not account for the differences in converting carbohydrates to a unit of volume at different densities. There was a suggestion that we did not appropriately account for the different sites on which each forest type is traditionally found (i.e., aspen is found on more northerly and colder sites than the sugar maple-beech-yellow birch plots) (Frelich 2007). Although we believed that by restricting our analyses to Ecoregion 212 we might have sufficiently controlled for climate, further analysis with a larger data set might allow us to better explore this issue.

We did observe trends, although not statistically significant, that suggested middle-level diversity in both species and height provided an optimal situation for stand-level productivity in sugar maple-beech-yellow birch stands. Further investigation with a larger data set, perhaps one with linked remeasurements, is needed.

We tried to control for age by eliminating portions of the data at either end of each forest type's age-class distribution. Nevertheless, we did observe trends where diversity increased with age. These trends were particularly visible in aspen stands.

Other Measures Available

Other methods are available to explore some of the relationships we examined in this study.

Diameter Based. The Shannon index for diameters has the advantage of historical validation (deLiocourt 1898), ease of measurement compared to height (Avery and Burkhart 2002), and correlation with other measures of tree size. Although diameter is easier to measure than tree height and is frequently correlated with this variable, diameter growth is lower in priority

than height growth for a tree's allocation of carbohydrates (Kozlowski *et al.* 1991) and therefore presents an imperfect mirror to the structural (height) diversity we seek. In a sense, diameter diversity is more a result of partitioning of growing space than a creator of these partitions (i.e., it is a "post-hoc" measure).

Distance Dependent-Quantitative. Analysis of distance-dependent relationships provides an apparently more precise evaluation of tree-to-tree influences. Some sort of pair-correlation function might be called for here, but an index does not tell us about within-plot clumping. (Are the trees grouped with equal numbers of near and far trees, or are they equally spaced?) Furthermore, edge correction, particularly on the circular FIA plots, is quantitatively challenging and, depending on the radius of the plot, a source of considerable bias.

Clumping. These methods provide measures of clumping, but do not consider differences in particular species or structural diversity. To do this, one must run the calculation separately for different species or size classes.

Aggregation Index. (Clark and Evans 1954) This index compares the mean distances between each tree in a forest stand with an expected mean distance in a stand with random tree locations.

Neighborhood or Contagion Index. (v. Gadow 1999).The neighborhood index looks at the expected equal size of the number of angles framed by *n* trees, and determines the index as a proportion of the angles less than the equal size. For example, a neighborhood with four trees would have a standard angle size of 90 degrees. The contagion index would calculate the proportion of the angles that are less than 90 degrees. This method necessitates the use of a distance-dependent measure and/or an estimate of density (Graz 2006).

Acknowledgments

The authors thank Lee Frelich, Manfred Mielke, Will McWilliams, and Cynthia Moser for their thoughtful analyses and comments on earlier versions of this manuscript.

Literature Cited

Assmann, E. 1970. The principles of forest yield study. New York: Pergamon Press. 506 p.

Avery, T.E.; Burkhart, H.E. 2002. Forest measurements. Boston: McGraw-Hill. 456 p.

Baker, F.A. 1949. A revised tolerance table. Journal of Forestry. 47(3): 179-181.

Blackman, V.H. 1919. The compound interest law and plant growth. Annals of Botany. 33: 353-360.

Caspersen, J.P.; Pacala, S.W. 2001. Successional diversity and forest ecosystem function. Ecological Research. 16: 895-903.

Clark, P.J.; Evans, F.C. 1954. Distance to nearest neighbor as a measure of spatial relationships in populations. Ecology. 35(4): 445-453.

Convention on Biological Diversity. 2006. Convention on biological diversity: implementation in the European Union. Luxembourg: Office of official publications of the European Union. 28 p.

Crow, T.R.; Haney, A.; Waller, D.M. 1994. Report on the scientific roundtable on biological diversity convened by the Chequamegon and Nicolet national forests. Gen. Tech. Rep. NC-166. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station. 55 p.

deLiocourt, F. 1898. De l'amenagement des sapinieres. Bulletin trimestriel, Société forestière de Franche-Comtè et Belfort, juillet 1898, pp. 396-409. English translation Nygren, M.; Nygren, P.; and Larsen, D.R. 2001. On the amelioration of fir forests. (Revised 27 February 2002.)

Edgar, C.B.; Burk, T.E. 2001. Productivity of aspen forests in northeastern Minnesota, U.S.A., as related to stand composition and canopy structure. Canadian Journal of Forest Research. 31(6): 1019-1029.

Eyre, F.H., ed. 1980. Forest cover types of the United States and Canada. Washington, DC: Society of American Foresters. 148 p.

Farrar, R.M. 1996. Fundamentals of uneven-aged management in southern pine. In: Moser, W.K.; Brennan, L.A., eds. Misc. Pub. 9. Tallahassee, FL: Tall Timbers Research Station. 66 p.

Franklin, J.F.; Spies, T.A.; Van Pelt, R.; Carey, A.B.; Thornburgh, D.A.; Berg, D.R.; Lindenmayer, D.B.; Harmon, M.E.; Keeton, W.S.; Shaw, D.C.; Bible, K.; Chen, J. 2002. Disturbances and structural development of natural forest ecosystems with silvicultural implications, using Douglas-fir forests as an example. Forest Ecology and Management. 155: 399-423.

Frelich, L. 2005. Personal communication. February 21, 2007.

Graz, F.P. 2006. Spatial diversity of dry savanna woodlands. Biodiversity and Conservation. 15: 1143-1157.

Kelty, M. 1986. Productivity of New England hemlock/hardwood stands as affected by species composition and canopy structure. Forest Ecology and Management. 28(3-4): 237-257.

Kozlowski, T.T.; Kramer, P.J.; Pallardy, S.G. 1991. The physiological ecology of woody plants. San Diego: Academic Press. 390 p.

Laidly, P.R. 1990. *Populus tremuloides*. In: Burns, R.M.; Honkala, B.H., tech. coords. Silvics of North America, vol. 2, Hardwoods. Agriculture Handbook No. 654. Washington, DC: U.S. Department of Agriculture, Forest Service. 877 p.

MacArthur, R.H.; MacArthur, J.W. 1961. On bird species diversity. Ecology. 42(3): 594-598.

Magurran, A.E. 1988. Ecological diversity and its measurement. Princeton, NJ: Princeton University Press. 179 p.

Miller, J.H.; Zutter, B.R.; Zedaker, S.M.; Edwards, M.B.; Newbold, R.A. 2003. Growth and yield relative to competition for loblolly pine plantations to midrotation: a southeastern United States regional study. Southern Journal of Applied Forestry. 27: 1-16.

Moser, W.K.; Jackson, S.M.; Podrázský, V.V.; Larsen, D. 2002. Examination of stands on quail plantations in the Red Hills region of Georgia and Florida managed by the Stoddard-Neel System: an example for forest managers. Forestry (Oxford). 75(4): 443-449.

Odum, E.P.; Marshall, S.G.; Marples, T.G. 1965. The caloric content of migrating birds. Ecology. 46(6): 901-904.

Oliver, C.D.; Larson, B.C. 1996. Forest stand dynamics. Updated edition. New York: McGraw Hill. 540 p.

Patel, A.; Rapport, D.J.; Vanderlinden, L.; Eyles, J. 1999. Forests and societal values: comparing scientific and public perception of forest health. Environmentalist. 19: 239-249.

Perala, D.A. 1990. *Populus tremuloides*. In: Burns, R.M.; Honkala, B.H., tech. coords. Silvics of North America, vol. 2. Hardwoods. Agriculture Handbook No. 654. Washington, DC: U.S. Department of Agriculture, Forest Service. 877 p.

Pretzsch, H. 1997. Analysis and modeling of spatial stand structures. Methodological considerations based on mixed beech-larch stands in Lower Saxony. Forest Ecology and Management. 97: 237-253.

Probst, J.R.; Crow, T.R. 1991. Integrating biological diversity and resource management. Journal of Forestry. 89(1): 12-17.

Rapport, D.J. 1998. Biodiversity and saving the earth. Environmental Monitoring and Assessment. 49: 169-175.

Spies, T.A. 1998. Forest structure: a key to the ecosystem. Northwest Science. 72(2): 34-39.

v. Gadow, K. 1999. Waldstruktur and Diversität (Forest structure and diversity). Allg. Forst-u Jagd-Zeitung. 170: 117-122.

Woods, K.D. 1979. Reciprocal replacement and the maintenance of codominance in a beech-maple forest. Oikos. 33: 31-39.

Additional Reading

Pommerening, A. 2002. Approaches to quantifying forest structures. Forestry (Oxford). 75(3): 305-324.

Estimating Number and Size of Forest Patches From FIA Plot Data

Mark D. Nelson[1], Andrew J. Lister[2], and Mark H. Hansen[1]

Abstract.—Forest inventory and analysis (FIA) annual plot data provide for estimates of forest area, type, volume, growth, and other attributes. Estimates of forest landscape metrics, such as those describing abundance, size, and shape of forest patches, however, typically are not derived from FIA plot data but from satellite image-based land cover maps. Associating image-based land cover metrics with FIA plot-based attributes is problematic due to differences in definitions between FIA land use and image-based land cover, temporal inconsistencies between plot and image acquisitions, and spatial misregistration between plots and map pixels. We assess an existing approach for using FIA field plot data directly for estimating the number and mean size of forest patches within estimation units typically reported by FIA (e.g., counties, States, or other geographic extents of moderate to large area). Comparison analyses reveal that FIA plot-based estimates of mean patch size are larger, and estimates of the number of patches are smaller than estimates derived from a satellite image-based land cover map.

Introduction

Scientists in the Forest Inventory and Analysis (FIA) program developed a standard plot and sample design to produce nationally consistent estimates of forest area, volume, and other attributes over moderate- to large-sized estimation units, (e.g., counties and States) (Bechtold and Patterson 2005). These estimates provide useful information about many forest resources, including amount of wildlife habitat. The suitability of forest land for wildlife habitat is dependent not only upon local characteristics and total area but also on the landscape pattern of forest habitat. For example, some neotropical migratory bird species require patches of forest habitat measuring at least several hundred hectares in area (Wenny *et al.* 1993). The number and size of individual forest patches and other metrics of landscape pattern typically are not estimated from FIA plot data, however (Riemann *et al.* 2003); such metrics are routinely derived from land cover maps by using spatial pattern analysis computer programs (Lister *et al.* 2003, McGarigal and Marks 1995). Associating independently acquired satellite image-based landscape metrics with FIA plot attributes is problematic due to differences in definitions between FIA forest land use and image-based forest land cover, temporal inconsistencies between plot and image acquisitions, and spatial misregistration between plots and map pixels. Thus, it is necessary to obtain estimates of landscape metrics that are more consistent with FIA estimates of standard forest attributes.

Although not designed specifically for estimating spatially explicit attributes such as landscape metrics, FIA plots do contain spatial information inherent in inventory cluster plot designs. This spatial information occurs within subplots, between subplots, and between plots (Kleinn 2000). Using between-subplot spatial information, inventory cluster plots can provide for estimates of the number and size of forest patches and the length and area of forest/nonforest buffers (Kleinn 2000). Kleinn (2000) described an approach for estimating metrics of landscape pattern, involving the assignment of subplot center points to forest or nonforest class, or other classes of interest. These estimates are most meaningful when the size of cluster plots is smaller than the size of forest patches and smaller than the distance between patches (Kleinn 2000), a requirement generally met by FIA's plot design. Van Deusen (2005) described a similar approach for estimating average patch size; he stated, "The possibility of making patch size estimates is unique to the mapped plot design." Kleinn (2000), however, previously demonstrated that nonmapped cluster plots provide for estimates of landscape metrics, including patch size.

[1] Research Forester, U.S. Department of Agriculture (USDA) Forest Service, Northern Research Station, St. Paul, MN.
[2] Research Forester, USDA Forest Service, Northeastern Research Station, Newtown Square, PA.

Because FIA cluster plots are mapped, they can be used with either Kleinn's (2000) or Van Deusen's (2005) approach, both of which require an assumption of typical patch shape. Kleinn's (2000) approach also allows for the analysis of other inventory data obtained from nonmapped cluster plots.

In this article, we apply Kleinn's (2000) approach to produce estimates of forest patch size and number, using nonmapped data from FIA annual inventory cluster plots. We then compare these FIA plot-based estimates with satellite image-based pixel estimates of forest patch size and the number of patches.

Data and Methods

FIA Plot Data

The study was conducted in Michigan, USA, a State that has 83 counties and 14.7 million ha of land area. According to FIA estimates from Michigan's first annual inventory (2000 to 2004), approximately 7.8 million ha of forest land occur within the State and constitute about half of the total land area.

At base Federal sampling intensity, one FIA annual inventory plot is established per approximately 2,400 ha. Each FIA ground plot comprises a cluster of four points, each surrounded by a 7.32-m (24-ft) fixed radius subplot within which subplot center points are assigned a land use class, land use conditions are mapped, and trees are measured (Bechtold and Scott 2005). FIA's four subplots consist of one central subplot and three peripheral subplots. Peripheral subplot centers are spaced 36.58 m (120 ft) from the center of the central subplot, at azimuths of 0, 120, and 240 degrees from the central subplot. Thus, the outer three subplots of FIA cluster plots form an equilateral triangle with a circumcircle through subplot centers of 73.152 m (240 ft) in diameter.

During Michigan's first annual inventory, FIA plots were sampled at an intensity three times greater than the base Federal sample intensity, resulting in 18,952 measured plots. Of these plots, 18,233 were fully sampled and had land use conditions recorded for all four subplots. Forty-five percent of these plots (8,211) had forest condition at all four subplot centers and 44 percent (8,015) had nonforest condition at all four subplot centers. Plots containing a mixture of forest and nonforest conditions at subplot centers constituted 11 percent (2,007) of the total number of plots, ranging from 3 to 22 percent across Michigan counties.

Landscape metrics were estimated following the approach defined by Kleinn (2000).

In short, Kleinn's (2000) approach uses an inventory cluster plot's shape and size to determine the probability of intersecting an imaginary forest/nonforest buffer, from which one can estimate mean patch size and the number of patches from observations of subplot center conditions. In this study, we knew FIA's cluster plot shape and circumcircle diameter, but we did not know typical patch shape; therefore, we calculated metrics using various assumptions of patch shape.

The triangular shape of FIA's three peripheral subplots determines a conditional probability of plot intersection with a forest/nonforest border, $P_{is} = \frac{3\sqrt{3}}{2\pi} = 0.82699$. The proportion of area in an estimation unit within an imaginary forest/nonforest buffer is $\hat{p}_s / P_{is}$, where $\hat{p}_s$ is the proportion of plots having a mixture of forest and nonforest subplot center conditions. A relationship between forest area and buffer area is defined as $\hat{k} = \frac{\hat{F}_{Forest}}{\hat{F}_{Buffer}} = \frac{\hat{p}}{\hat{p}_s} P_{is}$, where $\hat{p}$ is the estimated forest cover proportion. An estimate of mean forest patch size follows, $\hat{a} = \frac{1}{v}\hat{k}^2 d^2$, given the constant v, which relates to typical patch shape, and the FIA cluster plot circumcircle diameter d = 73.152 m , which was converted to d = 0073152 km to maintain the consistency of units of patch area estimates. We applied multiple constants of v, assuming shapes of forest patches to be circular, square, or one of several rectangular shapes with various length/width ratios. A standardized measure of relative patch size, independent of patch shape, is defined as $\hat{k}^2_{s\tan d} = \hat{k}^2 d^2$. The mean number of forest patches was estimated as $\hat{b} = \frac{\hat{p}F}{\hat{a}}$, where F is the known total area of an estimation unit. See Kleinn (2000) for additional explanation and for metrics of forest/nonforest perimeter length, buffer area, and estimates of precision.

Satellite Image-Based Land Cover Data

We used the National Land Cover Dataset (NLCD) of 1992 as our satellite image-based source of forest landscape metrics (fig. 1). The NLCD of 1992 has the following characteristics: thematic land cover data set of the conterminous United States, 30-m spatial resolution, derived from Landsat Thematic Mapper satellite imagery circa 1992, comprising 21 land cover classes, and produced by the U.S. Geological Survey (Vogelmann *et al.* 2001). Prior to calculating patch metrics, we updated the NLCD to capture recent development that had occurred since 1992 and we conducted additional processing to more closely conform with FIA's definitions of forest patch minimum area and width. To accomplish these preprocessing steps, we performed the following five sets of activities: (1) converted a vector data set of road networks to a 30-m raster data set; calculated road density using a convolution filter (moving window) with a circular, 7-pixel radius circle; and reassigned NLCD pixels within areas of high road density into one of five new "developed" classes, following the procedure of Lister *et al.* (2005); (2) aggregated the NLCD classes and five new developed classes into either forest or nonforest class; (3) clumped and eliminated isolated forest and nonforest pixel clusters such that the resulting data set contained no pixel clusters with fewer than four 30-m pixels; (4) bisected forest pixel clusters with Topologically Integrated Geographic Encoding and Referencing System (TIGER) roads data such that resulting forest patches would be bounded and constrained in size by roads; and (5) masked nonforest pixels to exclude them from subsequent analyses. Hereafter, this enhanced NLCD 1992 data set is referred to as NLCD+.

Figure 1.—*Landsat Thematic Mapper-based 30-m forest cover data set, Michigan, USA.*



Estimates of mean patch size and the number of forest patches were produced as follows. Unique patches of forest were identified using the REGIONGROUP command in the ArcInfo GRID software package. The command works by grouping adjacent pixels of the same class into discrete regions, assigning a unique number to each region, assigning a region's unique number to every pixel within that region, and producing a summary table that lists all regions and the count of pixels within each region. The number of forest patches is equivalent to the number of regions. The size of each patch was calculated as the product of the number of pixels per region multiplied by the size of each 30-m (0.09-ha) pixel. Per-county estimates of the number and mean size of forest patches were obtained by assigning each region to one county, based on the location of the geographic centroid of each region, and summarizing the regions within each county.

Comparison of Plot- and Image-Based estimates

Plot- and image-based estimates of per-county forest area, mean patch size, and the number of patches were compared using linear regression analyses. Estimates of the relative number of patches and relative patch size were computed as the ratio of county estimates to the statewide estimate for both patch size and the number of patches. Regressions of relative patch size were produced before and after applying a lognormal transformation to the per-county estimates.

Results

Michigan forest area estimates from FIA and NLCD+ were 78,155 and 79,392 km^2, respectively. The FIA estimate of Michigan forest land was about 52 percent of total land area, which includes noncensus water and inland census water, ranging from 9 to 88 percent among counties. The comparison of per-county forest area estimates from FIA and NLCD+ is described by the regression equation y = 0.9456x + 66.129, with R^2 = 0.9909. About 11 percent of Michigan FIA plots have a mixture of forest and nonforest among the three peripheral subplot center conditions, ranging across counties from 3 to 22 percent. The FIA estimate of Michigan area within a forest/nonforest buffer of 73.152 m (the diameter of the FIA plot circumcircle) in width was 13 percent, ranging across counties from 4 to 26 percent. FIA and NLCD+ estimates of mean size and the number of Michigan forest patches and corresponding county minimums and maximums are reported in table 1. When patch shape was assumed to be circular, FIA and NLCD+ estimates of patch size were 0.91 and 0.29 km^2, respectively, and estimates of the number of patches were 86,095 and 276,143, respectively. FIA estimates grew substantially larger for patch size and substantially smaller for the number of patches as patch shape parameters diverged from an assumption of circular patch shape (table 1). An estimate of $\hat{k}^2_{s\tan d}$, the standardized metric related to mean patch size, was 0.0722 for Michigan, ranging from 0.0042 to 0.7254 among Michigan counties when circumcircle diameter was measured in kilometers.

FIA estimates of per-county relative forest patch size (fig. 2) were moderately positively correlated with NLCD+ estimates before applying a lognormal transformation, y = 0.6783x + 0.4761, R^2 = 0.483, and more strongly positively correlated after applying a lognormal transformation, y = 0.7176x + 0.1207, R^2 = 0.8457.

Figure 2.—*Ratio of Forest Inventory and Analysis plot-based per-county mean forest patch size to statewide mean forest patch size, Michigan, USA.*

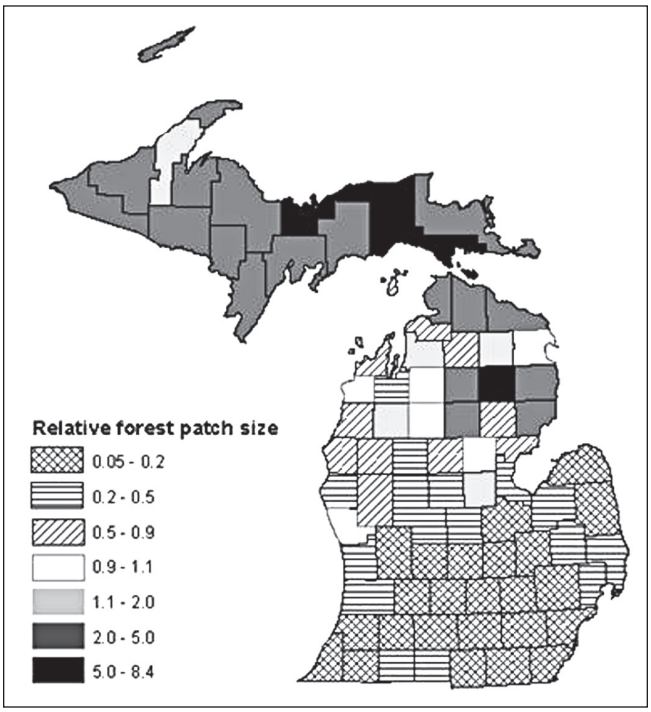


Table 1.—*Mean size and number of forest patches estimated from three peripheral subplot centers of FIA annual inventory cluster plots (per Kleinn 2000) and from an enhanced NLCD data set, MI, USA.*

Data source	Patch shape	Length/width factor	Patch size (km2)		Number of patches	
			MI	(County range)	MI	(County range)
FIA	circular	NA	1.06	(0.05 – 9.11)	73,719	(216–4303)
FIA	square	1	1.35	(0.07 – 11.61)	57,897	(169–3379)
FIA	rectangular	2	1.52	(0.08 – 13.06)	51,464	(150–3004)
FIA	rectangular	4	2.11	(0.11 – 18.13)	37,054	(108–2163)
FIA	rectangular	8	3.42	(0.17 – 29.38)	22,873	(67–1335)
FIA	rectangular	10	4.08	(0.20 – 35.11)	19,140	(56–1117)
FIA	rectangular	20	7.44	(0.37 – 63.98)	10,503	(31–613)
FIA	rectangular	50	17.56	(0.88 – 150.93)	4,452	(13–260)
NLCD	pixel region	NA	0.29	(0.04 – 2.98)	276,143	(429–7958)

FIA = Forest Inventory and Analysis. MI = Michigan. NA = Not applicable. NLCD = National Land Cover Dataset.

Discussion

In this study, we produced estimates of Michigan county mean forest patch area and the number of patches from FIA annual inventory subplot center condition data. Compared with NLCD+ estimates summarized by county, FIA estimates were comparable in total forest area, fewer in number of forest patches, larger in size of forest patches, and similar in relative size of forest patches. These results are similar to those of Riemann *et al.* (2003), who reported that NLCD-based estimates of Massachusetts forest patch size were smaller than estimates derived from aerial photo interpretation.

Although FIA estimates were closest to NLCD+ estimates when patch shape was assumed to be circular, these estimate pairs for both patch size and the number of patches still differed from each other by a factor of three. Explanations for this discrepancy are yet unknown but following are possible reasons. First, it was observed that the smallest NLCD+ forest patch size was 4 pixels for all Michigan counties, equivalent to 0.36 ha or 0.0036 km^2, which is the smallest region of pixels constrained by our preprocessing steps. These 4-pixel patches are slightly smaller than FIA's 0.4-ha (1-acre) definition of minimum forest area, resulting in a possible source of bias. Furthermore, except for 2x2-pixel configurations, 4-pixel clusters are narrower (30 m) than FIA's minimum width requirement of 36.576 m (120 ft). Forest patches 1 pixel wide and more than 4 pixels in length exceed FIA's minimum size requirement, but not the minimum width requirement. Within larger patches of forest, unimproved roads and nonforest strips narrower than 36.576 m or smaller than 1 ha are considered to be forest land, according to FIA definitions. Improved roads of any size are considered to be nonforest. We assumed all roads in the TIGER data set to be improved roads, thus nonforest. Further investigation is required to determine the distribution of these small pixel clusters and their effect on the discrepancy with FIA estimates of forest patch size and the number of patches.

Both Kleinn (2000) and Van Deusen (2005) require prior knowledge of patch shape, a parameter that is not derived from the inventory plot data themselves. Estimates of mean patch size and the number of patches vary greatly with patch shape. Sources of patch shape could include other field data (e.g., forest stand maps), analysis of remotely sensed data, or literature. A metric of relative patch size that is independent of patch shape was described by Kleinn (2000); such metrics allow for comparison between estimation units and inventories, independent of cluster plot configuration or patch shape, although they do not provide the specific information needed for assessing minimum area requirements of wildlife habitat.

FIA cluster plots contain only four subplots and results reported here are for estimates produced when using only the three peripheral subplots. Although estimates of precision are not reported here, we speculate that lower precision of patch size estimates may occur when using Kleinn's (2000) approach with FIA cluster plots than could be achieved with Van Deusen's (2005) mapped plot or with alternate cluster plot designs having more subplots. The comparison of estimates and the precision of those estimates from multiple approaches are recommended. Additional investigations are needed to develop operational protocols for integrating metrics of landscape and FIA attributes, which would provide for additional data, information, knowledge of our Nation's forests and their spatial pattern, and suitability for wildlife habitat and other ecological functions. Additionally, these metrics could be integrated with non-FIA inventory data obtained from nonmapped cluster plots of various designs.

Acknowledgments

We thank Rachel Riemann for providing road density data and enhanced National Land Cover Dataset land cover classes in developed areas. We thank Dr. Christoph Kleinn for previously developing a methodology using unmapped cluster plots and for his recent guidance and encouragement.

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.

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. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station: 27-42.

Kleinn, C. 2000. Estimating metrics of forest spatial pattern from large area forest inventory cluster samples. Forest Science. 46: 548-557.

Lister, A.; Riemann, R.; Lister, T.; Hoppus, M. 2003. Techniques and considerations for FIA forest fragmentation analysis. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C.; Moser, J.W., eds. Proceedings, third annual forest inventory and analysis symposium. Gen. Tech. Rep. NC-230. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station: 156-161.

Lister, A.; Riemann, R.; Lister, T.; McWilliams, W. 2005. Northeastern regional forest fragmentation assessment: rationale, methods, and comparisons with other studies. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C.; McWilliams, W.H., eds. Proceedings, fifth annual forest inventory and analysis symposium. Gen. Tech. Rep. WO-69. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station: 13-17.

McGarigal, K.; Marks, B.J. 1995. FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. Gen. Tech. Rep. PNW-GTR-351. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 122 p.

Riemann, R.; Lister, A.; Hoppus, M.; Lister, T. 2003. Fragmentation statistics for FIA: designing an approach. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C.; Moser, J.W., eds. Proceedings, third annual forest inventory and analysis symposium. Gen. Tech. Rep. NC-230. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station: 146-155.

Van Deusen, P.C. 2005. Mapped plot patch size estimates. In: McWilliams, W.H., ed. Proceedings, fifth annual forest inventory and analysis symposium. Gen. Tech. Rep. WO-69. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station: 111-116.

Vogelmann, J.E.; Howard, S.M.; Yang, L.; Larson, C.R.; Wylie, B.K.; Vandriel, N. 2001. Completion of the 1990s National Land Cover Data Set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. Photogrammetric Engineering & Remote Sensing. 67: 650-662.

Wenny, D.G.; Clawson, R.L.; Faaborg, J.; Sheriff, S.L. 1993. Population density, habitat selection and minimum area requirements of three forest-interior warblers in Central Missouri. The Condor. 95: 968-979.

Using Forest Service Multiple Species Inventory and Monitoring Protocols To Count Birds at Forest Inventory and Analysis Plots on the Caribbean Landscape: Results, Observations, and Challenges From Year 1 of a 2-Year Study

Sonja N. Oswalt[1], Thomas J. Brandeis[2], David W. Steadman[3], and Scott K. Robinson[4]

Abstract.—We conducted double-observer point counts of birds from December 3 to December 31, 2005, on preestablished permanent Forest Inventory and Analysis (FIA) plots and National Park Service System trails within the Virgin Islands National Park, St. John, U.S. Virgin Islands. We had three objectives: (1) to collect abundance and distribution data for wintering land birds, particularly neotropical migrants, in the subtropical dry and moist forests of St. John; (2) to test for differences between data collected on the FIA systematic grid and data collected using point counts established along randomly selected National Park System trails and; (3) to evaluate the effectiveness of Multiple Species Inventory and Monitoring (MSIM) protocols in the tropics. We recorded all species of birds (resident and migrant) observed or heard, and the distance from observers to each individual. Broad habitat descriptions were made for trails where habitat data were not already available. In year 1, we identified 47 indigenous species of birds, including 17 species of neotropical migrants. The most frequently recorded species was the resident songbird Caribbean Elaenia, with 517 registrations. The most frequently recorded migrant songbird was the American redstart (98 registrations). We present an overview of the Forest Service, U.S. Department of Agriculture FIA and MSIM protocols as well as results, observations, and challenges from year-1 sampling efforts on St. John.

Introduction

Monitoring plant and animal species at landscape levels can be cost and time intensive to public agencies and, ultimately, taxpayers. Although various intra- and interagency missions differ in the Federal government, there exists a degree of overlap among many, particularly in the natural resources community, where baseline assessments and continued monitoring are necessary and expected goals. The Forest Service Forest Inventory and Analysis (FIA) program is a nationally consistent inventory and monitoring program designed to assess the extent and condition of public and private forest lands in the United States, including outlying islands and U.S. territories. The FIA program collects a wealth of environmental data on an unbiased, systematic sample grid. Much of the data collected by FIA for assessing forest health and condition are very useful to wildlife biologists for assessing habitat extent and condition.

National Forest System and other public lands (e.g., National Park Service, U.S. Fish & Wildlife Service) stand to benefit from FIA monitoring when efforts to sample terrestrial vertebrate and invertebrate species are structured to coincide with FIA plots so that cooperative sampling can occur within or between agency staff. Resulting research may then benefit all

[1] Forester, U.S. Department of Agriculture (USDA) Forest Service, Southern Research Station, Forest Inventory and Analysis, 4700 Old Kingston Pike, Knoxville, TN 37919. E-mail: soswalt@fs.fed.us.

[2] Research Forester, USDA Forest Service, Southern Research Station, Forest Inventory and Analysis, 4700 Old Kingston Pike, Knoxville, TN 37919. E-mail: tbrandeis@fs.fed.us.

[3] Curator of Ornithology, University of Florida Museum of Natural History, Dickinson Hall, Gainesville, FL 32611. E-mail: dws@flmnh.ufl.edu.

[4] Katharine Ordway Professor of Ecosystem Conservation, University of Florida Museum of Natural History, Dickinson Hall, Gainesville, FL 32611. E-mail: srobinson@flmnh.ufl.edu.

involved agencies and thus help to lower the overall cost to any one entity, thereby lowering overall costs to taxpayers. In addition, colocating wildlife sampling with a nationally developed forest inventory permits standardization over large geographic extents, allowing for cross-regional modeling and comparisons.

Building on that concept, the Forest Service developed a Multiple Species Inventory and Monitoring (MSIM) protocol designed to yield a "consistent and efficient method for obtaining basic presence/absence data and associated habitat condition data for a large number of individual species at sites that represent a probabilistic sample" (Manley *et al.* 2006). The MSIM protocol layers common presence/absence species survey methodologies over the established FIA systematic grid to colocate sample points, thus taking advantage of data already collected by FIA. Protocols from both survey programs allow for intensification, both increasing the statistical accuracy of FIA data and increasing the usefulness of MSIM data where funding is available.

Methodologies contained in the MSIM program were developed and compiled by a team of biologists across the continental United States using well-known and tested protocols. The MSIM program, in conjunction with FIA, has received limited implementation, however, resulting in a paucity of information about the ability to perform the suggested protocols across a variety of landforms and habitat types. To date, MSIM, in conjunction with FIA sampling, has been tested and used primarily on national forest land in the Pacific Southwest, which is not representative of all United States landscapes, particularly insular landscapes occurring at subtropical and tropical latitudes.

An FIA sample plot framework was established on St. John, U.S. Virgin Islands, in the summer of 2004 to record forest resources and health. The Caribbean West Indies, including the island of St. John, are home to not only resident land birds year-round but also a host of neotropical migratory songbirds each winter. This last group of species requires an intercontinental effort for conservation goals to succeed. The Caribbean islands are the "most important (and sometimes exclusive) wintering ground" for many declining species of warblers, including the Cape May warbler, black-throated blue warbler, and prairie warbler (BirdLife International 2004). These species face continuing pressures on their wintering grounds. Habitat destruction, fragmentation, and invasive species are primary threats to avian diversity of the Caribbean. Arendt (1992) stressed the need for more long-term monitoring of birds on islands throughout the West Indies, particularly areas in which forested habitat persists.

Previous avian studies on St. John (Askins and Ewert 1999, Askins *et al.* 1992) have been conducted primarily along established trail systems and roadways, which may result in forest-edge, forest-type, or landform-related biases. Given the need for up-to-date information regarding migratory and resident land birds on St. John, combined with the need to test MSIM protocols on FIA plots in tropical regions, we initiated a 2-year avian study designed to accomplish three goals: (1) collect abundance and distribution data for wintering land birds, particularly neotropical migrants, in subtropical dry and moist forests; (2) test for differences between data collected on the FIA systematic grid and data collected using point counts established along randomly selected National Park System trails and; (3) evaluate the effectiveness of MSIM protocols in the tropics. Here, we discuss the results and challenges from the first year of avian sampling.

Methodology

Vegetation Sampling

Habitat data were collected on the island of St. John, U.S. Virgin Islands, in June through July of 2004 using the Forest Service FIA sample design (USDA Forest Service 2002). Twenty plots were arranged on an unbiased, systematic hexagonal sample grid across the island. Hexagons covered about 200 ha each, with one sample plot located within each hexagon. Plots were located and mapped using Global Positioning System equipment to ensure a high degree of accuracy. Plots were installed and measured in locations where at least 10-percent tree canopy coverage and a minimum forested area of 0.4 ha around each plot center was present, as designated by Forest Service FIA sample design guidelines. Plots consisted of a cluster of four subplots, each with a 7.3-m radius. Each subplot was 167 m^2, for a total sampled area of 670 m^2 (0.017 ha)

in a fully forested plot. Plots were assigned to life zones sensu Holdridge (1967).

Field crews collected forest inventory, understory structure and composition, and physiographic data on each subplot. Diameter at breast height (d.b.h.) (taken at 1.37 m), total height, crown ratio, crown width, and other parameters were measured on all trees with d.b.h. greater than or equal to 12.5 cm within the subplots; d.b.h. and total height were measured on saplings with d.b.h. greater than or equal to 2.5 cm within a 2.1-m radius microplot nested in each subplot (Bechtold and Scott 2005, USDA Forest Service 2002).

Avian Sampling

A team of six qualified observers from the University of Florida conducted double-observer, unlimited-radius circular plot point counts from December 3 to December 31, 2005. A total of 64 point-count stations were associated with FIA plots. At each FIA plot, two to five stations were arrayed around a central location slightly offset from the FIA plot center, as recommended by MSIM protocols. Each point-count station was located at least 200 m apart, and all points associated with one FIA plot were sampled on the same day using 10-minute point counts made during times of little or no rain, little or no wind, and within 4 hours of daybreak. Ninety-three additional points were established at 200-m intervals along randomly selected, maintained National Park Service trails systems to augment the FIA sample, for a total of 157 point-count stations. All species of birds detected during point counts were recorded by detection method (sight or sound) and by distance. Birds observed flying over count stations were recorded separately.

Statistical Methods

Absolute abundances of species at each point-count station were used to populate a species-by-station matrix (Wunderle and Waide 1993). The abundance matrix was then subject to Indicator Species Analysis in the statistical software PC-ORD, using two groups: (1) FIA (point-count stations colocated with FIA plots), and (2) non-FIA (point-count stations independent of FIA plots) (McCune and Mefford 1999). Indicator species were selected based on species importance values in the predefined groups, determined as a combination of relative abundance and relative frequency of species. Monte Carlo statistics were used to test the significance of species as indicators of a given group. Species with Monte Carlo significance at $P \leq 0.05$ in a given group were considered "more likely" to be encountered in that particular group than in the opposing group, even if the species was present to some degree in both groups.

Analysis of variance was used to evaluate if raw species counts differed between FIA and non-FIA groups. Multiresponse permutation procedures with Euclidean distance measures were used to test for differences in species composition between groups. A chi square was used to evaluate whether bias existed in the representation of life zones on FIA compared with non-FIA groups.

Results

Observers detected 2,317 individuals of 47 species on St. John (including flyovers), or an average of 14.8 individual birds per station. Seventeen of the species detected were neotropical migrants (table 1). The resident songbird Caribbean Elaenia *(Elaenia martinica)* was the most frequently recorded species, with 519 registrations. Detected at 85 percent of point-count stations, an average of 3.8 Caribbean Elaenias were observed at each station where the species was recorded. Although the Caribbean Elaenia was common in both life zones, it was observed more frequently in the subtropical dry forest ($P = 0.007$).

The most frequently observed neotropical migrant was the American redstart (*Setophaga ruticilla*; table 1). American redstarts were recorded a total of 98 times on 34 percent of point-count stations, with an average of 1.8 individuals observed at stations where the species was detected. American redstarts were most frequently observed in subtropical moist forests ($P = 0.001$).

The total number of species detected did not differ between FIA and non-FIA groups ($P = 0.26$). Point-count stations in the FIA group averaged 13.9 ($\pm$ 0.65) individuals per point whereas stations in the non-FIA group averaged 15.1 ($\pm$ 0.81) individuals per point. Species composition differed between groups ($T = -10$; $P < 0.0001$) but within-group heterogeneity was very high ($A = 0.04$), a phenomenon that was not surprising given that

Table 1.—*Number of year-1 observations of avian species at point-count stations, December 2005, St. John, U.S. Virgin Islands.*

Species	Number of observations	Species	Number of observations
Neotropical migrants		Bridled quail dove (*Geotrygon mystacea*)	47
American redstart (*Setophaga ruticilla*)	98	Green-throated carib (*Eulampis holosericeus*)	31
Yellow warbler (*Dendroica petechia*)	61	Zenaida dove (*Zenaida aurita*)	27
Northern parula (*Parula americana*)	51	Mangrove cuckoo (*Coccyzus minor*)	17
Black-and-white warbler (*Mniotilta varia*)	48	Smooth-billed ani (*Crotophaga ani*)	8
Ovenbird (*Seiurus aurocapillus*)	47	Common ground dove (*Columbina passerina*)	7
Prairie warbler (*Dendroica discolor*)	44	Scaly naped pigeon (*Columba squamosa*)	5
Northern waterthrush (*Seiurus noveboracensis*)	35	American kestrel (*Falco sparverius*)	3
Hooded warbler (*Wilsonia citrina*)	10	Red-tailed hawk (*Buteo jamaicensis*)	3
Worm-eating warbler (*Helmitheros vermivorus*)	10	Black-whiskered vireo (*Vireo altiloquus*)	2
Unknown warbler	5	Northern mockingbird (*Mimus polyglottos*)	2
Yellow-rumped warbler (*Dendroica coronata*)	4	Hummingbird sp.	2
Barn swallow (*Hirundo rustica*)	4	Puerto Rican flycatcher (*Myiarchus antillarum*)	1
Blackpoll warbler (*Dendroica striata*)	3	Puerto Rican flycatcher (Myiarchus antillarum)	1
Magnolia warbler (*Dendroica magnolia*)	2		
Black-throated blue warbler (*Dendroica caerulescens*)	1	**Others—data not analyzed**	
Black-throated green warbler (*Dendroica virens*)	1	Black-necked stilt (*Himantopus mexicanus*)	9
Cape May warbler (*Dendroica tigrina*)	1	Common moorhen (*Gallinula chloropus*)	6
Rose-breasted grosbeak (*Pheucticus ludovicianus*)	1	Unknown bird	4
Rose-breasted grosbeak (Pheucticus ludovicianus)	1	Belted kingfisher (*Ceryle alcyon*)	4
		Brown pelican (*Pelecanus occidentalis*)	3
Resident songbirds		Spotted sandpiper (*Actitis macularia*)	2
Caribbean Elaenia (*Elaenia martinica*)	519	Domestic chicken (*Gallus gallus*)	1
Pearly eyed thrasher (*Margarops fuscatus*)	403	Falcon sp.	1
Bananaquit (*Coereba flaveola*)	368	American coot (*Fulica americana*)	1
Lesser Antillean bullfinch (*Loxigilla noctis*)	151	Great blue heron (*Ardea herodias*)	1
Gray kingbird (*Tyrannus dominicensis*)	135	Great egret (*Casmerodius albus*)	1
Antillean crested hummingbird (*Orthorhyncus cristatus*)	74	Green heron (*Butorides striatus*)	1
Black-faced grassquit (*Tiaris bicolor*)	49	Magnificent frigatebird (*Fregata magnificens*)	1
		Sora (*Porzana carolina*)	1

groups were predetermined based on a nonbiological criterion rather than by environmental factors. Significant indicator species on FIA plots included the American kestrel, bananaquit, gray kingbird, lesser Antillean bullfinch, and yellow warbler. Significant indicator species on non-FIA plots included the American redstart, bridled quail dove, northern parula, and pearly eyed thrasher. Eleven species (20 individuals) were detected only at FIA related stations, whereas 9 species (18 individuals) were detected only at non-FIA related stations.

The distribution of FIA compared with non-FIA point-count stations differed by Holdridge life zone ($X^2 = 13.6$; $P < 0.001$). Sixty-two percent of FIA associated stations fell in the dry forest life zone, while 68 percent of non-FIA associated stations fell into the subtropical moist forest life zone. When FIA and non-FIA stations were considered together, distribution was relatively even between life zones, with 56 percent of stations in subtropical moist forest and 44 percent in dry forest.

Discussion and Conclusions

The results of this study suggest that neither the FIA sample grid nor the non-FIA–related trail systems adequately sample both life zones by themselves. Wintering neotropical migrants recorded during 1 year of study on St. John appear to prefer the subtropical moist forest life zone, while resident species generally appear to be more abundant in subtropical dry forests. Therefore, if only one method is used to sample songbirds (FIA or established park trails), the sample may need to be augmented to include additional points in underrepresented forest types.

Although methodologies differ slightly, our first year of sampling effort performed well when compared to other bird surveys on St. John and elsewhere in the Caribbean. Wunderle and Waide (1993) counted 28 species (12 migrants) at 60 point-count locations (generally along trails or roads) on St. John in 1993. Our study included 97 more stations, and we detected nearly twice the number of total species and nearly 1.5 times the number of migrant species.

The 20 FIA plots appear to undersample the subtropical moist forests, which, on St. John, occur on mountain tops at higher elevations and along gullies, ravines, and streams. Although the FIA grid is intensified on St. John, the sampling scale is still such that a systematic grid is less likely to capture rare or linear events on the landscape than would a stratified sample. Non-FIA sample points, in contrast, tend to be concentrated in this forest type because the same landscape features that can cause undersampling may influence the design of the park's trail systems and roadways. For example, ephemeral streams and moist forests are aesthetically pleasing and provide protective shade to adventurous tourists. Streams and trails follow similar "pathways of least resistance" up and down mountainous terrain. Also, the northern aspects where moist forests are prevalent provide more photoready viewing opportunities of sandy white beaches than are found on the drier, south sides of the island's mountains, which may affect road placement. In contrast, the FIA grid system results in a sample less affected by potential edge-effect biases associated with roads and trails. In addition, detailed quantitative vegetation information is available for FIA associated stations, whereas non-FIA station vegetation descriptions are vague and qualitative (although that could be remedied given additional time and money and using protocols similar to FIA for measuring vegetation information).

Although FIA-related point counts of birds have less bias and have the additional advantage of detailed habitat information, they also presented challenges to which non-FIA trail stations were not subjected. Terrain on St. John is steep and, in some areas, treacherous. A year-round growing season results in thick vegetation, including cacti and *Agave* species that make cross-country travel difficult. When conducting point counts of birds, researchers are generally restricted to a sampling period from dawn to 10 a.m., when birds are most active. On St. John, merely hiking to a given plot may take as much as 3 to 4 hours. In addition, thick vegetation and rapidly changing forest types restrict the number of point-count stations that may be associated with any given FIA point. In contrast to the Pacific Southwest, where MSIM was initially developed, our team was only able to establish, on average, two to three stations per FIA plot, as opposed to the suggested five to seven stations. Often, our team was only able to implement one station in the dawn-to-10 a.m. sampling window due to the time required to access the site. These complications suggest that sampling avian species on FIA plots, in difficult to access, rapidly changing terrain and vegetation types, may require modifications to the existing MSIM protocol. Nevertheless, the amount of information acquired by including FIA related stations serves as adequate justification for considering MSIM on FIA points in addition to traditional trail counts when developing avian monitoring plans.

Acknowledgments

The authors thank the U.S. Fish & Wildlife Service, Forest Service Southern Research Station, Forest Service International Institute for Tropical Forestry, the University of Florida, and Florida Museum of Natural History for project funding. They also thank the Virgin Islands National Park and personnel for their professional assistance. Special thanks are due Humfredo Marcano, Luis Ortíz-Vega, Iván Vicéns, Britta Dimick, Andy Kratter, Matt Reetz, Wendy Schelsky, Jill Jankowski, Gustavo Londono, and Natalie Wright for their expertise in the field. The authors thank Eileen Helmer and Ariel Lugo for their assistance throughout the project.

Literature Cited

Arendt, W.J. 1992. Status of North American migrant landbirds in the Caribbean region: a summary. In: Hagan, J.M., III; Johnston, D.W., eds. Ecology and conservation of neotropical migrant landbirds. Washington, DC: Smithsonian Institution Press. 608 p.

Askins, R.A.; Ewert, D.N. 1999. Impact of Hurricane Hugo on bird populations on St. John, U.S. Virgin Islands. Biotropica. 23(4): 481-487.

Askins, R.A.; Ewert, D.N.; Norton, R.L. 1992. Wintering migrants in fragmented and continuous forests, Virgin Islands. In: Hagan, J.M., III; Johnston, D.W., eds. Ecology and conservation of neotropical migrant landbirds. Washington, DC: Smithsonian Institution Press: 197-205.

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. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station: 27-42.

BirdLife International. 2004. The insular Caribbean: a biodiversity conservation priority. http://www.birdlife.org/regional/caribbean/overview.html. (20 October 2004).

Holdridge, L.R. 1967. Life zone ecology. San Jose, Costa Rica: Tropical Science Center. 206 p.

Manley, P.N.; Van Horne, B.; Roth, J.K.; Zielinski, W.J. *et al.* 2006. Multiple species inventory and monitoring technical guide. Gen. Tech. Rep. WO–73. Washington, DC: U.S. Department of Agriculture, Forest Service. 204 p.

McCune, B.; Mefford, M.J. 1999. PC-ORD. Multivariate analysis of ecological data. Ver. 4.0. Glendon Beach, OR: MjM SoftwareDesign.

U.S. Department of Agriculture (USDA) Forest Service. 2002. Supplement C to SRS regional manual 1.62: field procedures for Puerto Rico and the Virgin Islands. Knoxville, TN: U.S. Department of Agriculture, Forest Service, Southern Research Station, Forest Inventory and Analysis. http://srsfia2.fs.fed.us/data_acquisition/manual.shtml. [Date accessed unknown].

Wunderle, J.M., Jr.; Waide, R.B. 1993. Distribution of overwintering nearctic migrants in the Bahamas and Greater Antilles. The Condor. 95: 904-933.

Predicting Bird Habitat Quality From a Geospatial Analysis of FIA Data

John M. Tirpak[1], D. Todd Jones-Farrand[2], Frank R. Thompson, III[3], Daniel J. Twedt[4], Mark D. Nelson[5], and William B. Uihlein, III[6]

Abstract.—The ability to assess the influence of site-scale forest structure on avian habitat suitability at an ecoregional scale remains a major methodological constraint to effective biological planning for forest land birds in North America. We evaluated the feasibility of using forest inventory and analysis (FIA) data to define vegetation structure within forest patches, which were delineated from independent geospatial data sets of ecological subsection, forest type, and landform class. We used Swainson's warbler **(***Limnothlypis swainsonii*, Audubon) as a model to demonstrate how to integrate FIA data with geospatial data sets to estimate and monitor habitat suitability for a priority bird species in the West Gulf Coastal Plain/Ouachita Mountains Bird Conservation Region.

Introduction

The goal of the North American Landbird Conservation Plan is to create landscapes capable of sustaining bird populations at prescribed levels (Rich *et al.* 2004). To achieve this goal, the plan identified a three-step process:

1. Develop rangewide population objectives for each bird species.
2. Allocate these objectives to specific regions (e.g., Bird Conservation Regions (BCRs).
3. Translate these population objectives to habitat objectives within each region.

The first two steps of this process have been completed for most of the land birds breeding in the United States and Canada (Panjabi *et al.* 2005) and it is at the third step where conservation planning efforts stand today for most species.

Translating population targets to habitat objectives requires the development of models that explicitly state the relationship between bird numbers and habitat conditions. Given the number of land bird species covered by the North American Landbird Conservation Plan (448 species), the diversity of habitats they occupy, and the range in quantity and quality of available information for these species, it is not surprising that many modeling approaches are being explored to establish these relationships. Statistical models (e.g., hierarchical models, neural networks, regression procedures) provide an objective assessment of patterns in data, but are generally most suitable when large representative data sets are available to parameterize them. Statistical models are data hungry because they require relatively complex functions to compensate for the biases inherent in counting wild bird populations (Morrison 1998). When sufficient data do not exist (e.g., rare, nocturnal, or hard-to-sample species), a Habitat Suitability Index (HSI) framework provides one of the few practical alternatives for modeling species-habitat relationships. HSI models use *a priori* information to identify variables that affect the quality of a habitat for a given species, and this information is also used to create functions that relate habitat suitability to these key habitat requirements (Schamberger *et al.* 1982). HSI models have the desirable properties of scalability (landscape-scale parameters such as percent forest in the landscape can be easily incorporated into these models), intuitiveness,

[1] Postdoctoral Researcher, University of Missouri–Columbia, Vicksburg, MS 39180. E-mail: john_tirpak@fws.gov. (Note: The author was affiliated with the University of Missouri but physically stationed in Vicksburg, MS, during the time of this research.)

[2] Postdoctoral Researcher, University of Missouri–Columbia, Columbia, MO 65211. E-mail: FarrandD@missouri.edu.

[3] Project Leader, U.S. Department of Agriculture (USDA) Forest Service, Northern Research Station, Columbia, MO 65211. E-mail: frthompson@fs.fed.us.

[4] Research Wildlife Biologist, U.S. Geological Survey, Patuxent Wildlife Research Center, Vicksburg, MS 39180. E-mail: Dan_Twedt@usgs.gov.

[5] Research Forester, USDA Forest Service, Northern Research Station, St. Paul, MN 55108. E-mail: MDNelson@fs.fed.us.

[6] Assistant Joint Venture Coordinator, U.S. Fish & Wildlife Service, Lower Mississippi Valley Joint Venture, Vicksburg, MS 39180. E-mail: Bill_Uihlein@fws.gov.

and portability across sites, which make them useful in many circumstances (Larson *et al.* 2003).

Whether using a statistical or an HSI approach, current ecoregional-scale land bird planning relies mainly on landscape-scale spatial data (e.g., Dettmers *et al*. 2002). This is more a matter of necessity than preference, because spatially explicit site-level information (e.g., basal area) does not exist at the scale of an entire BCR (tens of millions of hectares), whereas landscape-scale data are much more readily available (e.g., landscape metrics derived from satellite image-based land cover classification); therefore, planning is limited to species that respond most strongly to landscape-scale factors (e.g., forest patch size preferences of the wood thrush [*Hylocichla mustelina*, Gmelin]; Driscoll *et al*. 2005). Because most species select habitat at multiple scales, our limited ability to assess habitat conditions at finer scales inhibits analyses of the suitability of an ecoregion for multiscale sensitive species.

Forest inventory and analysis (FIA) offers a potential data source to address these limitations. FIA data are collected to estimate the volume, growth, and condition of forest resources within large geographic extents (e.g., counties, states, ecological units, watersheds, or BCRs). FIA data could also provide information on habitat-specific forest structure attributes across large areas, which can be used to assess the suitability of an area for various bird species. Our objectives were to assess the utility of FIA data (1) to characterize avifaunal habitat structure in a spatially explicit manner at an ecoregional scale, and (2) to use this information to assess the sustainability of priority bird species in these landscapes.

Methodology

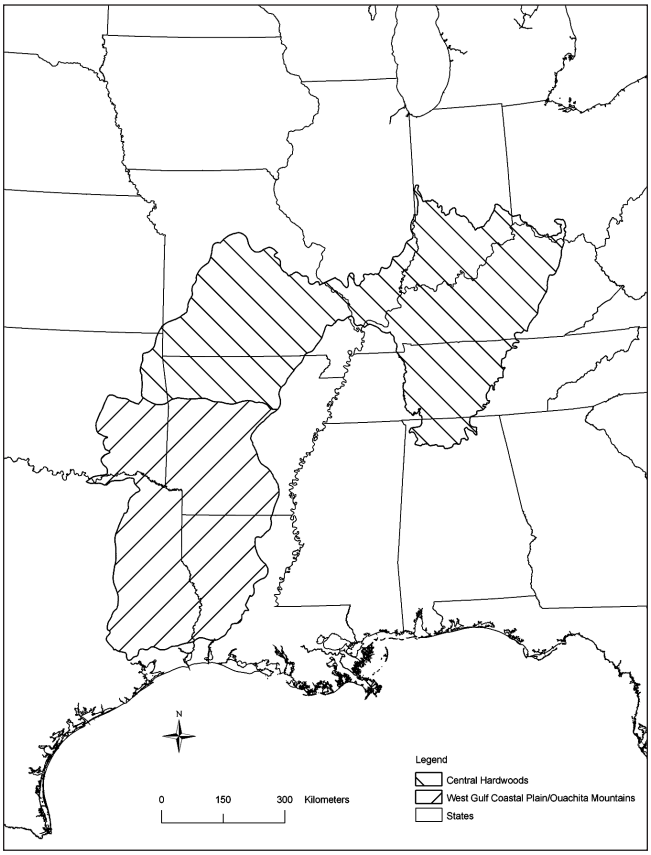
Avian Models

We developed HSI models for 40 priority bird species in the Central Hardwoods and the West Gulf Coastal Plain/Ouachita Mountains BCRs (fig. 1). We identified priority birds as species that utilize forested habitats with a total Partners in Flight regional combined score ≥ 20 (see Rich *et al*. 2004) or species designated as a Bird of Conservation Concern by the U.S. Fish and Wildlife Service in either BCR. To develop HSI models, we first performed a thorough literature review to identify site- and landscape-scale habitat factors that affected the occupancy, density, and/or productivity of each species. Empirical data derived from these sources formed the basis for individual suitability functions. We combined these suitability functions in biologically meaningful ways to produce overall habitat suitability estimates for density and productivity. Once initial models were developed, we solicited reviews from two to five experts for each species and revised models based on reviewer comments.

Geographic Information System Data

We constrained potential model variables to those available via nationally consistent geodata sets to maintain a uniform classification system across state boundaries within BCRs and to ensure our methodology was easily transferable to other forested biomes. We selected four nationally available geodata sets to define site and landscape conditions: ecological subsec-

Figure 1.—*Location of Central Hardwoods and West Gulf Coastal Plain/Ouachita Mountains Bird Conservation Regions, 2006.*



tions, the National Land Cover Dataset (NLCD), the National Elevation Dataset (NED), and the National Hydrography Dataset (NHD).

Our map of ecological subsections was based on the National Ecological Unit Hierarchy (Keys *et al.* 1995), which depicts relatively homogenous regions of topography, geology, climate, and potential natural communities. We therefore assumed subsection boundaries would capture a large amount of the variation in the broad-scale abiotic features that affect the composition and structure of the avian community within a BCR.

We used the NLCD 1992 to define the spatial location of forests and categorize forestlands into broad classes. NLCD 1992 delineates 21 land cover classes at 30-m resolution; 7 of these classes represent wooded land cover types that we used to define specific avian cover types: transitional, deciduous, evergreen, mixed, shrubland, orchard/vineyard, and woody wetlands (Vogelmann *et al*. 2001). Additionally, we included low-density residential as a forested land cover to capture the suburban shade tree habitats that are used by some priority species (e.g., orchard oriole [*Icterus spurious*, Linnaeus]).

Landforms (e.g., ridges, valleys) are local topographic features that can have a profound effect on both the flora and fauna of a forest community. Because no nationally consistent data set exists for this feature, we created our own classification from the nationally available NED, which maps elevation in meters at 30-m resolution (Gesch *et al*. 2002). We generated a landform geodata set from five NED-derived variables: relief, slope, aspect, local topographic position index (TPI), and landscape TPI. We separated areas of high and low relief by examining the standard deviation (SD) of elevation values within a 500-m radius moving window. We considered areas with an SD < 2 to be low relief and areas with an SD ≥ 2 to be high relief. We used a 5-percent threshold to separate high slope and low slope locations. We defined high-exposure (drier) slopes as those with aspects between 157.5 and 292.5 degrees (i.e., south by southeast to west by northwest) and all other aspects as low-exposure (moister) slopes. We placed areas lacking an aspect into a third category (flat). Derivation of TPI was based on a protocol developed by Jenness Enterprises (Jenness 2006), where the elevation at a pixel is compared to the mean elevation within a user-defined neighborhood. We calculated two separate TPI functions to highlight both local (500-m radius) and landscape (1500-m radius) effects and categorized the resulting spatial products into three classes: > 1 SD above the mean, > 1 SD below the mean, and within 1 SD of the mean. We defined 6 landform classes (floodplains, valleys, mesic slopes, terraces, xeric slopes, and ridges) based on the 108 unique combinations of values from the previously mentioned 5 variables.

Lastly, we used medium-resolution NHD (USGS 1999) to define the location of streams and other small water bodies that were not adequately captured by the NLCD but were important habitat cues for many priority species (e.g., Louisiana waterthrush [*Seiurus motacilla*, Vieillot]).

FIA Data

The geodata sets described previously allowed us to characterize landscape composition and structure, but we relied on FIA data to provide information about site-level forest structure. Staff from FIA's Spatial Data Services (SDS) center in St. Paul, MN, queried plot data to obtain unique plot numbers and location coordinates for 20,522 plots located within the two BCRs. These plots were sampled between 1986 and 1995, the years associated with the periodic inventories closest in date to the NLCD 1992 data set. Although true plot location coordinates were not made available to us, we were able to download publicly available PLOT, COND, and TREE tables for each state intersecting the BCRs (http://www.ncrs.fs.fed.us/4801/tools-data/data/). The three FIA tables for each state were imported into an Access (Microsoft, Redmond, WA) relational database, then combined and queried to generate tables containing plot-level summaries of the variables needed for our habitat models.

FIA does not measure all the key forest attributes for avian habitat selection on all phase 2 plots; therefore, we fit a regression equation to predict small (< 2.54 cm diameter at breast height [d.b.h] woody stem density (a derived phase 3 plot variable) from basal area and tree density (phase 2 plot variables). Similarly, we estimated overstory canopy cover from tree diameter and pole and sawtimber tree density based on an equation developed by Law *et al*. (1994). All other forest structure attributes were summarized directly from FIA data. We created

a summary table containing all forest structure variables and joined this table to the forest patch attribute table via the FIA plot identification number common to both tables.

The sampling intensity of periodic and annual FIA plots is not adequate for spatial interpolation (e.g., kriging) on forest structural attributes because the distance between plots is much greater than the distance over which these attributes are spatially correlated (Coulston *et al*. 2004). The spatial limitations of FIA's sampling design, coupled with privacy protections that restrict public access to exact plot locations, necessitated the development of an ecologically meaningful protocol for populating each forest patch in our landscape with FIA plot attributes. To accomplish this, we devised a stratification procedure that first defined our BCRs as patches of unique combinations of variables (i.e., strata), then identified the FIA plots within each of these unique combinations, and, finally, assigned each patch an FIA plot that had the same strata characteristics.

We stratified each BCR by ecological subsection, NLCD forest class, and landform type because we believed that these variables accounted for the greatest amount of variation in forest structure at the landscape scale. To avoid creating singular strata (i.e., unique combinations of variables associated with only one FIA plot) that would prohibit data accessibility, we used a reduced number of strata. Thus, we aggregated NLCD into six classes: deciduous, mixed, evergreen, woody wetland, transitional-shrubland, and nonforest. This stratification produced a map that contained 36 unique strata combinations (6 NLCD classes × 6 landform classes) in each ecological subsection.

SDS personnel spatially joined actual FIA plot locations to these strata and returned an attribute table containing plot identification numbers (but not coordinates) and the values for each of the three strata. Plot identification numbers allowed us to link each plot and its known strata values to our summary table of FIA and derived forest structure attributes. Due to potential security issues associated with some linkages, strata values for a small proportion of plots were not provided to us. Nonetheless, our stratification scheme allowed us to associate approximately 97 percent of FIA plots on private land and 60 percent of FIA plots on public land with our geospatial strata.

An inherent artifact of this approach is the wide range of FIA data plots associated with each strata; common strata combinations contained > 200 plots whereas rare combinations contained ≤ 1. To prevent all patches in a single strata combination from being represented by a relatively small number of plots, we established a six-plot minimum threshold for definition of all strata combinations and developed decision rules to guide aggregation of strata to achieve a minimum of 6 FIA plots. First, we identified all strata combinations within an ecological subsection that contained < 6 FIA plots and determined the proportion of the subsection represented by that unique landform-NLCD combination. If a stratum covered < 5 percent of the subsection, we considered it a rare strata and combined it with a similar NLCD class within the same landform (e.g., plots from floodplain woody wetlands would be aggregated with plots from floodplain deciduous). If a stratum covered > 5 percent of the subsection, we combined strata among similar landforms within the same NLCD classes (e.g., plots from floodplain woody wetlands would be aggregated with plots from valley woody wetlands). Through iterative applications of these rules, we combined strata across similar NLCD and landform classes to achieve the six-plot threshold. In some small and predominately nonforested subsections, we combined strata from different subsections to reach the six-plot threshold. In these cases, we combined subsections within the same ecological section before combining between different ecological sections. Once all strata were assigned at least six FIA plots, we assigned an FIA plot to every forested patch in our study area. We used a modified random number generator to assign an FIA plot identification number to each patch from the corresponding pool of plots associated with each unique combination of strata.

Spatial Assignment of FIA and Derived Forest Structure Variables

To spatially map FIA variables, we created individual geodata sets of each forest structural variable by reclassification on the variable of interest. This produced geodata sets wherein every pixel in a forest patch received the attribute value (e.g., basal area) measured on the plot assigned to that patch. Because all attributes of a plot are assigned together, the covariance structure of the FIA data was maintained and improbable combinations

of attributes (e.g., high sawtimber tree density and low basal area) were avoided. We caution that the final product of this procedure is a spatially explicit depiction of forest structure attributes; however, it is not spatially exact (i.e., each pixel has a value, but it is not necessarily the value that would be observed at that location). Despite this, because the final model outputs will be summarized by subsection and FIA data are representative of forest conditions within subsections, spatial exactness of these attributes within a subsection is not required.

Results/Discussion

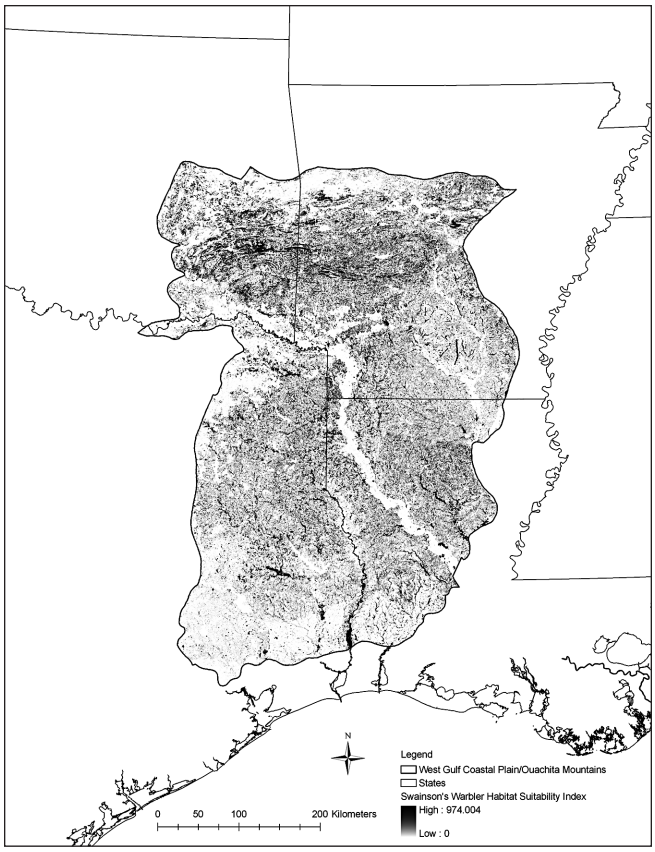
To illustrate our methodology, we present an example of a habitat suitability assessment for Swainson's warbler in the West Gulf Coastal Plain/Ouachita Mountains BCR. Swainson's warbler is a neotropical migrant that breeds in a variety of habitat types including canebrakes and palmetto thickets in mature bottomland hardwood stands in the Southeast, rhododendron thickets in the southern Appalachian Mountains, and 7- to 10-year-old pine plantations in eastern Texas (Brown and Dickson 1994). What these habitats have in common is a high density of small stems in the understory. Graves (2002) observed Swainson's warblers in habitats with a mean of 34,773 small stems per hectare and routinely found this warbler on sites with > 70,000 small stems per hectare. Regardless of location, Swainson's warblers are typically found in predominantly forested landscapes, and Eddleman *et al.* (1980) suggested contiguous forest tracts > 350 ha may be needed for the species to occur consistently.

Based on this information, we constructed an HSI model that contained six parameters related to bird density: landform, land cover, age class, forest patch size, percent forest in the local (1 km) landscape, and small stem density. The first suitability index (SI1) combined six landform (derived from NED), seven land cover (derived from NLCD), and five age (grass-forb, shrub-seedling, sapling, pole, and sawtimber; derived from FIA) classes into a single matrix that defined unique combinations of these variables. We directly assigned SI values to these combinations based on habitat suitability data from Hamel (1992). We also included forest patch size (SI2; derived from NLCD) in our model because of the presumed preference of Swainson's warblers for interior forest sites. We assumed forest patch sizes > 350 ha were adequate for Swainson's warblers and based a logistic function on data from Eddleman *et al.* (1980) describing this relationship. Nonetheless, forest patch size requirements are likely influenced by the percentage of forest in the local (1-km radius) landscape (SI3; derived from NLCD). Warblers in predominantly forested landscapes may use smaller forest patch sizes that may not be occupied in predominantly nonforested landscapes (Rosenberg *et al.* 2003). We considered landscapes with < 20 percent forest to be poor habitat and landscapes with > 80 percent forest to be excellent habitat (Donovan *et al*. 1997) regardless of forest patch size. We used the maximum of SI2 and SI3 where patch size and landscape composition influences were competing. Lastly, we included small stem density (SI4; derived from FIA) as a variable due to the affinity of this species for thick and dense understories in all occupied habitats. Site-level factors (landform, land cover, age class matrix, and small stem density) were weighed evenly in the overall calculation of habitat suitability by calculating the geometric mean of the individual SI values from the site-level factors. Furthermore, we weighed site- and landscape-scale factors evenly in the final suitability index score by calculating a geometric mean of site-level factors multiplied by the occupancy value from the maximum of the landscape factors (equation 1).

$$HSI = (((SI1 * SI4)^{0.500}) * Max (SI2 \text{ or } SI3))^{0.500}$$

By applying these SI functions to the appropriate data layers, we derived estimates of habitat suitability for Swainson's warblers throughout the West Gulf Coastal Plain/Ouachita Mountains (fig. 2). By subsequently relating the SI values to known abundances from point counts and other surveys, we estimated abundance of Swainson's warblers within individual subsections and, in aggregate, the entire BCR. This process has permitted a transparent mechanism to assess the comparability of "bottom-up" population estimates to the "top-down" target population numbers in the North American Landbird Conservation Plan. Additionally, we have used this same approach to assess the suitability of the West Gulf Coastal Plain/Ouachita Mountains BCR for productivity. By coupling density and productivity SIs, we are able to more accurately estimate the sustainability of Swainson's warbler populations in these habitats.

Figure 2.—*Distribution of habitat suitability index values for Swainson's warblers, West Gulf Coastal Plain/Ouachita Mountains Bird Conservation Regions, circa 1992.*



These results have important implications for ecoregional land bird planning, particularly in predominantly forested landscapes. In agricultural landscapes where forest cover is limited, such as the Mississippi Alluvial Valley, forest patch size and structure have an overriding influence on the suitability of a particular site for forest birds (Twedt and Loesch 1999). Conversely, in forested landscapes such as the Central Hardwoods or the Gulf Coastal Plains, forest blocks are relatively large, and the main determinant of habitat suitability for many species is the structure of the forest within the patch (Conner and Dickson 1997).

The HSI approach described here was applied by combining FIA periodic inventory data with independently available geospatial data sets, all of which are benefiting from recent enhancements. Horizontal accuracy of FIA plot location coordinates continues to improve with advances in field protocols and Global Positioning System technology. The NED data set, used for modeling landform, is being edited and enhanced, with spatial resolutions of 10 m in some locales. Ecoregion and NHD delineations continue to be refined and integrated at multiple scales. The imminent completion of an updated NLCD (2001) not only will provide temporal benefits, but also will deliver per-pixel estimates of percentage of forest canopy and percentage of impervious surface. New geospatial data sets are on the horizon too, including a nationwide 30-m data set of forest stand height derived from digital elevation model and Shuttle Radar Topography Mission data. While these revisions and new data sets are expected to provide further support for wildlife habitat assessments, they still do not provide the detailed forest structure attributes required for many HSI models. In the absence of wall-to-wall mapping of forest structure via remote sensing (e.g., Light Detection and Ranging and high spatial resolution optical sensors), integration of FIA with geospatial strata offers a viable solution to assessing forest structure attributes over large areas. With the advent of annual FIA surveys, the HSI approach offers a cost-effective habitat monitoring tool for a variety of forest species over broad areas.

Acknowledgments

This research was funded by a U.S. Geologic Survey/Fish & Wildlife Service Science Support Partnership grant. The authors thank M. Hatfield and R. McRoberts for invaluable guidance on technical issues related to FIA data acquisition and stratification; B. Dijak and H. He provided programming expertise for FIA data extraction and extrapolation; C. Baxter, J. Fitzgerald, A. Kiester, L. O'Brien, S. Shifley, H. Whiffen, and D. Yocum provided additional logistical and technical support.

Literature Cited

Brown, R.E.; Dickson, J.G. 1994. Swainson's warbler (*Limnothlypis swainsonii*). In: Poole, A.; Gill, F., eds. The birds of North America. Number 126. Philadelphia, PA: The Academy of Natural Sciences. 20 p.

Conner, R.N.; Dickson, J.G. 1997. Relationships between bird communities and forest age, structure, species composition and fragmentation in the West Gulf Coastal Plain. Texas Journal of Science. 49(3) Supplement: 123-138.

Coulston, J.W.; Reams, G.A.; McRoberts, R.E. *et al*. 2004. Practical considerations when using perturbed forest inventory plot locations to develop spatial models: a case study. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C. *et al*., eds. Proceedings, sixth annual forest inventory and analysis symposium. Gen. Tech. Rep. WO-70. Washington, DC: U.S. Department of Agriculture, Forest Service: 81-86.

Dettmers, R.; Buehler, D.A.; Bartlett, J.B. 2002. A test and comparison of wildlife-habitat modeling techniques for predicting bird occurrence at a regional scale. In: Scott, J.M.; Heglund, P.J.; Morrison, M.L. *et al*., eds. Predicting species occurrences: issues of accuracy and scale. Washington, DC: Island Press. 54: 607-616.

Donovan, T.M.; Jones, P.W.; Annand, E.M. *et al*. 1997. Variation in local-scale edge effects: mechanisms and landscape context. Ecology. 78(7): 2064-2075.

Driscoll, M.J.L.; Donovan, T.; Mickey, R. *et al*. 2005. Determinants of wood thrush nest success: a multi-scale, model selection approach. Journal of Wildlife Management. 69(2): 699-709.

Eddleman, W.R.; Evans, K.E.; Elder, W.H. 1980. Habitat characteristics and management of Swainson's warbler in southern Illinois. Wildlife Society Bulletin. 8(3): 228-233.

Gesch, D.; Oimoen, M.; Greenlee, S. *et al*. 2002. The National Elevation Dataset. Photogrammetric Engineering & Remote Sensing. 68(1): 5-11.

Graves, G.R. 2002. Habitat characteristics in the core breeding range of the Swainson's warbler. Wilson Bulletin. 114(2): 210-220.

Hamel, P.B. 1992. Land manager's guide to the birds of the South. Chapel Hill, NC: The Nature Conservancy, Southeastern Region. 437 p.

Jenness, J. 2006. Topographic Position Index (tpi_jen.avx) extension for ArcView 3.x. Version 1.2. Flagstaff, AZ: Jenness Enterprises. http://www.jennessent.com/arcview/tpi.htm. (8 January 2007).

Keys, J.E., Jr.; Carpenter, C.A.; Hooks, S. *et al*. 1995. Ecological units of the eastern United States: first approximation. Atlanta, GA: U.S. Department of Agriculture, Forest Service, Southern Region. 83 p. [Map and booklet of map unit tables].

Larson, M.A.; Dijak, W.D.; Thompson, F.R., III; Millspaugh, J.J. 2003. Landscape-level habitat suitability models for twelve wildlife species in southern Missouri. Gen. Tech. Rep. NC-233. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station. 51 p.

Law, J.R.; Johnson, P.S.; Houf, G. 1994. A crown cover chart for oak savannas. Tech. Brief NC-2. Columbia, MO: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station. 2 p.

Morrison, M.L. 1998. Modeling wildlife-habitat relationships. In: Morrison, M.L.; Marcot, B.G.; Mannan, R.W., eds. Wildlife-habitat relationships: concepts and applications. 2nd ed. Madison, WI: The University of Wisconsin Press. 309-349. Chapter 10.

Panjabi, A.O.; Dunn, E.H.; Blancher, P.J. *et al*. 2005. The partners in flight handbook on species assessment. Version 2005. Partners in Flight Technical Series No. 3. http://www.rmbo.org/pubs/downloads/Handbook2005.pdf. (8 January 2007).

Rich, T.D.; Beardmore, C.J.; Berlanga, H. *et al*. 2004. Partners in flight North American landbird conservation plan. Ithaca, NY: The Cornell Lab of Ornithology. 84 p.

Rosenberg, K.V.; Hames, R.S.; Rohrbaugh, R.W. *et al*. 2003. A land manager's guide to improving habitat for forest thrushes. Ithaca, NY: The Cornell Lab of Ornithology. 29 p.

Shamberger, M.A.; Farmer, A.H.; Terrel, J.W. 1982. Habitat suitability index models: introduction. FWSI-OBS-82/10. Fort Collins, CO: U.S. Department of the Interior, Fish & Wildlife Service. 2 p.

Twedt, D.J.; Loesch, C.R. 1999. Forest area and distribution in the Mississippi alluvial valley: implications for breeding bird conservation. Journal of Biogeography. 26(6): 1215-1224.

U.S. Geological Survey (USGS). 1999. The National Hydrography Dataset. U.S. Geological Survey Fact Sheet 106-99. Reston, VA: U.S. Department of the Interior, U.S. Geological Survey. 2 p.

Estimating Dead Wood During National Forest Inventories: A Review of Inventory Methodologies and Suggestions for Harmonization

Christopher W. Woodall[1], Jacques Rondeux[2], Pieter J. Verkerk[3], and Göran Ståhl[4]

Abstract.—Efforts to assess forest ecosystem carbon stocks, biodiversity, and fire hazards have spurred the need for comprehensive assessments of forest ecosystem dead wood (DW) attributes around the world. Currently, information regarding the prevalence, status, and methods of DW inventories occurring in the world's forested landscapes is scattered. The goal of this study is to describe the current status of DW, including DW attributes measured, sample methods employed, and DW attribute thresholds used by national forest inventories (NFI) that currently inventory DW around the world. Study results indicate that most countries do not inventory forest DW. Only 13 percent of all countries inventory DW, and sample methods and DW component definitions are diverse. The major commonality among DW inventories was that most countries have only just begun DW inventories and employ very low sample intensities. Harmonizing NFI DW inventories will be a major hurdle due to differences in population definitions, lack of clarity on sample protocols and estimation procedures, and sparse availability of inventory data and reports are some of the inconsistencies. Increasing database and estimation flexibility, developing common dimensional thresholds, publishing inventory procedures and protocols, releasing inventory data and reports to international peer review, and increasing communication (e.g., workshops) among countries inventorying DW are suggestions forwarded by this study to increase NFI DW harmonization.

Importance of National-Scale Inventories of Dead Wood

Dead wood (DW) is typically defined as all nonliving tree biomass (excluding duff and litter), including woody debris that is standing or lying along with stumps (FAO 2006). National-scale inventories national forest inventories (NFIs) of forest ecosystem DW are critical to four broad scientific pursuits: carbon accounting, fire/fuels, biodiversity, and wildlife habitat. Carbon (C) sequestration is becoming an increasingly important estimate derived from NFIs because of the link between greenhouse gases accumulation in the atmosphere and possible climate change (Smith *et al.* 2004a). In 1992, 150 countries signed the United Nations Framework Convention on Climate Change, which requires annual reports of greenhouse gas inventories, including C in forests. In 2006, approximately 11 percent of all greenhouse gas emissions in the United States were sequestered annually in forests and forest products (Smith *et al.* 2004b, EPA 2006). In the United States, 35 percent of the total forest C pool is in live vegetation, 52 percent in the soil, and 14 percent in dead organic material (excluding fine woody debris [FWD]) (Heath *et al.* 2003). Therefore, accurately estimating baseline forest DW carbon stocks and monitoring stock changes over time is essential. Even so, estimates of DW have been omitted from some large-scale C assessments (Goodale *et al.* 2002) due to the lack of sufficient inventory data.

[1] Research Forester, U.S. Department of Agriculture, Forest Service, Northern Research Station, St. Paul, MN 55108. Phone: 651–649–5141; fax: 651–649–5140. E-mail: cwoodall@fs.fed.us.

[2] Professor, Gembloux Agricultural University, Gembloux, Belgium.

[3] Researcher, European Forest Institute, Joensuu, Finland.

[4] Professor, Swedish University of Agricultural Sciences, Umeå, Sweden.

Concerns about the increase of forest fire occurrences at national scales have brought attention to the critical role that DW plays in large-scale fire hazards. Estimates of DW are integral to numerous fire behavior models (for examples, see Albini 1976, Burgan and Rothermal 1984, Finney 1998, Reinhardt *et al.* 1997, Rothermal 1972). NFIs of DW can be used to estimate fuel loads and fire dangers at national scales (for example, see Woodall *et al.* 2005).

DW components, such as standing dead trees and coarse woody debris (CWD), increase a forest's structural heterogeneity and serve as critical habitat for numerous flora and fauna. Flora uses the microclimate of moisture, shade, and nutrients provided by CWD to establish regeneration (Harmon *et al.* 1986). Both standing and down DW provide a diversity (e.g., stages of decay, size classes, and species) of habitat for fauna (ranging from large mammals to invertebrates) (Bull *et al*. 1997, Heilmann-Clausen and Christensen 2005, Harmon *et al.* 1986, Maser *et al.* 1979, Siitonen 2001). Due to the possibility of dwindling habitat for many native species across many countries, inventories of DW are important for habitat assessments and wildlife conservation efforts (for examples, see Ohmann and Waddell 2002, Tietje *et al.* 2002). Volume of standing and lying DW has also been adopted as a pan-European indicator for sustainable forest management related to forest biodiversity (MCPFE 2002).

Given the importance of NFI DW inventories, the goal of this article is to broadly describe the current DW NFI methods used around the world and suggest opportunities for harmonization. Specific goals include the following: (1) to describe the current status (e.g., year of first inventory, number of plots and transects, publicly available data) of DW NFIs, (2) to describe the dead wood attributes (e.g., standing dead trees or FWD) inventoried in current DW NFIs, (3) to describe briefly the DW sample techniques (e.g., fixed area plots or line-intersect transects) used in NFIs, and (4) to suggest opportunities for international harmonization of DW NFIs.

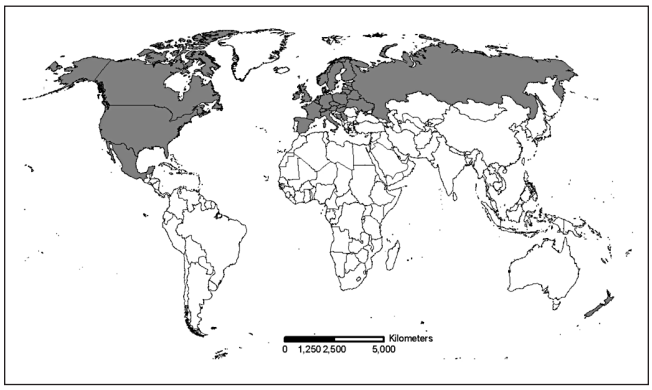
Study Survey Methods

Nearly 50 countries that were deemed most likely to have a DW inventory based on expert knowledge were contacted. For example, countries such as Libya or Afghanistan were unlikely to have DW inventories. Contacts for each country were based on advice from forestry colleagues and a list of participants in the Food and Agriculture Organization (FAO) of the United Nations's Global Forest Resource Assessment (FAO 2006). It was assumed that countries that did not respond to the DW survey either did not have a DW inventory, did not want to participate in the survey, or did not understand the English survey. Despite the authors' best attempts to accurately estimate the prevalence of DW sampling around the world, the results of this study's survey most likely underestimate the intensity of global DW sampling. The survey consisted of 21 questions in spreadsheet format grouped into four sections, which included (1) current status of inventory, (2) DW attributes inventoried, (3) inventory methods, and (4) attribute thresholds. All submitted surveys were summarized broadly so that individual countries could not be identified. For the purposes of the survey, in order to differentiate FWD from CWD, it was defined as downed and dead woody debris with a diameter less than 7 cm.

Current Status of Dead Wood National Inventories

This survey identified only 30 countries that currently inventory DW, representing only 13 percent of the world's 229 countries (FAO 2006) (fig. 1). Because most countries that inventory DW are located in the heavily forested regions of Europe and North America, this survey found that more than 41 percent of the forestland of the earth is inventoried for DW. DW inventories are a relatively recent phenomenon for most countries. More than 77 percent of DW inventories were initiated from 2000. The sample intensity (number of forested acres in any country divided by the number of DW inventory plots) varied widely with most countries having an intensity greater than one plot per 10,000 forested ha. Almost all countries had

Figure 1.—*Countries that currently have a national forest inventory of dead wood trees (shaded gray).*



an interval between plot remeasurement less than or equal to 10 years. Approximately 80 percent of countries have not publicly released their DW inventory data, while nearly 87 percent have not summarized the inventory in an official report. Overall, it appears that the majority of forests on earth are not inventoried for DW. In almost all countries where there is a history of forest inventories (e.g., Sweden and Germany), DW is inventoried. The analysis, dissemination, and review of these DW inventories is lacking given the dearth of publicly available DW inventory data and reports.

Dead Wood Components Measured in National Forest Inventories

Almost all countries that had a DW inventory compiled information on both standing dead and down dead trees (e.g., CWD). Of countries that had a DW inventory, 60 percent inventoried stumps, 73 percent inventoried residue piles, and 47 percent inventoried FWD. Almost all countries measured the species and decay class of DW. Notably, 68 percent of countries had a four- or five-decay-class rating system for DW. Overall, it appears that most DW inventories sample CWD, standing dead trees, stumps, and residue piles. Fine woody debris was only inventoried by 46 percent of surveyed countries. This lack of fine wood inventories may be because this DW component is being partially inventoried as part of the forest floor and because its contribution to the total forest biomass is relatively minor.

Dead Wood Sample Methods and Attribute Thresholds

Almost all countries used fixed-area plots for inventorying standing dead trees, but sample methods for dead and downed woody debris were more varied. Sixty-three percent of countries used fixed-area plots for CWD and 19 percent used line-intersect sampling. The remainder of countries used variable-radius plots or ocular estimation (i.e., expert observation/classification). The sample technique for FWD was evenly split between fixed-area plots (quadrat or fixed-radius) and line-intersect sampling. At the country-level, fixed-area plots were the most common DW inventory method, regardless of DW component. In terms of global forestland area, however, nearly 16 percent of the world's forest CWD is inventoried using line-intersect sampling techniques. In contrast, only approximately 3 percent of the world's forest CWD is inventoried using fixed-area plots. It appears that in countries with relatively large expanses of forest area (e.g., Canada and the United States) line-intersect sampling is the method of choice.

The definitions of DW variables, predominantly defined by measurement thresholds, varied among countries that inventory DW. Common minimum diameters at breast height for standing dead trees were 5, 7, 10, and 12 cm. A minimum diameter of 10 cm was the most common minimum diameter, with 19 percent of all countries having that minimum diameter. For CWD, 33 percent of countries inventoried CWD with a minimum diameter of 7.0 or 7.6 cm. The frequent minimum diameter was still 10 cm, however, with 27 percent of countries using that threshold. The threshold of 7.0 or 7.6 cm relates to a common break point between fine and CWD. A diameter of 7.6 cm is close to the English measurement unit of 3 in, which is used to differentiate between fine and heavy fuels in fuel and fire behavior models (Deeming *et al.* 1977). Minimum heights or lengths for standing and dead downed trees were overwhelmingly either 1.0 or 1.3 m, respectively. Dimensional thresholds for stumps were the most varied with minimum diameters appearing to be larger than for the standing and downed dead trees. Some countries that had a minimum diameter of 12 cm for standing dead trees had a minimum stump diameter of 30 cm. This result indicates that because stumps contain less biomass per cm of

diameter than standing dead trees, a larger stump diameter is needed to justify the effort to measure it. Most countries (53 percent) did not inventory FWD or even define it as a separate class of DW. As mentioned previously, the minimum diameter for CWD often defines the maximum diameter for FWD. Thus, 7.0 and 7.6 cm was the most common maximum diameter for FWD. Just a few countries specified minimum diameters for FWD, often 1.5 or 2.5 cm. Overall, the thresholds for DW components, in most cases, appear to be based on the relationship between sampling efficiency and the relative contribution of the DW component to overall stand biomass. Because the sampling of standing dead trees is probably the most efficient, along with being a major contributor to stand biomass, the population definition was the most inclusive (i.e., smallest minimum diameter). In contrast, either FWD was often not measured or its population was narrowly defined.

The Future of Dead Wood National Inventories

When viewing DW NFIs holistically, numerous similarities appear among them. First, standing dead and downed trees are often measured in unison. Rarely does a country inventory standing dead trees but not downed trees. Second, the size, species, and decay class of dead trees are ubiquitously measured. Most countries recognize the need to measure these parameters in order to more accurately estimate dead tree attributes such as volume, biomass, or carbon. Third, most countries have only recently started inventorying DW. This phenomenon can be most likely attributed to the relatively recent focus on national forest carbon stocks and indicators of biodiversity related to international agreements (e.g., greenhouse gas offset accounting, the Montreal Process indicators of sustainability, Pan-European Ministerial Conference on the Protection of Forests in Europe indicators for sustainable forest management). Fourth, fixed-radius sampling techniques were the most common technique for inventories of both standing and downed dead trees. Fixed-radius techniques were most likely adopted as efficient and logical extensions of fixed-radius techniques commonly used to inventory standing live trees. Fifth, most countries conducting DW inventories have neither publicly released their data nor summarized findings in a national report. These DW inventories are a recent activity for many countries, so it is likely that datasets are not complete or analytical expertise has not yet been developed. Finally, the remeasurement periods for DW NFIs is almost always 10 years or less, indicating countries' dedication to monitoring DW resources.

Despite the broad similarities among countries that inventory DW, even slight differences can cause problems with combining and comparing estimates in a regional/global context such as those required by global greenhouse gas offset accounting programs. The most prominent difference that can inhibit DW estimate comparison among countries is that of DW component population definitions. If countries use separate minimum diameters for either standing or downed dead trees then their resulting estimates are for different populations. At least two solutions can resolve this issue: common thresholds and database or estimation flexibility. Another apparent discrepancy was that of the number of DW components measured. Not all countries that inventory standing and downed dead trees also inventory stumps, residue piles, or FWD. Thus, national DW estimates may be incomparable. Total DW resource estimates may only be compared if the same DW components are measured or a common reporting framework is explicitly defined. The inherent nature of DW resources is that of decay and transition from standing dead trees, to coarse/FWD, to soil organic matter. Not only are dimensional thresholds (i.e., minimum diameter) important to define DW populations, defining the transition from standing dead to downed dead trees is important. How close to horizontal does a standing dead tree need to lean to be considered a downed dead tree? Finally, the force that may be driving DW inventories in different directions is the diversity of user groups demanding DW inventory information. For some countries, the main purpose of a DW inventory may be to assess fuel loadings, while in other countries it may be carbon accounting or biodiversity assessment. If inventory sample protocols are a reflection of inventory objectives, then the diversity of DW sample protocols reflects diversity in budgetary constraints and inventory objectives.

Forest DW inventories have expanded tremendously around the world during the past decade. Although numerous similarities bespeak a basis on ecological fundamentals (e.g., DW components of standing dead trees and CWD) and extensions

of historic standing live tree inventories (e.g., fixed-radius sample protocols), it is the inventory details that confound attempts to efficiently compare and combine DW resource estimates among countries. Differences in sample intensity, remeasurement period length, or sample technique (e.g., fixed-radius or line intersect) are not the major culprits in restricting global assessments. Almost all countries used defensible, peer-reviewed sample techniques for DW that should result in compatible estimates. It is the differences in population definitions, lack of sample protocol or estimation procedure clarity, and sparse availability of inventory data and reports that are the largest hurdles to harmonizating DW NFIs. Possible solutions to these problems include (1) increasing database or estimation flexibility to accommodate varying population definitions, (2) developing common dimensional thresholds, (3) publishing inventory procedures and protocols, (4) releasing inventory data and reports to international peer review, and (5) increasing communication (e.g., workshops or initiatives such as COST E43 in Europe) among countries inventorying DW. Given the substantial progress with DW inventories during the past years, there is little doubt that, with more effort and communication, these inventories may be more closely harmonized in the future.

Acknowledgments

The authors thank all national forest inventory correspondents that provided us with information on dead wood inventory methods. In addition, some survey responses were based on surveys already conducted by COST E43: "Harmonization of National Forest Inventories in Europe: Techniques for Common Reporting." Pieter J. Verkerk was financially supported by the European Commission through the SENSOR project (project no. 003874).

Literature Cited

Albini, F.A. 1976. Estimating wildfire behavior and effects. Gen. Tech. Rep. INT-30. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 92 p.

Bull, E.L.; Parks, C.G.; Torgersen, T.R. 1997. Trees and logs important to wildlife in the Interior Columbia River Basin. Gen. Tech. Rep. PNW-391. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 55 p.

Burgan, R.E.; Rothermel, R.C. 1984. BEHAVE: fire behavior prediction and fuel modeling system—FUEL subsystem. Gen. Tech. Rep. INT-167. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 126 p.

Deeming, J.E.; Burgan, R.E.; Cohen, J.D. 1977. The national fire-danger rating system. Res. Pap. INT-226. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 17 p.

Finney, M.A. 1998. FARSITE: Fire area simulator— model development and evaluation. Res. Pap. RMRS-RP-4. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 47 p.

Food and Agriculture Organization (FAO) of the United Nations. 2006. Global forest resources assessment, 2005. Forestry Pap. 147. Rome, Italy: Food and Agriculture Organization of the United Nations.

Goodale, C.L.; Apps, M.J.; Birdsey, R.A.; Field, C.B.; Heath, L.S. *et al.* 2002. Forest carbon sinks in the northern hemisphere. Ecological Applications. 12: 891-899.

Harmon, M.E.; Franklin, J.F.; Swanson, F.J. *et al.* 1986. Ecology of coarse woody debris in temperate ecosystems. Advanced Ecology Research. 15: 133-302.

Heath, L.S.; Smith, J.E.; Birdsey, R.A. 2003. Carbon trends in U.S. forest lands: 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 US forest soils to sequester carbon and mitigate the greenhouse effect. Boca Raton, FL: CRC Press: 35-45.

Heilmann-Clausen, J.; Christensen, M. 2005. Wood-inhabiting macrofungi in Danish beech-forests—conflicting diversity patterns and their implications in a conservation perspective. Biological Conservation. 122: 633-642.

Maser, C.; Anderson, R.G.; Cromack, K., Jr. *et al.* 1979. Dead and down woody material. In: Thomas, J.W., tech. ed. Wildlife habitats in managed forests: the Blue Mountains of Oregon and Washington. Agric. Handb. 553. Washington, DC: U.S. Department of Agriculture, Forest Service: 78-95.

Ministerial Conference on the Protection of Forests in Europe (MCPFE). 2002. Improved pan-European indicators for sustainable forest management as adopted by the MCPFE Expert Level Meeting. Vienna: Ministerial Conference on the Protection of Forests in Europe, Liaison Unit.

Ohmann, J.L.; Waddell, K.L. 2002. Regional patterns of dead wood in forested habitat of Oregon and Washington. In: Laudenslayer, W.F.G., Jr.; Shea, P.J.; Valentine, B.E.; Weatherspoon, C.P.; Lisle, T.E., tech. coords. Proceedings, symposium on the ecology and management of dead wood in western forests. Gen. Tech. Rep. PSW-GTR-181. Albany, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station: 535-560.

Reinhardt, E.D.; Keane, R.E.; Brown, J.K. 1997. First order fire effects model: FOFEM 4.0, user's guide. Gen. Tech. Rep. INT-344. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station. 65 p.

Rothermel, R.C. 1972. A mathematical model for predicting fire spread in wildland fuels. Res. Pap. INT-115. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 40 p.

Siitonen, J. 2001. Forest management, coarse woody debris and saproxylic organisms: Fennoscandian boreal forests as an example. Ecological Bulletins. 49: 11-41.

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: 25-31.

Smith, J.E.; Heath, L.S.; Woodbury, P.B. 2004b. Forest carbon sequestration and products storage. In: U.S. agriculture and forestry greenhouse gas inventory. Tech. Bull. 1907. Washington, DC: U.S. Department of Agriculture, Office of the Chief Economist, Global Change Program Office: 80-93.

Tietje, W.D.; Waddell, K.L.; Vreeland, J.K.; Bolsinger, C.L. 2002. Coarse woody debris in oak woodlands of California. Western Journal of Applied Forestry. 17: 139-146.

U.S. Environmental Protection Agency (EPA). 2006. Inventory of U.S. greenhouse gas emissions and sinks: 1990–2004. EPA 430-R-06-002. http://yosemite.epa.gov/oar/globalwarming.nsf/content/ResourceCenterPublicationsGHGEmissionsUSEmissionsInventory2006.html. (13 June).

Woodall, C.W.; Charney, J.J.; Liknes, G.C.; Potter, B.E. 2005. What is the fire danger now? Linking fuel inventories with atmospheric data. Journal of Forestry. 103: 293-298.

Estimating the Quadratic Mean Diameter of Fine Woody Debris for Forest Type Groups of the United States

Christopher W. Woodall[1] and Vicente J. Monleon[2]

Abstract.—The Forest Inventory and Analysis program of the Forest Service, U.S. Department of Agriculture conducts a national inventory of fine woody debris (FWD); however, the sampling protocols involve tallying only the number of FWD pieces by size class that intersect a sampling transect with no measure of actual size. The line intersect estimator used with those samples requires a measurement of the quadratic mean diameter (QMD). Published information regarding appropriate QMDs by FWD diameter class by species or forest types across the United States is lacking. Thus, the objective of this study is to employ a technique known as the graphical estimation (GE) method for estimating FWD QMDs for major forest types across the United States. Results indicate that the GE method, along with adaptations proposed in this study, allows for rapid estimation of FWD QMDs without additional fieldwork. The value (–1.81) of the scaling power function between the number of woody pieces by size class closely matches published scaling values (–2.00) between bole and fine branch diameters. Validation results indicate that estimates of QMD from the GE method differ substantially from published QMDs; however, published QMD results differ from one another to an even greater extent. Given the mixing of FWD from a diversity of tree and shrub species in the Nation's forests, this study concludes that a set of general QMDs for FWD across the United States should be used in the absence of empirically derived or power function-based QMD estimates.

A National Inventory of Fine Woody Debris

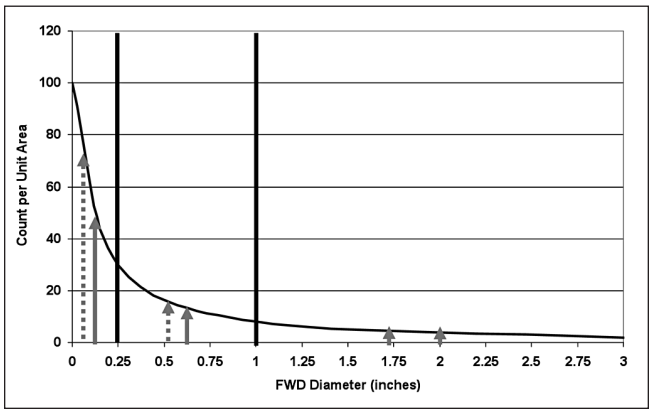
The Forest Service's Forest Inventory and Analysis (FIA) program conducts a national inventory of fine woody debris (FWD). FWD is defined by the FIA program as dead and downed woody debris in forests that is less than 3 in in diameter (USDA Forest Service 2006). Estimates of FWD volume and biomass based on the FIA inventory are highly desired by the fire and fuel and carbon science communities. FWD is a substantial component of fuel loadings and determines, to a large extent, fire behavior (Albini 1976, Burgan and Rothermel 1984, Deeming *et al.* 1977). FWD, as a component of the forest floor, is often reported as a forest carbon stock (IPCC *et al.* 1997, Smith *et al.* 2006).

FIA's sample protocol for FWD involves tallying the number of FWD pieces by size class that intersect a sampling plane. FIA protocols define three FWD size classes: small (0.01 to 0.24 in), medium (0.25 to 0.99 in), and large (1.00 to 2.99 in). The sample protocols do not require the measurement of the actual diameter of each FWD piece, only the number of pieces in each diameter class. In order to estimate FWD volume, the FWD quadratic mean diameter (QMD) for each size class is required. QMD is defined as the square root of the average squared diameter. Since the FWD volume estimator requires a QMD to be squared, any substantial errors in QMD approximation can lead to even larger errors in FWD volume estimates. Because the distribution of the number of FWD pieces as a function of FWD diameter is typically a negative exponential, using FWD size class midpoints in lieu of QMD may be inappropriate (solid arrows, fig. 1). Empirically derived FWD QMDs are available in the literature and offer a more defensible estimate of diameter. Unfortunately, published FWD QMDs are species specific and are only available for a very limited number of species and sites in the United States (e.g., see Nalder *et al.* 1997,

[1] Research Forester, U.S. Department of Agriculture (USDA) Forest Service, Northern Research Station, St. Paul, MN 55108. Phone: 651–649–5141; fax: 651–649–5140. E-mail: cwoodall@fs.fed.us.

[2] Mathematical Statistician, USDA Forest Service, Pacific Northwest Research Station, Portland, OR.

Figure 1.—*Hypothetical distribution of fine woody debris diameters.*



Note: Solid vertical lines define fine woody debris (FWD) size classes, solid arrows denote FWD diameter class midpoint, and dotted arrows denote midpoint of area under the FWD distribution curve.

1999; Roussoupolos and Johnson 1973; Van Wagtendonk *et al.* 1996). Given the predominantly mixed species condition of forests across the Nation, along with hundreds of tree and shrub species, empirically deriving FWD QMDs for every FIA FWD transect may be unfeasible. Hence, a substantial knowledge gap exists regarding FWD QMDs in forests of the United States.

An Alternative: The Graphical Estimation Method

Van Wagner (1982) was one of the first to identify the limitation of empirically derived FWD QMDs. As an alternative to conducting fieldwork for every population of interest (unique forest type across the country), Van Wagner (1982) proposed the graphical estimation (GE) method for deriving QMDs that involve no additional fieldwork. Van Wagner (1982) assumed that the distribution of FWD diameters follows a power law:

$$y = ax^b \quad (1)$$

where y is the number of FWD pieces, x is the diameter, and a and b are constants. Then, the theoretical QMD for a diameter class can be calculated analytically as (equation 5 of Van Wagner 1982):

$$(QMD)^2 = \int_{x_1}^{x_2} ax^{b+2}dx \Big/ \int_{x_1}^{x_2} ax^b dx = \frac{(b+1)\left(x_2^{b+3} - x_1^{b+3}\right)}{(b+3)\left(x_2^{b+1} - x_1^{b+1}\right)} \quad (2)$$

where QMD is the quadratic mean diameter and x_1 and x_2 are the lower and upper diameter class limits, respectively.

To estimate the QMD for each diameter class, the coefficient b has to be estimated from the data. To this effect, Van Wagner (1982) log-transformed the simple power equation (equation 1). He estimated b as the slope of the regression of the log of Y, the number of intersections per unit sample line and diameter class, on the log of X, the midpoint of the diameter class. The diameter classes were normalized to a unit width by dividing the frequency by the width of the diameter class. Thus, the slope of the regression model is used to determine the midpoint of the area under the estimated FWD diameter distribution (dotted arrows, fig. 1). Van Wagner's original hypothesis (1982) was only demonstrated for a few selected species, assuming that users would only be able to manually fit the GE models.

The goal of this study was to explore the possibility of using Van Wagner's GE method to estimate FWD QMDs across the United States with specific objectives including: (1) estimating FWD QMDs by forest type in the United States using a nonlinear power function to estimate b (equation 1) in Van Wagner's GE method (1982), (2) validating FWD QMD estimates where published FWD QMD estimates are available, and (3) discussing future options for approximating FWD QMDs when estimating FWD volumes in a national inventory of FWD.

Data and Analysis

A total of 9,788 observations from single forest condition FIA plots across the United States were included in this study. Each observation consisted of FWD and coarse woody debris (CWD) tally counts along 10 ft of the 150-degree transect (for further sample protocol information, see USDA 2006, Woodall and Williams 2005). FWD with transect diameters less than 0.25 in and 0.25 in to 1.00 in (1 and 10 hr, respectively) were tallied separately on a 6-ft slope-length transect (14 ft to 20 ft on the 150-degree transect) and standardized to a 10-ft length by multiplying counts by 1.666. FWD with transect diameters of 1.00 to 2.99 in (100-hr) were tallied on a 10-ft slope-length transect (14 ft to 24 ft on the 150-degree transect) (for more information on fuel class definitions, see Deeming *et al.* 1977). CWD, along the 10-ft section of the 150-degree transect with transect diameters within the size classes of 3.0 to 8.9 in and 9.0 to 27.0 in, were included as additional woody debris size

classes. Tally counts, standardized to a 10-ft sample transect by woody size class, were averaged by forest type group.

A simple power function formulated as a nonlinear model (equation 3) was fitted by forest type group in order to estimate QMD^2 by forest type group across the United States:

$$E(y) = ax^b \quad (3)$$

where E(.) is an expected value, y is the number of FWD pieces tallied in a particular size class, x is the diameter of the size class, and a and b are coefficients to be estimated. The b in equation 3 was used to solve for QMD in equation 2.

QMDs were validated by comparing selected forest type group QMD^2s to empirically derived QMD^2s (ponderosa pine, Douglas-fir, western larch; see table 1) and determining a mean absolute and relative difference. Because more than one published QMD^2 was available by species, a mean and coefficient of variation among published empirical estimates (Brown and Roussopoulos 1974, Roussopoulos and Johnson 1973, Ryan and Pickford 1978, Sackett 1980, Van Wagtendonk *et al.* 1996) were determined for each validation species.

Estimates of QMD^2 by Forest Type Group

The approximate R-square for the fitted nonlinear model exceeded 0.80 for all forest type groups. Estimates of b (equation 2) used to determine the midpoint of the area under the nonlinear curve (Van Wagner 1982) did not vary greatly by forest type group (table 2). Consequently, estimated QMD^2s did not vary greatly by forest type group (table 2). QMD^2s for the smallest FWD did not vary to any degree, while QMD^2s for the largest FWD varied from 2.7 to 3.3. The mean QMD^2s for all forest type groups in the United States was 0.018, 0.262, and 3.101 for increasing FWD size classes, respectively. If the square of the midpoint of the diameter class is used, the estimates of QMD^2 would have been 0.016, 0.384, and 3.98.

Validation of QMD^2s

The differences between the selected forest type groups' QMD^2 estimate and the empirically derived QMD^2s for selected species were substantial (table 3). The differences in estimates may be attributed to the possible inability of the GE method to adequately represent natural interspecific variation in FWD sizes. The predicted QMD^2s are based on an estimated mean. This conclusion, however, cannot be strongly supported by the validation results since it may be inappropriate to compare forest type group results from a wide geographic area to individual species from a particular site. Across the United States, the mixing of FWD from a diversity of tree and shrubs species obscures interspecific differences, thus a substantial difference in FWD QMDs between forest type groups should not be expected. Furthermore, the coefficients of variation between empirical studies themselves often exceeded differences between empirical studies and the GE method results (table 3). For example, although this study's QMD^2 estimates for small Douglas-fir FWD varied from the literature by 67 percent, the coefficient of variation among the published QMD^2s exceeded 115 percent.

Table 1.—*Means of published estimates of FWD QMD^2s used in validation procedures.*

Forest type	Small FWD		Medium FWD		Large FWD	
	Mean QMD^2	Citations	Mean QMD^2	Citations	Mean QMD^2	Citations
All types	0.045	All	0.451	All	3.976	All
Ponderosa	0.087	All	0.434	All	4.285	BR 1974, S 1980,V 1996
Douglas-fir	0.049	All	0.419	All	2.880	BR 1974, S 1980,V 1996
Western larch	0.078	BR 1974, RP 1978	0.335	BR 1974, RP 1978	1.472	BR 1974

FWD = fine woody debris. QMD = quadratic mean diameter.
Sources: Brown and Roussopoulos 1974 (BR 1974), Ryan and Pickford 1978 (RP 1978), Sackett 1980 (S 1980), Van Wagtendonk et al. 1996 (V 1996)

Table 2.—*Estimates of QMD² by forest type group across the United States based on the graphical estimation method.*

Forest type group	b	QMD²		
		Small	Medium	Large
White/red/jack pine	– 1.9106	0.017	0.254	3.043
Spruce/fir	– 1.8861	0.018	0.256	3.058
Longleaf/slash pine	– 1.2514	0.020	0.310	3.457
Loblolly/shortleaf pine	– 1.5277	0.019	0.286	3.279
Pinyon/juniper	– 1.8785	0.018	0.257	3.062
Douglas-fir	– 2.1937	0.016	0.233	2.879
Ponderosa pine	– 1.4478	0.019	0.293	3.330
Fir/spruce/mountain hemlock	– 2.0054	0.017	0.247	2.987
Lodgepole pine	– 1.9952	0.017	0.248	2.993
Hemlock/Sitka spruce	– 1.8624	0.018	0.258	3.072
Western larch	– 1.6936	0.018	0.272	3.175
Redwood	– 1.7481	0.018	0.267	3.141
Other western softwoods	– 1.6366	0.019	0.277	3.210
California mixed conifer	– 1.6981	0.018	0.271	3.172
Oak/pine	– 1.5927	0.019	0.280	3.238
Oak/hickory	– 1.6870	0.018	0.272	3.179
Oak/gum/cypress	– 1.5227	0.019	0.286	3.282
Elm/ash/cottonwood	– 1.5926	0.019	0.280	3.238
Maple/beech/birch	– 1.7763	0.018	0.265	3.124
Aspen/birch	– 1.6121	0.019	0.279	3.226
Alder/maple	– 1.6665	0.018	0.274	3.192
Western oak	– 1.9738	0.017	0.249	3.005
Tanoak/laurel	– 2.0071	0.017	0.247	2.986
Other western hardwoods	– 2.5979	0.015	0.206	2.661
All types	– 1.8143	0.018	0.262	3.101

QMD = quadratic mean diameter.

Table 3.—*Mean absolute (Abs) and relative difference (Rel. diff.) between estimates of FWD QMD²s derived from this study's GE method and mean of published QMD²s (n = number of empirically estimated QMD²s) along with the coefficient of variation between the published QMD²s (CV).*

Forest type group	Small FWD				Medium FWD				Large FWD			
	Abs. (inches)	Rel. diff. (percent)	n	CV	Abs. (inches)	Rel. diff. (percent)	n	CV	Abs. (inches)	Rel. diff. (percent)	n	CV
All groups	0.028	48.6	5	NA	0.190	39.4	5	NA	1.293	33.1	3	NA
Ponderosa	0.069	78.2	4	83.1	0.141	32.5	4	34.89	0.955	22.3	3	70.26
Douglas-fir	0.033	67.3	4	115.6	0.186	44.4	4	44.4	0.001	0.0	3	56.2
Western larch	0.060	76.9	2	98.8	0.063	18.7	2	64.9	1.703	115.7	1	NA

FWD = fine woody debris. GE = graphical estimation. QMD = quadratic mean diameter.

Future Options

Budgetary constraints most likely limit the ability of large-scale forest inventories to actually measure the line intersect diameter of FWD pieces. Therefore, tallying the number of FWD pieces by size class has become the widely accepted method for inventorying FWD and estimating volume. Van Wagner's (1982) GE method provides a theoretical approach for estimating FWD QMDs by defined populations (e.g., forest types). Our study demonstrated application of Van Wagner's approach for a national inventory of FWD. Substantial disparity exists between the estimated QMDs and the empirically derived QMDs, however. Therefore, three options for the future include (1) empirically measuring FWD QMDs by major forest

type group across the United States, (2) using the GE method to theoretically determine FWD QMDs across the United States, and (3) exploring the possibility of using stem or branch scaling factors to determine FWD QMDs. The value (–1.81) of the scaling power function between the number of woody pieces by size class across the United States closely matches published scaling values (–2.00 in Enquist 2002) between bole and fine branch diameters. Therefore, refining understanding of the scaling function between tree boles, limbs, and fine twigs may afford a new technique for estimating FWD QMDs.

Overall, although empirically derived QMDs may be superior in all cases, their determination is currently prohibitive for the entire United States because of budgetary limitations. Given the mixing of FWD from trees and shrubs in all the various forest types across the United States, it is suggested that generic FWD QMD defaults, such as those presented in this study, may be used in estimation procedures until either empirical or scaling information supplants these defaults.

Literature Cited

Albini, F.A. 1976. Estimating wildfire behavior and effects. Gen. Tech. Rep. INT-30. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 92 p.

Brown, J.K.; Roussopoulos, P.J. 1974. Eliminating biases in the planar intersect method for estimating volumes of small fuels. Forest Science. 20: 350-356.

Burgan, R.E.; Rothermel, R.C. 1984. BEHAVE: fire behavior prediction and fuel modeling system—FUEL subsystem. Gen. Tech. Rep. INT-167. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 126 p.

Deeming, J.E.; Burgan, R.E.; Cohen, J.D. 1977. The national fire-danger rating system. Res. Pap. INT-226. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 17 p.

Enquist, B.J. 2002. Universal scaling in tree and vascular plant allometry: toward a general quantitative theory linking plant form and function from cells to ecosystems. Tree Physiology. 22: 1045-1064.

Intergovernmental Panel on Climate Change (IPCC); United Nations Environmental Programme (UNEP); Organization for Economic Co-Operation and Development (OECD); International Energy Agency (IEA). 1997. Revised 1996 IPCC guidelines for national greenhouse gas inventories. Paris, France.

Nalder, I.A.; Wein, R.W.; Alexander, M.E.; de Groot, W.J. 1997. Physical properties of dead and downed round-wood fuels in the boreal forests of Alberta and Northwest Territories. Canadian Journal of Forest Research. 27: 1513-1517.

Nalder, I.A.; Wein, R.W.; Alexander, M.E.; de Groot, W.J. 1999. Physical properties of dead and downed roundwood fuels in the boreal forests of western and northern Canada. International Journal of Wildland Fire. 9: 85-99.

Roussopoulos, P.J.; Johnson, V.J. 1973. Estimating slash fuel loading for several lake states tree species. Res. Pap. RP-88. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Experiment Station.

Ryan, K.C.; Pickford, S.G. 1978. Physical properties of woody fuels in the Blue Mountains of Oregon and Washington. Res. Note PNW-315. Portland, OR: U.S. Department of Agriculture, Forest Service.

Sackett, S.S. 1980. Woody fuel particle size and specific gravity of southwestern tree species. Res. Note RM-389. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

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-GTR-343. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station.

U.S. Department of Agriculture (USDA) Forest Service. 2006. Forest inventory and analysis national core field guide: Washington, DC: U.S. Department of Agriculture, Forest Service, Northern Research Station. Vol. II.

Van Wagner, C.E. 1982. Graphical estimation of quadratic mean diameters in the line intersect method. Forest Science. 28: 852-855.

Van Wagtendonk, J.W.; Benedict, J.M.; Sydoriak, W.M. 1996. Physical properties of woody fuel particles of Sierra Nevada conifers. International Journal of Wildland Fire. 6: 117-123.

Woodall, C.W.; Williams, M.S. 2005. Sample protocol, estimation procedures, and analytical guidelines for the down woody materials indicator of the FIA program. Gen. Tech. Rep. GTR-256. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station. 47 p.

Attributes of Standing Dead Trees in Forests of the United States

Christopher W. Woodall[1], James E. Smith[2], and Patrick D. Miles[3]

Abstract.—Standing dead trees in forests of the United States serve as wildlife habitat, a fuel loading component, and carbon stocks. Although standing dead trees are a vital component of forest ecosystems, information regarding this resource across the Nation is lacking. The first annual inventory of standing dead trees across the United States was initiated in 1999, resulting in a comprehensive assessment of this resource. The goal of this study is to broadly summarize the attributes of standing dead trees across the United States using the first national inventory of standing dead trees. Study objectives were to examine volume and biomass estimates by geographic regions, diameter at breast height/decay class distributions, and species composition. Results indicate that a substantial number of standing dead trees exists in forests across the United States, exceeding 7 billion nationwide. Rocky Mountain and Pacific Northwest forests have some of the highest volume and biomass of standing dead trees, while southeastern forests have the least. The species composition of standing dead trees is quite diverse, with 26 species groups having more than a billion trees nationwide. Overall, standing dead trees are a prevalent component of forests across the United States.

Why Are Standing Dead Trees Important?

Standing dead trees, sometimes referred to as snags, are remnants of once living trees that are still self-supported and leaning less than 45 degrees from vertical (as defined by the Forest Service, U.S. Department of Agriculture's Forest Inventory and Analysis (FIA) program [USDA Forest Service 2006]). Standing dead trees are a substantial component of fuel loadings. The total biomass of standing dead trees in some forests may exceed that of downed and dead woody debris (Kirby *et al.* 1998); in such cases, standing dead fuels constitute a substantial fire hazard. In addition, standing dead trees fuels serve as a fuel ladder to upper crown fuels (Stephens 1998), and may be an important predictor of down woody debris through fuel succession (Schimmel and Granstrom 1997). On the other hand, standing dead trees are a component of healthy forest ecosystems, serving as wildlife habitat and increasing stand structural diversity. Standing dead trees serve as critical habitat for numerous wildlife species including a variety of avian species (Raphael and White 1984). In addition, the decaying substrate of standing dead trees provides critical habitat to forest invertebrate species (Harmon *et al.* 1986). Finally, dead wood is often a reporting component for forest carbon pools in national assessments. The Intergovernmental Panel on Climate Change of the United Nations calls for yearly reporting of dead wood carbon stocks, of which standing dead trees are a considerable component (e.g., see EPA 2004). Overall, standing dead trees are an integral component of forest ecosystems.

A National Inventory of Standing Dead Trees

Very little analysis regarding standing dead wood resources across the United States exists. In the past, most standing dead tree analyses were at local or regional scales (e.g., see Goodburn and Lorimer 1998, Cline *et al.* 1980, Healy *et al.* 1989) while national-scale forest resource analyses omitted dead tree attributes entirely (e.g., see Smith *et al.* 2004). The lack of national standing dead tree estimates was due to the lack of a nationally consistent standing dead tree inventory. Standing

[1] Research Forester, U.S. Department of Agriculture (USDA) Forest Service, Northern Research Station, St. Paul, MN 55108. Phone: 651–649–5141; fax: 651–649–5140. E-mail: cwoodall@fs.fed.us.

[2] Plant Physiologist, USDA Forest Service, Northern Research Station, Durham, NH 03824.

[3] Research Forester, USDA Forest Service, Northern Research Station, St. Paul, MN 55108.

dead trees have been infrequently inventoried during periodic inventories across the United States or have been inventoried only within specific FIA regions since the early 20th century. Since many of these inventories were only for determination of growing stock mortality, older dead trees were possibly omitted during the inventory. With the inception of a national annual forest inventory across the United States in 1999 (Gillespie 1999), uniform standing dead tree inventory protocols have been adopted allowing the first ever national assessment of standing dead trees.

Due to the availability of a national inventory of standing dead trees, an analysis and interpretation of inventory estimates is highly warranted. Therefore, the specific objectives of this study were to (1) determine the average biomass of standing trees by geographic region across the United States, (2) determine the diameter at breast height (d.b.h.) (1.4 m) and decay class distribution of standing dead trees nationally, (3) determine the species composition distribution of standing dead trees, and (4) suggest opportunities for development of a forest health indicator using standing dead tree information.

Data and Analysis

The FIA program conducts a 3-phase inventory of forest attributes of the United States (Bechtold and Patterson 2005). The FIA sampling design is based on a tessellation of the United States into hexagons, approximately 2,420 ha in size with at least one permanent plot established in each hexagon. In phase 1, the population of interest is stratified and plots are assigned to strata, such as forest, nonforest, and edge, to increase the precision of estimates. In phase 2, tree and site attributes are measured in forested conditions for plots established in the 2,428-ha hexagons. Phase 2 plots consist of four 7.32-m fixed-radius subplots on which standing dead trees are inventoried with measurement of numerous individual tree variables such as species, diameter, and total height (for more information, see USDA Forest Service 2006, Bechtold and Patterson 2005).

All standing dead data were from the most current, publicly available inventory for each State in the coterminous United States. All inventory data were from annual inventories conducted since 1999 except for the following states where periodic inventories were used in the study analyses: Mississippi (1994), New Mexico (1999), North Carolina (2002), Oklahoma (1993), and Wyoming (2000). The number of FIA plots used in this study totaled 87,401.

Mean volume (m^3/ha) and dry biomass (tonnes/ha) of standing dead trees on forestland in the United States were determined by geographic region. The total number of standing dead trees was determined for 10 cm d.b.h. classes and decay class (five classes; USDA Forest Service 2006). Decay class is a subjective determination of the stage of decay of a standing dead tree (USDA Forest Service 2006). A decay class one tree still has an upper crown with sapwood intact and minimal decay; whereas a decay class five trees has no branches remaining, absent sapwood, and often a broken top. The total volume (gross cubic volume, m^3) of standing dead trees was determined by selected species group for the entire United States. Finally, the ratio of the number of standing live trees to standing dead trees was determined by State across the United States. Population estimation procedures are detailed by Bechtold and Patterson (2005).

Biomass by Region

The western regions of the United States (Pacific Coast and Rocky Mountains) had the largest estimates of mean standing dead tree biomass per hectare (table 1). The States of California, Washington, and Oregon in the Pacific Coast region had

Table 1.—*Mean dry biomass (tonnes/ha) of standing dead trees on forestland in the United States by geographic region.*

Geographic region	Constituent states	Biomass (tonnes/ha)
Northern	CT, DE, IL, IN, IA, KS, ME, MA, MS, MI, MN, MO, NE, NH, NJ, NY, ND, OH, PA, RI, SD, VT, WV, WI	8.62
Pacific Coast	CA, OR, WA	12.50
Rocky Mountain	AZ, CO, ID, MT, NV, NM, UT, WY	11.08
Southeastern	AL, AR, FL, GA, KY, LA, MS, NC, OK, SC, TN, TX, VA	0.98

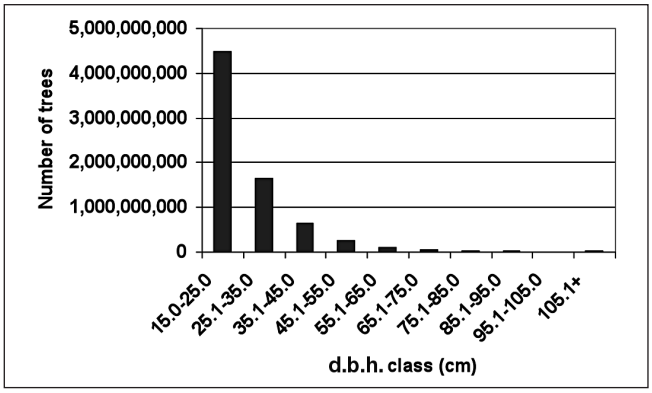
mean biomass/hectare estimates of 12.5 tonnes/ha compared to mean estimates of southeastern States that were less than 1 tonne/ha. These results indicate that standing dead trees are a prevalent component of forests across of the United States. The disparity in standing dead resources between eastern and western forests may be due to numerous speculative factors. First, the majority of eastern forests are privately owned where they may be more actively managed to reduce mortality and increase timber production. Western forests have a greater proportion of land area managed by Federal land agencies where timber production objectives may not take priority over other management objectives such as wildlife habitat maintenance (i.e., snag creation). Second, numerous areas of the western United States have endured prolonged drought (Cook *et al.* 2004) or experienced recent wildfires that may have increased mortality. Third, areas of the Pacific Northwest are some of the most productive forestland in the United States (Smith *et. al.* 2004); not only do these areas have high volumes of standing live trees but also standing dead trees. Ultimately, these disparities in standing dead tree biomass estimates are most likely due to a mix of biotic factors (e.g., insects/diseases), abiotic factors (e.g., droughts), and cultural and management practices (e.g., timber production versus wildlife habitat maintenance) that differ across the Nation.

Diameter and Decay Class Distribution

This study estimated more than 7 billion standing dead trees in forests of the United States. The d.b.h. distribution of this population is highly skewed towards smaller-sized trees (fig. 1). More than 4 billion standing dead trees have a d.b.h. between 15 and 25 cm in forests of the United States. In comparison, the total of standing dead trees with a d.b.h. of greater than 25 cm is less than 3 billion trees. Relative to the preponderance of the smaller-sized trees, the number of large-sized trees is considerably lower (d.b.h. greater than 55 cm). These trends are most likely indicative of suppression-related mortality and the natural "negative exponential" distribution of uneven-aged stands that can result in higher standing dead tree densities (Goodburn and Lorimer 1998). The observed trends are not indicative of any widespread forest health issues that may be detrimental to larger-sized trees (e.g., chestnut blight).

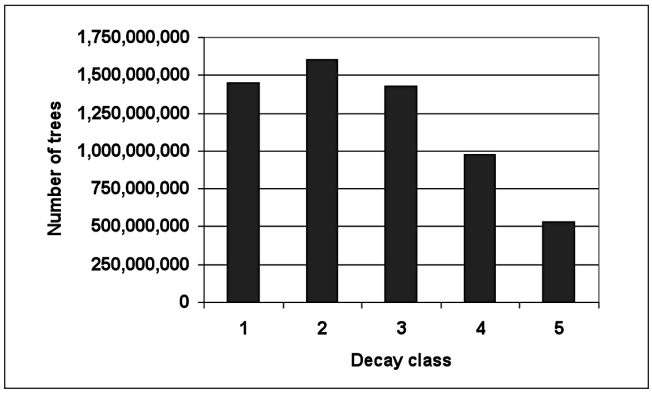
The distribution of standing dead trees by decay class is skewed toward less decayed trees (fig. 2). More than 1.5 billion standing dead trees in decay class two are spread across the Nation. Decay classes two and three are nearly equal, at nearly 1.4 billion trees nationwide. Comparatively, only 0.5 billion standing dead trees are in decay class five. These results are logical given the decay progression of deceased trees. A recently deceased tree will progress rather rapidly through decay class one, losing some bark and fine twigs. A standing dead tree may reside in decay classes two and three for some time depending on wind disturbances, microclimate, and abiotic factors (e.g., fungi and wildlife disturbance) (e.g., see Harmon 1982, Sun *et al.* 2004). Once a tree reaches the decay class of four and five, it is much more susceptible to windthrow with an inability to support its own weight. Overall, the current decay class distribution of standing dead trees across the United States appears to follow a natural progression of tree decay.

Figure 1.—*Distribution of standing dead trees by d.b.h. class (cm) across the United States.*



d.b.h. = diameter at breast height.

Figure 2.—*Distribution of standing dead trees by decay class across the United States.*



Species Composition

Given that western United States forest regions had relatively large amounts of standing dead tree biomass (table 1), western tree species groups dominate the species composition of standing dead trees nationally (table 2). With regard to total gross volume of standing dead trees across the United States, the top three species groups were true firs, Douglas-fir, and lodgepole pine, with more than 35 billion m^3 combined. Prevalent eastern tree species groups were that of other eastern soft hardwoods, other red oaks, and soft maple. Nationally, the species composition of standing dead trees is very diverse, with 26 species groups having an estimate of total gross volume exceeding 1 billion m^3. Whether mortality is the result of suppression mortality, drought, or insects/diseases, the question arises as to what amount of standing dead trees is a "healthy" amount? Standing dead trees are a valuable source of wildlife habitat and structural diversity, but, at the same time, they serve as an indicator of widespread tree mortality and fire hazard. If eastern tree species groups serve as an indicator, then amounts of standing dead trees may indicate regional forest health issues. The species group of other eastern soft hardwoods is the top eastern species group in terms of standing dead trees due to Dutch Elm disease killing so many American Elms in this species group. In addition, the second most prevalent eastern tree species group is that of other red oak—a species group suffering from regional oak decline for years (Thomas and Boza 1984). Although standing dead trees are a valuable resource across the country, they may serve as an indicator of accumulated forest health issues since they represent cumulative tree mortality reduced by site specific decay processes.

Standing Dead Trees as an Indicator of Forest Health

When examining standing dead trees across the country, it is apparent that certain regions may have more productive forests that are constantly experiencing a "turnover" of trees into standing dead trees. Through combination of both standing live and dead tree resources into a ratio, one might better ascertain an assessment of standing dead tree resources. The ratio of the number of standing live to standing dead trees ranged from 4 to 34 for States across the United States. Idaho had the lowest ratio at 4.3. The median ratio was approximately 11. In other words, the median forest in the United States has 11 live trees for

Table 2.—*Total gross volume (billions of cubic meters) of standing dead trees by selected species groups across the United States (species groups with total volume below 1.29 billion m3 not included)*

Species group	Constituent species examples*	Total volume (billions, m3)
True firs	Abies amabilis, Abies concolor, Abies procera, Abies grandis	13.15
Douglas-fir	Pseudotsuga menziesii	12.29
Lodgepole pine	Pinus contorta	10.25
Engelmann/other spruces	Picea engelmannii, Pivea breweriana	5.28
Other eastern soft hardwoods	Acer negundo, Aesculus glabra, Celtis occidentalis, Ulmus americana	4.33
Ponderosa/Jeffrey pines	Pinus jeffreyi, Pinus ponderosa	3.76
Other western soft hardwoods	Cupressus lawsoniana, Cupressus macrocarpa, Larix lyalli	3.71
Western cottonwood/aspen	Populus deltoids, Populus tremuloides	2.93
Western woodland soft hardwoods	Juniperus occidentalis, Pinus edulis	2.51
Other red oaks	Quercus coccinea, Quercus laurifolia	2.50
Spruce/balsam fir	Abies balsamea, Picea rubens	2.41
Western red cedar	Thuja plicata	2.09
Eastern cottonwood/aspen	Populus deltoids, Populus tremuloides	1.82
Soft maple	Acre rubrum, Acer saccharinum	1.67
Western hemlock	Tsuga heterophylla	1.58
Eastern white/red pine	Pinus resinosa, Pinus strobes	1.43
Select white oaks	Quercus alba, Quercus macrocarpa	1.37
Loblolly/shortleaf pines	Pinus echinata, Pinus taeda	1.37
Beech	Fagus grandifolia	1.36
Select red oaks	Quercus rubra, Quercus shumardii	1.29

** See Miles et al. (2001) for details.*

every standing dead tree. States with an extremely high ratio would indicate that constituent forests are heavily managed or are young with little potential for accumulation of wildlife habitat. States with extremely low ratios would indicate that constituent forests are dense or unmanaged (large wilderness areas) with the potential for catastrophic fires. Given the dual role that standing dead trees play in forest ecosystems (e.g., wildlife habitat versus fire hazard), the recent availability of national standing dead tree inventory data, and the confounding process of decay/turnover, continued exploration of standing dead tree data and subsequent development of forest health indicators is highly warranted.

Overall, standing dead trees are an abundant natural resource across the United States. This resource is not equally distributed, however, with western forests having more than five times as much standing dead tree biomass as the eastern forests. Amounts of standing dead trees are not necessarily indicative of unhealthy forests since they serve as critical wildlife habitat and increase forest structural diversity. Exploring the use of standing dead tree estimates in forest health indicators is strongly suggested for future research. A key research question to explore is, tree mortality is a natural process, but how much mortality is unnatural?

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, TN: U.S. Department of Agriculture, Forest Service. 85 p.

Cline, S.P.; Berg, A.B.; Wight, H.A. 1980. Snag characteristics and dynamics in Douglas-fir forests, western Oregon. Journal of Wildlife Management. 44: 773-786.

Cook, E.R.; Woodhouse, C.A.; Eakin, C.M.; Meko, D.M.; Stahle, D.W. 2004. Long-term aridity changes in the western United States. Science. 306: 1015-1018.

Gillespie, A.J.R. 1999. Rationale for a national annual forest inventory program. Journal of Forestry. 97: 16-20.

Goodburn, J.M.; Lorimer, C.G. 1998. Cavity trees and coarse woody debris in old-growth and managed northern hardwood forests in Wisconsin and Michigan. Canadian Journal of Forest Research. 28: 427-438.

Harmon, M.E. 1982. Decomposition of standing dead trees in the southern Appalachian Mountains. Oecologia. 52: 214-215.

Harmon, M.E.; Franklin, J.F.; Swanson, F.J.; Sollins, P.; Lattin, L.D. *et al*. 1986. The ecology of coarse woody debris in temperate ecosystems. Recent Advances in Ecological Research. 15: 133-302.

Healy, W.M.; Brooks, R.T.; DeGraaf, R.M. 1989. Cavity trees in sawtimber-size oak stands in central Massachusetts. Northern Journal of Applied Forestry. 6: 61-65.

Kirby, K.J.; Reid, C.M.; Thomas, R.C.; Goldsmith, F.B. 1998. Preliminary estimates of fallen dead wood and standing dead trees in managed and unmanaged forests in Britain. Journal of Ecology. 35: 148-155.

Miles, P.D.; Brand, G.J.; Alerich, C.L.; Bednar, L.F.; Woudenberg, S.W. *et al*. 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.

Raphael, M.G.; White, M. 1984. Use of snags by cavity-nesting birds in the Sierra Nevada. Wildlife Monographs. 86: 1-66.

Schimmel, J.; Granstrom, A. 1997. Fuel succession and fire behavior in the Swedish boreal forest. Canadian Journal of Forest Research. 27: 1207-1216.

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

Stephens, S.L. 1998. Evaluation of the effects of silvicultural and fuels treatments on potential fire behavior in Sierra Nevada mixed-conifer forests. Forest Ecology and Management. 105: 21-35.

Sun, O.J.; Campbell, J.; Law, B.E.; Wolf, V. 2004. Dynamics of carbon stocks in soils and detritus across chronosequences of different forest types in the Pacific Northwest, USA. Global Change Biology. 10: 1470-1481.

Thomas, W.D.; Boza, C.A. 1984. The oak decline complex. Journal of Arboriculture. 10: 170-177.

U.S. Department of Agriculture (USDA), Forest Service. 2006. Forest inventory and analysis national core field guide. Version 3.1. St. Paul, MN: U.S. Department of Agriculture, Forest Service, Northern Research Station. Vol. 1

U.S. Environmental Protection Agency (EPA). 2004. Inventory of U.S. greenhouse gas emissions and sinks, 1990-2002. EPA 430-R-04-003. Washington, DC: U.S. Environmental Protection Agency.

Allometric Equations for Predicting Puerto Rican Dry Forest Biomass and Volume

Thomas Brandeis[1], Matthew Delaney[2], Larry Royer[3], and Bernard Parresol[4]

Abstract.—We used forest inventory data, intensive tree measurement, destructive sampling in the field, and subsequent laboratory analyses to develop regression equations that estimate tree biomass, merchantable volume, and total volume for upland forests in Puerto Rican subtropical dry forest. Most parsimonious and additive biomass equations for mixed, dry forest species were fitted to data from 30 destructively sampled trees, and most parsimonious volume equations were fitted for mixed species, *Bucida buceras*, and *Bursera simaruba*. Before undertaking this study, we were unable to confidently estimate biomass and volume for these forests despite having current, detailed forest inventory data.

Introduction

The subtropical dry forest life zone on the Caribbean islands that constitute the Commonwealth of Puerto Rico is found in areas with rainfall between 600 to 1,100 mm per year at elevations less than 300 m (Ewel and Whitmore 1973). This forested life zone covers 15 percent of Puerto Rico, mainly along the south coast and over most of the outlying islands of Culebra and Vieques (Ewel and Whitmore 1973). Native tree species typically found in the upland deciduous dry forest in Puerto Rico and much of the Caribbean include *Bursera simaruba* (L.) Sarg., *Bucida buceras* L., *Gymnanthes lucida* Sw., *Exostema caribaeum* (Jacq.) J.A. Schultes, *Guaiacum officinale* L., *Guaiacum sanctum* L., *Pisonia albida* (Heimerl) Britt. ex Standl., *Pictetia aculeata* (Vahl) Urban, *Acacia macracantha* Humb. and Bonpl., *Capparis spp.,* and *Coccoloba spp.*, among many others (Ewel and Whitmore 1973, Little and Wadsworth 1989, Murphy *et al.* 1995).

Besides natural variability, disturbance has dominated subtropical dry forest development since European colonization of the islands; on mainland Puerto Rico, only 4 percent of the original dry forest remains (Murphy *et al.* 1995). Human uses of dry forest areas in Puerto Rico have included sugar cane production, livestock grazing, irrigated agricultural crops and fruit trees, urbanization, industrialization, and live-fire military exercises in parts of Vieques and Culebra. Naturalized species often dominate highly disturbed Caribbean subtropical dry forests and include *Prosopis juliflora* (Sw.) DC., *Parkinsonia aculeata* L., *Tamarindus indica* L., *Acacia farnesiana* (L.) Willd., *Melicoccus bijugatus* Jacq., and *Leucaena leucocephala* (Lam.) DeWit. (depending on authority, this *L. leucocephala* is considered either native or naturalized) (Ewel and Whitmore 1973, Little and Wadsworth 1989).

Carbon accounting, forest health monitoring, and sustainable management of these forests require an accurate assessment of the tree biomass and wood volume. Early inventories of Puerto Rico excluded dry forests even though they make up a substantial portion of the island's forests because they were not considered to have the productive capacity to support commercial wood production (Birdsey and Weaver 1982, Franco *et al.* 1997). Current inventories include all forest types regardless of their productive capacities, so resource reports will include estimates of live tree aboveground biomass (AGB) and wood volume and for forest types where these resources were not previously considered.

[1] Research Forester, U.S. Department of Agriculture (USDA) Forest Service, Southern Research Station, Forest Inventory and Analysis Unit, 4700 Old Kingston Pike, Knoxville, TN 37919. E-mail: tjbrandeis@fs.fed.us.

[2] Delaney Forestry Services, LLC, 1216 Calapooia Street, Albany, OR 97321. E-mail: matt@delaneyforestry.com.

[3] Forester, USDA Forest Service, Southern Research Station, Asheville, NC 28802. E-mail: lroyer@fs.fed.us.

[4] Mathematical Statistician, USDA Forest Service, Southern Research Station, Asheville, NC 28802. E-mail: bparresol@fs.fed.us.

No locally developed allometric equations for estimating AGB in Puerto Rico's dry forests have been available, so estimates were made using equations developed from international data sets, principally Brown (1997), whose equation uses diameter at breast height (d.b.h.) to estimate AGB, and Martínez-Yrízar *et al.* (1992), which uses basal area as the explanatory variable. The lack of predictive equations hinders accurately estimating subtropical dry forest wood volume as well. In the Puerto Rican forest inventories conducted in 1980 and 1990 (Birdsey and Weaver 1982, Franco *et al.* 1997), field crews took multiple diameter and height measurements along the bole of each tree so that merchantable stem volume inside bark could be calculated by applying a geometric formula to different bole sections.

Despite having detailed and current forest inventory data, we have been unable to accurately estimate live tree AGB and merchantable stem volume for subtropical dry forest life zone in the forests of Puerto Rico. In this study, we outline the procedures taken to develop regression equations that use measurements of d.b.h. and H_T (total height) in the best parsimonious and additive models to estimate subtropical dry forest leaf, branch, bole, and total AGB. Volume equations were developed to estimate merchantable (stem volume to a 10-cm upper stem diameter) and total (stem volume in the entire stem) volume, inside bark, from measurements of d.b.h. and H_T.

Methods

Description of Study Area

We measured and destructively sampled trees in subtropical dry forest near the city of Ponce, PR (latitude 17°58'37.73" by longitude 66° 40'18.23") at a site about to be cleared for road construction. The site has an average rainfall of about 650 mm per year, according to the Southeast Regional Climate Center. A field crew visited the study area and installed a forest inventory plot nearby before destructive sampling. The study area was categorized as mature secondary forest and based on a single inventory plot installed at the site (see Brandeis [2003] for details on forest inventory sampling). A total of 7.1 m^2 per hectare of basal area on 2,749 stems per hectare of trees greater than or equal to 2.5 cm in d.b.h. was present. Tree species found on the inventory plot included *Bursera simaruba*, *Bucida buceras*, *Gymnanthes lucida*, *Bourreria succulenta* Jacq., *Krugiodendron ferreum* (Vahl) Urban, *Thouinia striata* Radlk, *Zanthoxylum caribaeum* Lam., *Reynosia uncinata* Urban, *Coccoloba microstachya* Willd., and *Thrinax morrisii* H. Wendl..

Sample Tree Selection, Harvesting, and Measurement

Preference was given to sampling trees that would be classified as growing stock according to the U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis guidelines. (Guidelines and measurement descriptions appear in the *FIA Field Data Collection Manual, Version 1.62, Supplement C for Puerto Rico and the Virgin Islands*, which can be downloaded at http://www.srs.fs.usda.gov/fia/manual/2.0_P2Manual.htm). To ensure sampling of healthy trees with more characteristic growth forms, we avoided trees with abnormal form, and trees classified as rough or rotten cull were sampled only when growing-stock trees were not present.

After each tree was felled, the field crew took detailed measurements following the methodology described in Cost (1978) used for developing volume equations in the southern United States. The measurements included diameter at the base or stump, d.b.h., diameters every 60 to 150 cm along the length of the stem depending on bole form, branch diameter and lengths, and total tree height. Bark thickness was sampled each time bole diameter was measured so that inside and outside bark merchantable volume could be estimated. The merchantable bole was defined as the main stem from a 30-cm-tall stump to a 10-cm upper stem diameter.

After the felled measurements were taken, the tree was separated into components and each component was weighed. The first component was the tree's crown, which consisted of leaves, small branches, and sometimes seeds and fruits. The second component was large branches (with a diameter greater than or equal to 2.5 cm). The third component was the main bole.

Subsamples were collected to determine the fresh-weight/dry-weight ratio for leaves and branches. The samples were then oven dried and a fresh-weight/dry-weight ratio was determined. For the large branches, small branches, and leaves, average fresh-weight/dry-weight ratios were applied to the total fresh-weight biomass of each of those components. Wood

disc subsamples were cut from the main bole; one disc was collected from the top, one disc was collected from the middle, and a third was collected at the base of the bole. After all three discs were collected, detailed measurements of disc volume were taken, which consisted of two measurements of diameter and two measurements of thickness. After the measurements were taken, each of the discs was placed in a cloth bag and labeled. Oven-dry weight was determined by a commercial laboratory. To determine oven-dry weight of the bole, the density results from each of the three discs were applied to the fresh-weight values. After all of the oven-dry weights were determined, the total biomass of the tree was calculated by summing all components.

Biomass Modeling

We took two approaches to modeling these data. In the first, we sought the best parsimonious models for the leaf, woody, and total tree biomass. In the second, we invoked the principle of additivity, meaning the predicted leaf and woody biomass sum to give exactly the total biomass predicted from the total biomass equation. This process is accomplished by imposing across equation constraints and fitting the resultant system of equations with nonlinear, seemingly unrelated regressions (NSUR).

Only two independent variables were available for modeling: diameter *(D)* and height *(H)*. This situation leads to expressions having the functional form Y = f (D, H). We selected the best biomass component equations based on scatterplots of the data and running all possible regressions on combinations of the variables.

For the mixed, dry forest species data ($N = 26$), we settled on the following equations:

$$\begin{aligned} \hat{w}_{\text{leaf}} &= b_1 D^{b_2} \\ \hat{w}_{\text{woody}} &= b_1 (D^2 H)^{b_2} \\ \hat{w}_{\text{total}} &= b_1 (D^2 H)^{b_2} \end{aligned} \tag{1}$$

where
$\hat{w}$ = biomass in oven-dry kilograms.

Scatterplots of the residuals revealed significant heteroscedasticity (as expected). One may want to assume equation errors that are multiplicative to derive log-linear models. The logarithmic transformation tends to stabilize heteroscedastic variance (if σ_ε is proportional to $E[y]$; Neter *et al.* [1985]) and is an alternative to deriving weights for each equation. Thus, for the dry forest data, we have the following biomass equations:

$$\begin{aligned} \ln \hat{w}_{\text{leaf}} &= b_1' + b_2 \ln D \\ \ln \hat{w}_{\text{woody}} &= b_1' + b_2 \ln (D^2 H) \\ \ln \hat{w}_{\text{total}} &= b_1' + b_2 \ln (D^2 H) \end{aligned} \tag{2}$$

where
ln = the natural logarithm,
$\hat{w}$ = biomass in oven-dry kilograms,

and $b_1' = \ln b_1$.

Parresol (1999, 2001) has discussed the problem of forcing additivity on a set of tree biomass functions. Because of the additivity restriction, the inherent model for $\hat{w}_{\text{total}}$ cannot be linearized. Thus, NSUR must be used as opposed to linear, seemingly unrelated regressions. The resulting system of additive equations for the mixed, dry forest species data includes following:

$$\begin{aligned} \ln \hat{w}_{\text{leaf}} &= \ln b_{11} + b_{12} \ln D \\ \ln \hat{w}_{\text{woody}} &= \ln b_{21} + b_{22} \ln (D^2 H) \\ \ln \hat{w}_{\text{total}} &= \ln \left[b_{11} D^{b_{12}} + b_{21} (D^2 H)^{b_{22}} \right] \end{aligned} \tag{3}$$

This system was fitted with PROC MODEL in SAS software (SAS Institute 1993).

Volume Modeling

Using the detailed measurements taken on each tree, volume estimates (in cubic meters) were directly calculated by applying the formula for the volume of a conic frustum to bole sections and summing these section volumes for total and merchantable volumes. Total volume is defined here as the inside bark portion of the tree's stem between a 30-cm-tall stump and the tip of the tree's stem without a minimum upper diameter merchantability limit. Merchantable volume is defined here as the inside bark portion of the tree's stem between a 30-cm-tall stump and a 10-cm upper stem diameter (outside bark). Both total and

merchantable volumes exclude wood volume in branches and only refer to main bole volume.

To estimate inside bark merchantable stem volume, the diameters for all sections were converted from outside bark to inside bark. Diameters inside bark at stump, breast, saw-log top, and pole-top heights were then calculated with the following formula for hardwoods:

$$BR = \frac{D_{BH} - D_{BT}}{D_{BH}} \quad (4)$$

where
BR = bark ratio,
D_{BH} = d.b.h. outside bark,
and D_{BT} = double bark thickness.

Section heights and inside and outside bark diameters were used in the formula for a conic frustum to calculate the wood volume of individual sections:

$$V_{SEC} = \{H_{SEC} \cdot [D_{IB1}^{\ 2} + (D_{IB1} \cdot D_{IB2}) + D_{IB2}^{\ 2}] \cdot 0.00007854\}/3 \quad (5)$$

where
V_{SEC} = section volume in m^3,
H_{SEC} = section height in meters,
D_{IB1} = diameter in cm inside bark at one end of section,
D_{IB2} = diameter in cm inside bark at other end of section,
and 0.00007854 is a constant derived from the expression

$$D_i^{\ 2} \cdot [\pi/(4 \cdot 10{,}000)] \quad (6)$$

where
D_i = section diameter in centimeters (Husch *et al.* 1993).

Results

Figure 1 shows total AGB (in oven-dry kilograms) by stem d.b.h. for each species sampled. The coefficients and fit statistics for the most parsimonious, natural logarithm-transformed models for mixed, dry forest species are given in table 1. The coefficients and fit statistics for additive models fitted to the mixed, dry forest species are presented in tables 2 and 3. The number of trees sampled, equation coefficients, mean square error (MSE) and r-squared statistic for estimating total and merchantable volume in cubic meters for mixed, dry forest species, *Bucida buceras*, and *Bursera simaruba* for models using tree d.b.h. and total tree height are given in table 4.

Figure 1.—*Total aboveground biomass (in oven-dry kilograms) by stem diameter at breast height (d.b.h.) for each species sampled.*

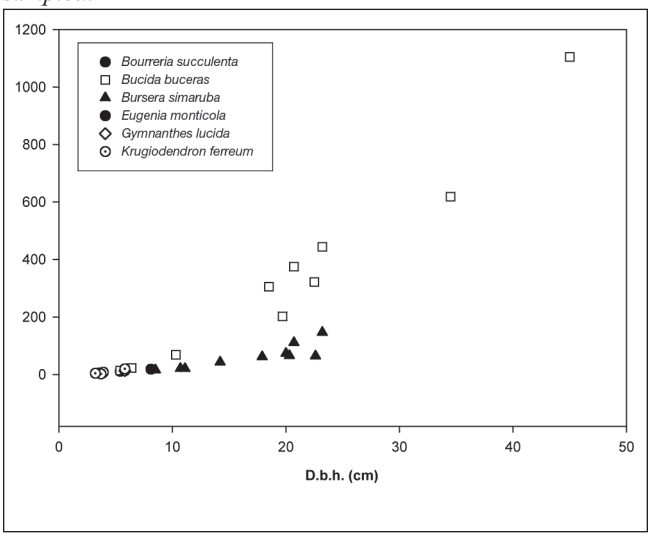


Table 1.—*Number of trees sampled, equation coefficients, mean square error (MSE), and r-squared statistic for the most parsimonious leaf, woody, and total aboveground biomass models for mixed, dry forest species.*

Component	Trees	a	b	MSE	r^2
	Number				
Leaf	26	– 1.75242	1.71833	0.57199	0.75980
Woody	26	– 2.87503	0.92900	0.27738	0.92500
Total	26	– 1.94371	0.84134	0.25252	0.91750

Table 2.—*Additive equation for aboveground biomass models for mixed, dry forest species.*

b_{11}	b_{12}	b_{21}	b_{22}
0.307631	1.540044	0.072847	0.899279

Table 3.—*Number of trees sampled, mean square error (MSE), and r-squared statistic for additive leaf, woody, and total aboveground biomass models for mixed, dry forest species.*

Component	Trees	MSE	r^2
	Number		
Leaf	26	0.5938	0.7403
Woody	26	0.2729	0.9232
Total	26	0.2488	0.9153

Table 4.—*Number of trees sampled, equation coefficients, mean square error (MSE), and r-squared statistic for estimating total and merchantable volume in cubic meters for mixed, dry forest species,* Bucida buceras*, and* Bursera simaruba *for models using tree diameter at breast height (in centimeters) and tree height (in meters).*

Species	Trees	Inside bark				Outside bark			
		a	b	MSE	r^2	a	b	MSE	r^2
	Number								
Total volume[a]									
Mixed species	17	– 0.0173	0.000047	0.0011	0.9906	– 0.0114	0.000046	0.0010	0.9916
Bucida buceras	9	0.0001	0.000047	0.0010	0.9952	0.0002	0.000046	0.0010	0.9950
Bursera simaruba	8	0.0370	0.000028	0.0003	0.8536	0.0354	0.000030	0.0004	0.8312
Merchantable volume[b]									
Mixed species	17	– 0.0593	0.000046	0.0020	0.9826	– 0.0576	0.000045	0.0019	0.9828
Bucida buceras	9	– 0.0697	0.000046	0.0024	0.9888	– 0.0688	0.000046	0.0023	0.9890
Bursera simaruba	8	0.2890	0.000025	0.0009	0.5553	0.0285	0.000023	0.0009	0.5620

[a] *Inside bark portion of the tree's stem between a 30-cm-tall stump and the tip of the tree's stem without a minimum upper diameter merchantability limit.*
[b] *Inside bark portion of the tree's stem between a 30-cm-tall stump and a 10-cm upper stem diameter.*

Conclusions

Although our ability to accurately estimate merchantable stem volume and live tree AGB for upland forests in the subtropical dry forest life zone in Puerto Rico and the U.S. Virgin Islands has been improved, much work remains to be done. This sample is small and needs to be expanded to include more species and more of the larger (d.b.h. greater than 30 cm) trees. This work should be seen as part of an ongoing process, and the equations presented here will be refined as additional sampling adds to the data available.

Acknowledgments

We thank James Bentley, Vincent Few, Tony Johnson, Humfredo Marcano, and Luis Ortíz of the Forest Service, Southern Research Station, Forest Inventory and Analysis program; Ross Hammons of the Tombigbee National Forest; Eileen Helmer and Ivan Vicéns of the Forest Service International Institute of Tropical Forestry; and Esther Rojas of the Puerto Rican Conservation Foundation.

Literature Cited

Birdsey, R.A.; Weaver, P.L. 1982. The forest resources of Puerto Rico. Res. Bull. SO-85. New Orleans, LA: U.S Department of Agriculture, Forest Service, Southern Forest Experiment Station. 59 p.

Brandeis, T.J. 2003. Puerto Rico's forest inventory: adapting the Forest Inventory and Analysis program to a Caribbean island. Journal of Forestry. 101: 8-13.

Brown, S. 1997. Estimating biomass and biomass change of tropical forests: a primer. FAO For. Pap. 134. Rome, Italy: Food and Agriculture Organization of the United Nations. 55 p.

Cost, N.D. 1978. Multiresource inventories—a technique for measuring volumes in standing trees. Res. Pap. SE-196. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 18 p.

Ewel, J.J.; Whitmore, J.L. 1973. The ecological life zones of Puerto Rico and the U.S, Virgin Islands. Res. Pap. ITF-18. Rio Piedras, PR: U.S. Department of Agriculture, Forest Service, Institute of Tropical Forestry. 72 p.

Franco, P.A.; Weaver, P.L.; Eggen-McIntosh, S. 1997. Forest resources of Puerto Rico, 1990. Res. Bull. SRS-22. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 45 p.

Husch, B.; Miller, C.I.; Beers, T.W. 1993. Forest Mensuration. 3rd ed. Malabar, FL: Krieger Publishing. 402 p.

Little, E.L.; Wadsworth, F.H. 1989. Common trees of Puerto Rico and the Virgin Islands. 2nd ed. Agricultural Handbook No. 249. Washington, DC: U.S. Department of Agriculture, Forest Service. 556 p.

Martínez-Yrízar, A.; Sarukhán, J.; Pérez-Jiménez, A.; Rincón, E.; Maass, M.; Solís-Magallanes, A.; Cervantes, L. 1992. Above-ground phytomass of a tropical deciduous forest on the coast of Jalisco, México. Journal of Tropical Ecology. 8: 87-96.

Murphy, P.G.; Lugo, A.E.; Murphy, A.J.; Nepstad, D.C. 1995. The dry forests of Puerto Rico's south coast. In: Lugo, A.E.; Lowe, C., eds. Tropical forests: management and ecology. New York: Springer-Verlag: 178-209.

Neter, J.; Wasserman, W.; Kutner, M.H. 1985. Applied linear statistical models. 2nd ed. Homewood, IL: Irwin Professional Publishing: 137-138.

Parresol, B.R. 1999. Assessing tree and stand biomass: a review with examples and critical comparisons. Forest Science. 45: 573-593.

Parresol, B.R. 2001. Additivity of nonlinear biomass equations. Canadian Journal of Forest Research. 31: 865-878.

SAS Institute. 1993. SAS/ETS user's guide. Version 6. 2nd ed. Cary, NC: SAS Institute.

Forest Inventory Predictions From Individual Tree Crowns: Regression Modeling Within a Sample Framework

James W. Flewelling[1]

Abstract.—Remotely sensed data can be used to make digital maps showing individual tree crowns (ITC) for entire forests. Attributes of the ITCs may include area, shape, height, and color. The crown map is sampled in a way that provides an unbiased linkage between ITCs and identifiable trees measured on the ground. Methods of avoiding edge bias are given. In an example from a forest of young southern pine, the forest is delineated into several thousand stands. Forty stands are sampled, each with two 0.12 acre plots. The resultant estimator of a volume surrogate, tree basal area times height summed over all trees, is unbiased and has a 90-percent confidence interval of ± 4.1 percent. The root mean square errors for basal area and the volume surrogate at the stand level are estimated at 9.7 percent and 12.8 percent, respectively. That precision in basal area for individual stands is approximately the same as would have been achieved by ground sampling with ten 0.12 acre plots in each stand, making no use of the remotely sensed data.

Introduction

Methods to obtain and interpret high spatial resolution digital imagery are evolving rapidly. Forest inventory systems are increasingly making use of such imagery with the goals of improved precision, lower field costs, and faster completion of large inventories. A logical step towards greater precision in inventory is to identify and describe the visible individual tree crowns (ITCs). Gougeon and Leckie (2003) provided an overview of the methodologies. Until recently most crown segmentation was based on digital color and infrared (CIR) photography. CIR photography is now complemented by "fused" data from airborne Light Detection and Ranging (LIDAR), making it practical to obtain and analyze high resolution CIR photography data and LIDAR data for entire forests. Gougeon and Leckie (2003) opined that the use of ITC techniques in forest inventory would evolve slowly, starting with a few niche applications. Næsset *et al.* (2004) reported that LIDAR is starting to be used in large-area forest inventories, although applications involving the estimation of individual tree properties are still in research mode. Næsset and Nelson (2009) presented the status of an ongoing operational-scale project in Norway.

Sampling Within a Stand

The overall sample design presented here uses stands as the primary sample units. Data from stands selected for sampling are used to develop relationships that will be applied in many stands; that strategy is the primary focus of this article. The within-stand sampling methods for multistand relationships are the same as would be proposed for sampling within individual stands, one stand at a time. That single-stand case is addressed first.

Sample Frame

The sample frame for a stand is the map of ITCs, together with the remotely sensed characteristics assigned to each ITC. The map is assumed to have accurate scale and stand boundaries. Each ITC is required to have a specific spatial extent and a specifically identified center. ITCs may not overlap one another, but may overlap stand boundaries. Only the ITCs whose centers are within the stand boundary are included in the sample frame.

Other attributes associated with the ITCs may include any statistics derived from remote sensing. If CIR photography data are available, means or other summary statistics for the color intensities of the pixels associated with each ITC are used. If LIDAR data are available, they are summarized to provide one

[1] Consulting Biometrician, 9320 40th Avenue NE, Seattle, WA 98115. E-mail: jwflew-wmen@yahoo.com.

or more height statistics; for example, the height above ground of the highest LIDAR return is an obvious statistic to consider. Additionally, data obtained from a "window" in the vicinity of each subject ITC may be available; for example, a statistic similar to top height could be calculated from the LIDAR heights of nearby ITCs.

Sample Selection and Field Work

A fixed number of sample points (*n*) are sought for a stand. Coordinates are randomly selected from a two-dimensional region that fully encompasses the mapped boundaries of the stand. If a random coordinate falls within the stand, it is accepted as a sample point. In addition to the *n* desired points, several additional random points within the stand are kept in reserve. The coordinates of the randomly selected points become the preliminary estimates of the locations of the sample points.

Field work consists of traveling to the preselected coordinates and installing fixed-area field plots around each point. The plots are located on the ground with portable Global Positioning System (GPS) equipment. The field crew travels to a point near the desired location. At that point, the GPS equipment is allowed to stabilize to produce a better estimate of the current location. Using that improved estimate, a bearing and distance to the preselected coordinate is determined; the field crew makes the indicated traverse and then monuments a plot center. Subsequent analysis will determine whether or not the monumented plot center is actually in the target stand; if it is not, one of the reserve coordinate pairs is used to establish an alternative plot center.

The field plot is a fixed-area circular plot. Every tree within the plot, whose diameter at breast height (d.b.h.) (D) exceeds a certain threshold (D_{min}), is recorded for diameter, species or species group, and distance and bearing from the monumented plot center. A subset of the trees is measured for height.

Plot Registration

The field plot is not assumed to have been located perfectly. The determination of the location of the field plot on the crown map is accomplished with a computer-assisted system that overlays the field-determined stem map on a representation of the remotely sensed data. Initially this procedure is performed in the field. If it appears likely that the center of the field plot is outside the stand boundary, a replacement plot within the stand is required. The plot registration process can be refined as part of the overall analysis. After the registration is completed, the location of the field plot's center on the ITC map becomes the accepted location for the sample. The vector change in location of each sample point is due to the sum of errors in the map and in the GPS.

Crown and Tree Matching

A procedure is required whereby a mapping between ITCs and field trees is developed. An ITC may be matched to no trees or to any number of trees. A tree may be matched to one ITC or may be unmatched. A requirement of the matching process is that it identifies all of the trees associated with ITCs whose centers are within the boundaries of a crown analysis plot. That plot is a circular plot, centered on coordinates determined in the plot registration process. The size of the plot is less than that of the field plot. A secondary requirement is that all the unmatched trees within a ground analysis plot be identified. The ground analysis plot is a circular plot whose center is the same as the field plot and whose size is smaller than that of the field plot. In the example inventory described later in the article, the field plot size is 0.12 acres and the crown analysis plot and the ground analysis plot both are 0.08 acres.

The mechanism of crown and tree matching is almost entirely automatic. The boundaries of each ITC are extended by up to 1 m, but the boundaries are not allowed to overlap. This process is applied to all crowns within a processing area whose extent must be greater than that of the field plot. Trees whose locations fall within the extended boundaries for an ITC are tentatively associated with that ITC. Of those trees, the one or two with the largest diameters are accepted as being matched, and the remainder is considered to be unmatched. Some subjectivity in the matching is allowed for leaning trees.

Sampling Theory

Every tree within the stand is considered to belong to one of two subpopulations. The first, referred to as the associated trees, consists of all the trees that would be matched to an ITC if the ITC were sampled. The second subpopulation, referred to as the unassociated trees, consists of those trees that would be

unmatched even if all the ITCs in their vicinity were sampled. This imposes a constraint upon the matching process: the decision on matching for an ITC or for a field tree must not be affected by the location of the plot center relative to the ITC or field tree. This constraint is addressed in the discussion section. Separate estimators are developed for the two subpopulations.

Models are developed to predict numbers and sizes of trees associated with each ITC. Before dealing with that complexity, however, the requisite theory to calculate sample weights that allow the data to fairly represent the population of trees within the stand is presented. Although the weights are to be used within regression analysis of ITC data, a simpler application may offer better motivation. Consider how weights would be calculated if the stand's basal area due to associated trees were to be estimated as the product of a sample-determined ratio and the summed areas of the ITCs within a stand. A suitable estimator for that purpose is described. Separately, I address how the observed, unmatched tree data can be used to make an unbiased estimate of the stand's basal area due to the unassociated trees.

Three weight calculations all deal with avoiding biases related to sampling near stand boundaries. Relative weights are assigned to the multiple sample points within a stand. Also weights are associated with individual ITCs within the crown analysis plot, and weights are associated with individual unmatched trees within the ground analysis plot. Ideally, the sample plot weights would be the same for all plots within the stand, which would be the situation if the maps and the GPS equipment were perfect; each sample plot would have been centered at the coordinates of the originally chosen random location. Instead the established plot locations have errors. The distribution of errors is observed in the plot registration process. That distribution can be modeled, and the model can be used to determine the probability density function (*pdf*) for sample point locations within the stand. Since the complete stand geometry is available, the determination of the *pdf* could be a numeric process. Alternatively, for large stands with smooth boundaries, the relative value of the *pdf* can be approximated as a function of the shortest distance to the edge of the stand (d):

$$pdf(d) \propto 1 - .5 \times \exp(k \times d)$$

where k is a constant related to the variance of errors in plot location.

Weight calculations are required for individual ITCs within a crown analysis plot and for individual unmatched trees within a ground analysis plot. The weight computation depends on the location of the ITCs or the location of the unmatched trees; the same computation procedure applies to both. The most efficient, unbiased weighting scheme is the tree concentric method (Schreuder *et al*. 1993). This scheme considers an object-centered circular plot. Each object's weight is the inverse of the proportion of the plot's area within the stand. The plot size is that of the crown analysis plot or the ground analysis plot, depending on whether ITCs or trees are being addressed. The determination of plot area within the stand is a routine computation for GIS software.

A sample-based estimate of the unassociated trees in the stand is simply a weighed list of the unmatched trees in ground analysis plots. For a single plot, each observed tree represents the number of trees per acre calculated as the product of the tree's weight and the reciprocal of the size of the ground analysis plot. The overall estimate of the distribution of unmatched trees is the weighted average of the estimates derived from the individual plots, using the plot weights specified above. The resultant estimated distribution is an unbiased estimate of the true distribution.

For a simple ratio-of-means estimate of the basal area for ITC-associated trees, the computations are as follows. The sample ratio of basal area to crown area is the weighted sum of the ITC-associated basal areas divided by the weighted sum of the ITCs' areas. The weights here are the product of the plot weights and the weights associated with the individual ITCs. The estimator of stand basal area is the product of the sample ratio and the sum of the areas of all the ITCs in the stand. The estimator is asymptotically unbiased. This estimator is presented for the sole purpose of demonstrating that an ITC sampling approach can be used within the context of a familiar sample estimator, the ratio of means.

Sampling Within Multiple Stands

The situation to be addressed is for a large forest holding with many stands in one or more strata. The intent is that a single set of regression relationships between ITC characteristics and the associated trees should apply to all the stands in a stratum. No expectation exists that the yield statistics of the stands in a stratum are similar. All strata are to be analyzed separately from one another; hence, the discussion need only address a single stratum. That single stratum is assumed to have a large number of stands. All stands are to have remotely sensed imagery, and ITC's are to be delineated for the entire areas of all the stands. Only a small fraction of the stands are selected for sampling. The within-stand sampling protocol is the same as was presented for single-stand sampling.

Two sets of sample stands are used: a development set and a calibration set. The development set of sample stands may be chosen by any means. The calibration set is a random selection from all the stands in the stratum. For the latter, sampling with replacement with probability proportional to size is assumed.

Regression equations are developed to predict the number, species, and sizes of trees associated with each ITC. The form of the equations is selected by viewing the data from the crown analysis plots in the development set of sample stands. Two types of regressions are needed. Logistic regressions predict the number and species of the matched trees for each ITC, and conditional regression equations predict the diameter and height of each tree. After the equation forms are chosen, the calibration and development data sets are combined to fit the coefficients.

The prediction equations receive a final modification based on the calibration data only. The modification to tree counts is a simple multiplier. For the d.b.h. equation, the modification is similar to a multiplier on basal area per tree, with an offset such that trees with diameter D_{min} are unmodified. Another modifier equation predicts altered values of tree height. The coefficients of the modifier equations are set in this order: trees, diameters, and heights. The values of the coefficients are set such that the weighted sum of predictions for the sample ITCs exactly equals the weighted sums of statistics from the trees matched to the same set of ITCs. The statistics being matched are trees per acre, basal area per acre, and product of basal area and Lorey height. That product is computed as the sum over trees of basal area times tree height and expressed on a per-acre basis.

The tree predictions sum to stand predictions, which in turn sum to the prediction of a stratum total. The estimator of a stratum total is similar to a ratio of means estimator where the independent variable is the unmodified regression prediction and the dependent variable is the true value. In the present case, however, the calibration sample influences the regression predictions; hence, the usual properties of the ratio estimator cannot be assured. The generalized regression (GREG) estimator (Särndal *et al.* 1992) is similar to that proposed here in that both adjust the estimator of the total on the basis of the weighted errors in the sample. The GREG estimator does so with an additive term while the approach here uses a multiplicative term. The ratio estimator is asymptotically unbiased, and the GREG estimator is approximately asymptotically design unbiased. The estimator proposed here is also expected to be approximately asymptotically design-unbiased. Asymptotic unbiasedness could be assured by limiting the use of the calibration data to the final modification step; doing so is not recommended because of the diminution of the sample size available for fitting the regression models.

The unassociated trees are dealt with separately. One approach would be to simply calculate an average stand table of unassociated trees using the weighted data from the calibration set. Alternatively, characteristics of the distribution of unassociated trees can be related to the predictions for associated trees. A complex model is warranted and possible only if a fair number of unassociated trees are in the sample data, and if the unassociated component is of silvicultural significance.

Error statistics are sought at the per acre level not the tree level. The variances of the estimators of strata-level means for trees per acre, basal area per acre, and basal area times Lorey height can be estimated from the sample results by stand; the individual plot data are needed only for computing stand-level statistics. As a first approximation, the predicted statistics for the stand can be treated as having come from a ratio estimator where the data are the stand-level means for the plot data in the calibration sample, and the means of the corresponding predic-

tions before the final calibration. The variance of the ratio may be approximated with standard formula for a ratio estimator, making no finite population correction. Alternatively, the bootstrap variance estimator may be used.

Example Inventory

Population, Sample Design, and Data

The sampling and regression methodologies were applied in a commercial forest with two strata. The single stratum addressed here consists of young plantations of a southern pine species, with minor hardwood components in some stands. A lower limit on stand age was chosen so that most of the dominant trees would be expected to have diameters greater than 3 in. Some 2,456 stands met the age criterion and were included in the stratum. Color infrared photography was obtained with a Digital Mapping Camera with 8-bit pixel radiometric resolution, multiple charge-coupled device camera heads, and forward motion compensation. Pixel size was 0.6 m on a side. Airborne LIDAR data was from an Optech ALTM 3100 Sensor System (100,000 laser pulses per second) that had an average of five first-returns per square meter. The LIDAR data were pixelized, 0.5 m on a side, and smoothed.

ITCs were delineated with algorithms similar to those described by Hyyppä *et al.* (2001). Manual training processes were used to set parameters for the algorithms. Each delineated ITC consisted of a specific set of LIDAR-derived pixels. A center point was defined as the geometric mean of the pixels associated with an ITC. A color was obtained for each ITC as the average color over the ITC by overlaying the ITC shape pattern on the CIR data. The height of the ITC was calculated as the height of the highest pixel.

Management data were available for all the stands. The data that were used included stand age, year of most recent thinning, and stand area. The development sample consisted of data from three sample stands which were thinned and ten which were unthinned. All of the unthinned stands in the strata had been placed into a 3 × 3 × 3 orthogonal array based on age, height, and crown closure. Statistics for stand height and crown closure had been derived from the LIDAR data. One sample stand was drawn from each of the corner cells of the array, and two were drawn from the center cell. Sampling was proportional to stand area. The resultant sample broadly covered the array of stand conditions but was not a representative sample. The calibration sample consists of 25 stands selected from the stratum; the selections were independent, with replacement, and with probability proportional to stand area.

Each sample stand was sampled with two 0.12 acre circular plots. The plots were located using the registration protocol stated earlier, with the exception that plots, whose centers were determined to be outside the targeted stand, were excluded but not replaced. The lower diameter limit (D_{min}) was 2.95 in. All trees above that limit had their diameters, species, and coordinates recorded. Every fourth tree was measured for height; additional height measurements were taken in plots where the number of height trees per species group would otherwise have been low. The species groups were pine and hardwood. Height diameter curves were fit, usually by species group within a plot, to impute height for unmeasured trees. For stands with few hardwoods, height diameter data were sometimes shared between the plots in a stand.

Predictions and Error Analysis

The prediction process is summarized here and is not given in detail. CIR data from the matched ITCs from plots in the development data set, supplemented with other subjectively chosen trees, were used in a discriminant analysis to obtain a preliminary prediction of the probability that an ITC would have a pine as its largest matched tree. That preliminary probability estimate, together with other ITC data and stand variables, were used as inputs in logistic regressions to predict the probabilities that zero, one, or two trees were associated with the ITC and to predict the species of the associated trees. Hence, the discriminant analysis had, as its only purpose, the collapsing of the CIR data to a single probability value for each ITC. A series of least-squares regression equations predicted the d.b.h. and heights of trees, conditioned on species and whether the tree was the largest or second largest associated with the ITC. For each ITC, the predictions were a series of outcomes with associated probabilities; the probabilities sum to one. The final step in developing the equations was to bring in adjustment factors which forced the weighted mean predictions for the calibration

data to equal the weighted mean of the observations. This step was done separately by species. The first adjustment fixed trees per acre, the second adjustment fixed basal area per acre, and the third adjustment fixed the basal area height product. In application, the predicted probabilities of various outcomes for every ITC in the population were converted to expected values and summed to generate stand tables.

Error analyses, as briefly described earlier, were used to estimate the variance of strata means estimators. One approach was to make predictions for each sample stand as the average of the predictions for the two plots, use the corresponding observed data to construct the ratio of means estimator, and then calculate the variance of the ratio estimator. That approach used the simplifying assumption of a common ratio for both species. Another approach was the bootstrap; its implementation has separate factors by species. Both approaches yielded similar results. The 90 percent confidence interval for basal area per acre and basal area times height is ±4.1 percent and ±5.4 percent, respectively. Actual errors were not known for any stands. The magnitude of the stand level errors, however, could be inferred through the use of a mixed model with terms for plot error and for stand error. The mixed model implied that stand level results had a root mean square error of 9.7 percent for basal area and 12.8 percent for basal area times height. To put the basal area result into perspective, the same level of uncertainty at the stand level would have been achieved on average by a system that did not use crown data but instead estimated each stand's basal area as the mean basal area observed on ten 0.12 acre plots, randomly located within the stand.

Discussion

Much of this manuscript is focused on how maps of ITCs can be used for sampling in a way that allows for unbiased estimation of basal area and a few other statistics. Subject to a few caveats, unbiased estimation is possible in a stand which is being sampled. If the unbiasedness property were not to hold at the stand level, it would not hold at the strata level. Furthermore, without the unbiasedness property, there would be little hope of obtaining accurate error assessments. What has been demonstrated is that unbiased estimators are possible in spite of errors in ITC delineation, errors in matching, and errors in modeling. To the extent that the mentioned sources of errors can be reduced, the precision of the estimators can be improved.

The within-stand sampling techniques presented here are potentially better than most cruising techniques in terms of randomly locating plots and successfully avoiding edge bias. The avoidance of edge bias will become increasingly important as stands become more fragmented and as management systems seek to delineate more special features such as riparian zones. One substantive difference between sampling from a map and sampling on the ground is the determination of stand boundaries. The use of crown maps for this purpose may be ideal in situations where differences in the remotely sensed imagery prompted the decision to draw a stand boundary. An extreme example of this type of stand boundary is embodied in the Gougeon and Leckie (2003) system of drawing stand boundaries so as to separate regions of differing ITCs.

The method of weighting plots within a stand based on a numerically determined *pdf* requires some effort, but it is feasible. Implicit in that method is an assumption of no differential bias between map coordinates and GPS coordinates. If all the map coordinates were off by several meters in one direction, however, this would cause one edge of the stand to be underrepresented in the actual field sampling. A costly solution would be to expand the region in which random map points are located to include the stand plus a buffer with a width equal to the maximum anticipated coordinate bias. A less expensive solution would be to estimate the location bias before entering the stand and then compensate for the bias. Any accessible feature in or close to the stand that can be located on the ground and in the imagery could be used to estimate the location bias.

The principal weakness in the sampling system presented here relates to tree matching. A problem is that subjective alterations in matching are allowed under the presumption that leaning trees or minor errors in ITC positioning cause the automatic matching process to go astray. These subjective alterations, while infrequent, have the potential to introduce bias. Persson *et al.* (2002) discuss the matching problem in

some detail from the perspective of trying to identify the physically correct matches. In addition to considering relative size of potentially matched trees, they favor matches that have short distances between tree position and ITC position. Apart from any bias consideration, having trees lean out of position and become associated with the "wrong" ITCs leads to loss of precision. Gatziolis (2009) minimizes this problem by locating in the field the three dimensional location of the tree tops, thereby improving the physical accuracy of the tree matching, particularly for conifers. His methods also improve the plot registration process. Gatziolis suggests that submeter accuracies are attainable; results in the example inventory of southern pine support that conclusion. With sufficiently high LIDAR scanning density in a coniferous forest, spatial accuracies of 0.2 m or less would seem to be achievable. If lean were to be fully accounted for, however, the retention of unbiasedness might require that larger field plots be measured. In the example inventory, the distance between the outer edge of the analysis plot and the outer edge of the field plot was 2.28 m. With this short distance, it is likely that some crowns, whose centers were within the crown analysis plot, extended beyond the measurement plot; hence, the location of the sample point relative to an ITC could affect the matching of trees to an ITC. An analysis of distances between ITC centers and locations of the largest associated matched trees indicated that matches missed due to not having measured beyond the 0.12-acre field-plot boundary would occur for 1 out of 200 ITC's on the outer edge of the crown analysis plot and far less frequently for ITCs located closer to the center of the crown analysis plot. If matching were to be based on the coordinates for the tops of the field trees rather than the bole locations at breast height, either the field plot size would have to increase or bias due to missed matches would increase.

The overall sampling design of having separate development and calibration samples is unusual. The deliberate assignment of the development sample within a design matrix ensures that the full range of stand conditions is being sampled in a manner that is likely to be efficient for regression modeling. These data are used to select the forms of the regression models. Model-assisted survey sampling, and possibly all sample survey estimators which claim unbiasedness, assume that the form of the estimation equation is fixed before sampling. Since ITC regression models are in an early development stage, and would be expected to vary with forest type and sensing technology, they should not be presumed to be known before sampling. The procedures outlined here ensure that the model forms are known and fixed before the critical step of calibration. A motivation for using sampling with replacement in the calibration set is that this ensures that the stand-level observations are from identical and independent distributions. In turn, variance computations are greatly simplified, and the assumptions in standard bootstrap methodology are more easily met.

Acknowledgments

The example inventory described here was carried out by ImageTree Corporation of Morgantown, WV. Bob Pliszka coordinated the project and supervised the field work. Olavi Kelle developed and applied the image processing algorithms. The author served as a consultant to ImageTree Corporation in the design of the inventory and the analysis of the data.

Literature Cited

Gatziolis, D. 2009. Precise plot registration using field and dense LIDAR data. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C.; McWilliams, W.H., Proceedings, eighth annual forest inventory and analysis symposium: 243-250.

Gougeon, F.A.; Leckie, D.G. 2003. Forest information extraction from high spatial resolution images using an individual tree crown approach. Information Report BC-X-396. Victoria, British Columbia, Canada: Canadian Forest Service, Pacific Forestry Center. 29 p.

Hyyppä, J.; Kelle, O.; Lehikoinen, M.; Inkinen, K. 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. IEEE Transactions on Geoscience and Remote Sensing. 39: 969-975.

Næsset, E.; Gobakken, T.; Holmgren, J. *et al.* 2004. Laser scanning of forest resources: the Nordic experience. Scandinavian Journal of Forest Research. 19: 482-499.

Næsset, E.; Nelson, R.F. 2009. Sampling and mapping forest volume using airborne LIDARs. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C.; McWilliams, W.H., eds. Proceedings, eighth annual forest inventory and analysis symposium. Gen. Tech. Rep: 297-302.

Persson, Å.; Holmgren, J.; Söderman, U. 2002. Detecting and measuring individual trees using an airborne laser scanner. Photogrammetric Engineering and Remote Sensing. 68: 925-932.

Särndal, C.E.; Swensson, B.; Wretman, J. 1992. Model assisted survey sampling. New York: Springer. 694 p.

Schreuder, H.T.; Gregoire, T.G.; Wood, G.B. 1993. Sampling methods for multi-resource inventories. New York: John Wiley. 446 p.

External Validation of a Forest Inventory and Analysis Volume Equation and Comparisons With Estimates From Multiple Stem-Profile Models

Christopher M. Oswalt[1] and Adam M. Saunders[2]

Abstract.—Sound estimation procedures are desideratum for generating credible population estimates to evaluate the status and trends in resource conditions. As such, volume estimation is an integral component of the U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis (FIA) program's reporting. In effect, reliable volume estimation procedures are fundamental to FIA's mission. Currently, FIA in the Southern Research Station (SRS) uses linear regression models of the form $V = \alpha + \beta(dbh^2 Ht) + \varepsilon$ to derive volume estimates. Although copious data have been used to estimate region- and species-specific parameters, little attention has been paid to external validation of the developed models. We investigated the validity of the SRS FIA volume equation (SRSvol) for cherrybark oak *(Quercus pagoda)* through external validation procedures using a stem analysis data set of 49 separate stems. The aggregate volume of measured Smalian bolts for each stem were compared with volume estimates derived from SRSvol for gross cubic-foot volume (VOLCFGRS in FIA Database 2.0). In addition, estimates from four stem-profile models (MAXvol, CLRKvol, FARvol, and WALTvol) were compared to those of SRSvol. The SRSvol model provided the most accurate and precise estimates of cherrybark oak volume. In addition, the stem-profile model FARvol proved to be the superior stem-profile model. Although the performance of the SRSvol model can be viewed as positive and supportive of the current SRS FIA volume estimation procedures, this study reflects results from one species. As such, additional external validation of the SRS FIA volume equations across multiple species is warranted. In addition, the results of this study suggest that as FIA as a national program proceeds towards nationally consistent estimation procedures, stem-profile models such as the FARvol model form may provide reasonable estimates with the additional benefits of allowing for regional and species-specific flexibility and for the incorporation of changing utilization specifications.

Introduction

The Forest Inventory and Analysis (FIA) program of the Forest Service Southern Research Station (SRS) is currently charged with producing State-level estimates of numerous forest resource values, levels, and trends for 13 southern States. Sound estimation procedures are desideratum for generating credible population estimates to evaluate the status and trends in resource conditions. As such, volume estimation is an integral component of FIA reporting. In effect, reliable volume estimation procedures are fundamental to FIA's mission. Currently, the SRS FIA uses volume estimators of the conventional linear regression model form $V = \alpha + \beta(dbh^2 Ht) + \varepsilon$ (Spurr 1952) to derive volume estimates (Zarnoch *et al.* 2003). Although copious data have been used to estimate region and species-specific parameters, little attention has been paid to publishing external validation of the developed models. As a result, published volume estimates must be taken at face value by external and internal FIA users who lack the appropriate information necessary for assessing confidence in the estimates.

[1] Forest Resource Analyst, U.S. Department of Agriculture Forest Service, Southern Research Station, Forest Inventory and Analysis, 4700 Old Kingston Pike, Knoxville, TN 37919. E-mail: coswalt@fs.fed.us.

[2] Undergraduate Research Assistant, University of Missouri-Columbia, Department of Forestry and Department of Statistics, 146 Middlebush Hall, Columbia, MO 65211–6100.

Numerous model forms and methods are used to estimate standing tree volume. For example, the former North Central Research Station FIA program uses nonlinear models developed by Hahn and Hansen (1991) and linear regression models developed by Hahn (1984). Furthermore, the former Northeastern Research Station FIA program uses nonlinear methods (Scott 1981) to derive cubic-foot volume estimates of standing trees. The Rocky Mountain Research Station FIA program uses multiple published (Edminster *et al.* 1980, 1982; Hann and Bare 1978; Myers 1964; Myers and Edminster 1972) and unpublished equations, including both linear and nonlinear model forms.

Advances in analytical techniques have led to additional, more complex methods for estimating volume in standing trees. Stem-profile (or stem-taper) models have been of particular interest to many researchers. For example, Clark *et al.* (1991) developed stem-profile equations for 58 southern tree species, for possible use by the Forest Service. Although the authors appeared to be pleased with the accuracy of the results, no comparison with the linear regression equations used by the SRS FIA was included. Zarnoch *et al.* (2003) compared the SRS individual tree linear regression methodology to taper models similar to taper models produced by Clark *et al.* (1991), and found no appreciable difference between the two methods on both an individual tree and statewide scale. No external validation was included, however. Clark *et al.* (1991) did include a validation of the developed models through cross-validation; however, true external validation with an independent data set is always optimal (Neter *et al.* 1996).

Without external validation and comparisons with externally validated alternative models, it is difficult to completely defend model or equation choice or maintain confidence that the optimal methodology has been chosen for FIA reporting. Currently, to the authors' knowledge, no published external validation, using an independent data set, of the SRS linear regression equations exists. In addition, a direct comparison with alternative models, using the same validation data set, such as taper equations developed by Clark *et al.* (1991), Farrar and Murphy (1987), Max and Burkhart (1976), and Walters and Hann (1986), may suggest the relative accuracy of the currently used methodology and provide insight into potentially needed change in estimation procedures.

Here, we investigate the strength of the SRS FIA volume equations currently used, specifically for predicting gross volume of cherrybark oak (*Quercus pagoda Raf.*), through external validation procedures using a stem analysis data set of 49 separate stems. In addition, alternate volume estimators are evaluated with the same data set. The three objectives of this article are to (1) evaluate current SRS FIA volume estimators with external validation procedures; (2) evaluate a series of alternate volume estimators, including recently developed stem-profile equations; and (3) compare the performance of the SRS FIA estimators with that of each of the alternative models.

Methods

Volume Estimators

Volume estimators used for comparative purposes included the SRS FIA linear regression model (SRSvol) currently used in data processing in the South, Max and Burkhart's (1976) segmented-profile polynomial model (MAXvol), the form-class segmented-profile model of Clark *et al.* (1991) (CLRKvol), Farrar and Murphy's (1987) taper function model (FARvol), and the taper equations of Walters and Hann (1986) that include live crown ratio as a predictor variable (WALTvol).

Data

Separate independent data sets were used to parameterize the models (parameter data set) for comparisons and evaluate predictions (evaluation data set). The parameter data set was used to parameterize each model form when published coefficients for cherrybark oak were not available (the SRS FIA linear regression coefficients, although not published, were made available by Larry Royer of the Forest Service SRS FIA program). When published coefficients were available for both pole and sawtimber, coefficients were applied accordingly. When coefficients were estimated with the parameter data set, however, one set of coefficients was developed due to limited data.

Data for the parameter data set represent a species-specific (cherrybark oak) subset of a larger data set of detailed felled tree data from across the South, collected by field crews from the old Southeastern Forest Experiment Station (Larry Royer and Tony Johnson provided the larger data set from which the parameter data set was taken). The data were obtained from a total of 81 cherrybark oak trees and collected sensu Cost (1978) modified for felled trees, for the development of volume equations for southern tree species. Tree height varied from 34 to 108 ft and diameter at breast height (d.b.h.) varied from 4 to 38 in.

Data for the evaluation data set were obtained from 49 felled cherrybark oak trees evenly distributed among four separate stands. Three stands were located on the Natchez Trace State Forest in western Tennessee (latitude 35.83333 by longitude -88.25833) and one stand was located on privately owned land in Dixon Springs, TN (latitude 36.35804 by longitude -86.050209). D.b.h. varied from 3 to 20 in and height varied from 19 to 99 ft. Selected trees were single stemmed, free of sweep or crook, and visibly undamaged. After felling, radial disks were removed from each tree at the stump, defined as 15 cm above ground level d.b.h. and every 1.22 m thereafter for the reminder of the stem. Diameter inside bark (d.i.b.), diameter outside bark (d.o.b.), and bark width were collected for each disk. Stem volume inside bark and stem volume outside bark were calculated for each tree using Smalian's formula. To achieve d.b.h. agreement between the parameter and evaluation data sets, all trees in the evaluation data set with a d.b.h. below 4 in (the minimum d.b.h. of the parameter data set) were removed. The resultant data set contained a total of 422 observations on 31 trees.

Validation Procedures

After satisfactory coefficients were either found in the literature or estimated using the parameter data set, gross cubic-foot volume was estimated for trees in the evaluation data set with each of the proposed models. The specifications for calculating gross cubic-foot volume followed definitions of gross volume outlined by the FIA Statistics Band (2006). Volume was calculated from a 1-ft stump to a minimum 4-in top d.o.b. Estimated volume was then compared to the aggregate Smalian-bolt volume for each tree under the same specifications. Mean prediction error ($\bar{e}$), standard deviation of the prediction, mean absolute difference (MAD), mean squared prediction error (MSPR), and 10-percent confidence limits were calculated for each model.

The predictive capability of the chosen model was evaluated with the calculated MSPR (Neter *et al.* 1996):

$$\text{MSPR} = \frac{\sum_{i=1}^{n}\left(Yi - \hat{Yi}\right)^2}{n} \quad (1)$$

where

Yi is the value of the response variable in the ith validation case,

$\hat{Yi}$ is the predicted value for the ith validation case based on the development data set,

and n is the number of cases in the evaluation data set.

In addition, the relative error in prediction (RE%) and the prediction coefficient of determination (R_p^2) were calculated for each model:

$$R_p^2 = 1 - \frac{\left|\sum (Yi - \hat{Yi})^2\right|}{\sum (Yi - \bar{y})^2} \quad (2)$$

where

Yi is the value of the response variable in the ith validation case,

$\hat{Yi}$ is the predicted value for the ith validation case based on the development data set,

and $\bar{y}$ is the mean of n observations.

Finally, individual two-tailed t-tests were preformed to detect significant differences between the cumulative Smalian-bolt volume and each of the volume estimates.

Results and Discussion

Model Performance

Electing a "best" model purely from the interpretation of the calculated validation statistics and diagnostic tests appeared

straightforward. Most of the validation statistics supported the SRSvol model as the better performing estimator. In addition, the plurality of the diagnostic tests supported the SRSvol model. The lowest MAD, mean absolute percent prediction error (MAPPE), MSPR, RE%, $\bar{e}$ and standard deviation, and standard error of the $\bar{e}$ resulted from the application of the SRSvol model (table 1). The SRSvol model also produced the largest prediction coefficient of determination, or R_p^2 of 0.98, followed by the CLRKvol, FARvol, WALTvol, and MAXvol models (0.96, 0.92, 0.87, and 0.65, respectively). A simple nonweighted ranking of the calculated validation statistics and diagnostic tests used in this study indicated a separation between the SRSvol and the FARvol models from the remaining three models. The procedure ranked the SRSvol model as the best model, followed by the FARvol, CLRKvol, MAXvol, and WALTvol models.

Overall, the SRSvol and FARvol models appeared to produce the most accurate estimates with $\bar{e}$ of -1.21 and -1.62 and MAD of 1.42 and 1.71, respectively. In addition, the MAPPE was smallest for the SRSvol and FARvol models (table 1). All of the models had a tendency to overpredict, with the exception of the MAXvol model, which tended toward the underprediciton of individual stem volume. When comparing the predicted and measured volumes of the collective data set, the MAXvol model underpredicted total volume and the SRSvol, FARvol, WALTvol and CLRKvol models overpredicted it (fig. 1). The SRSvol, FARvol, and CLRKvol models appeared much more accurate than the WALTvol and MAXvol models.

Comparable to the SRSvol model's accuracy, the SRSvol model appeared to promulgate the greatest precision. Each of the estimates of model precision, MSPR, RE%, and R_p^2 indicated the SRSvol model as having greatest precision of the five tested models, followed by the FARvol model and then the CLRKvol model (table 1).

Although the validation statistics and the nonweighted ranking of the models suggested that the SRSvol and FARvol models were superior, the results from both the t-tests and f-tests concluded that the predicted volume was statistically different from the measured volume (table 1) for all models. However,

Table 1.—*External validation statistics for five volume estimation procedures for estimating cubic-foot volume in cherrybark oak.*

	Models				
	SRSvol	**MAXvol**	**CLRKvol**	**WALTvol**	**FARvol**
ē	– 1.21	2.26	– 1.76	– 3.49	– 1.62
Std Dev	1.34	1.61	2.38	2.98	1.53
Std Err	0.25	0.3	0.44	0.55	0.28
MSE	1.77	1.42	2.66	2.44	1.95
SSE	47.88	38.45	71.88	65.92	52.71
MPPE	– 13.19	32.42	– 3.66	– 19.62	– 13.54
MAPPE	15.09	25.92	17.99	23.25	15.89
MAD	1.42	2.29	2.07	3.59	1.71
MSPR	3.19	7.65	8.59	20.77	4.88
RE%	15.11	27.94	24.79	37.78	20.59
Rp^2	0.98	0.65	0.96	0.87	0.92
t-test	4.88	7.57	– 3.98	– 6.32	– 5.71
(P-value)	*< 0.0001*	*< 0.0001*	*0.0004*	*< 0.0001*	*< 0.0001*
f-test	1052.35	904.41	980.24	1250.84	1051.91
(P-value)	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*
X^2	4.18	5.4	4.89	4.07	4.16
(P-value)	*0.12*	*0.07*	*0.09*	*0.13*	*0.12*
Pearson	0.987	0.985	0.987	0.989	0.987
(P-value)	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*
AIC	18.54	12.18	30.32	27.81	21.33

AIC = Akaike's Information Criterion. ē = prediction error. MAD = mean absolute difference. MAPPE = mean absolute percent prediction error. MPPE = mean percent prediction error. MSE = mean square error. MSPR = mean square error of prediction. Pearson = Pearson correlation coefficient. (P-value) = associated P-value of the previous test statistic. RE% = relative error in prediction. Rp^2 = prediction coefficient of determination. SSE = error sum of squares. Std Dev = standard deviation of the prediction error. Std Err = standard error of the prediction error. t-test = paired t-test. X^2 = chi-square test of independence.

the X^2 test of independence was not significant for any of the five models. Furthermore, multiple thresholds for acceptable growth and yield model performance have been published. Freese (1960) suggested a plus or minus 7.16-percent error and Pilsbury *et al.* (1995) estimated that a plus or minus 12-percent error was an appropriate line on which to base model choice. Moreover, Huang *et al.* (2003) posited that a plus or minus 10- to 20-percent error implied a "reasonable" model. No model in this study preformed adequately, according to the standard set by Freese (1960) or Pilsbury *et al.* (1995), and only the SRSvol, FARvol, and CLRKvol models preformed reasonably, according to the Huang (2003) standard.

Composite validation statistics, such as previously discussed, do not always provide all necessary information with respect to model performance across the population of interest. For example, many of the composite statistics will not adequately convey how a volume model specifically operates as stem diameter changes. Graphical evaluation of model predictions provides a very effective method to gain this form of insight (Huang *et al.* 2003).

Fluctuations in prediction error for each of the five models varied as d.b.h. increased (fig. 2). The MAXvol model consistently underpredicted and the SRSvol and FARvol models appeared to consistently overpredict. The remaining models underpredicted for small diameter stems and overpredicted for large diameter stems (WALTvol and CLRKvol). The impact of changing diameter on prediction error, as a percentage of total stem volume, also varied drastically among the five tested models (fig. 3). The estimates produced by the CLRKvol, MAXvol, and WALTvol models appeared more sensitive to changes in diameter with errors ranging from approximately -10 to 60 percent, -10 to -35 percent, and -40 to 25 percent of the measured volume for the MAXvol, WALTvol, and CLRKvol models, respectively. The SRSvol and FARvol models appeared much less sensitive to stem diameter changes (fig. 3). The use of models that minimize the variability in prediction error among diameter classes will decrease the overall error in the

Figure 1.—*Total predicted volume (in cubic feet) for all cases by five volume models and associated difference between predicted and measured volume for cherrybark oak (*Quercus pagoda Raf.*).*

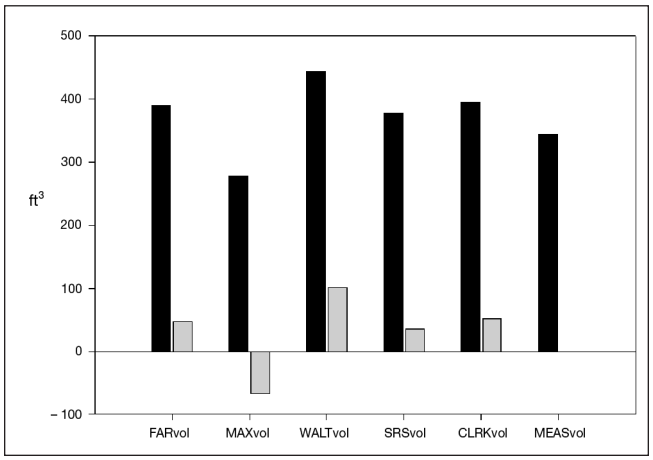


Figure 2.—*Regression of prediction error by diameter class for five volume estimation methods.*

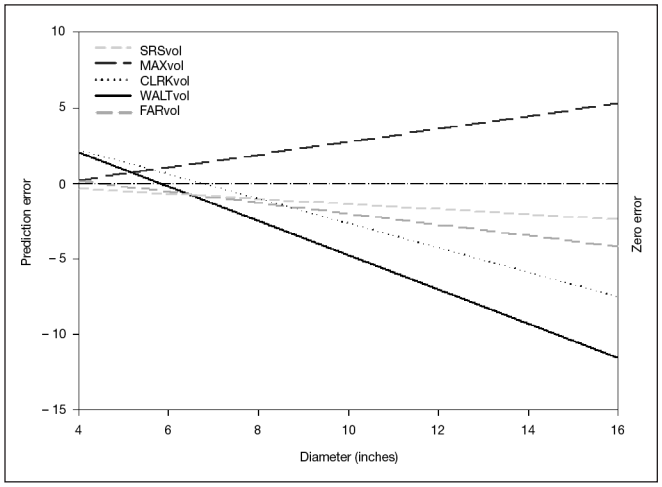
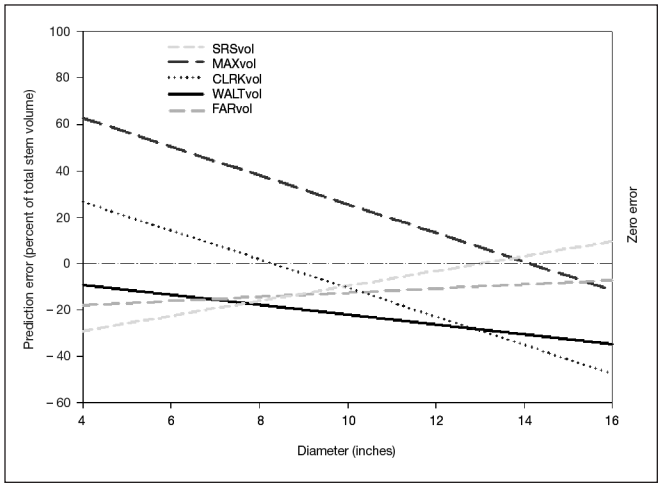


Figure 3.—*Regression of percent prediction error by diameter class for five volume estimation methods.*

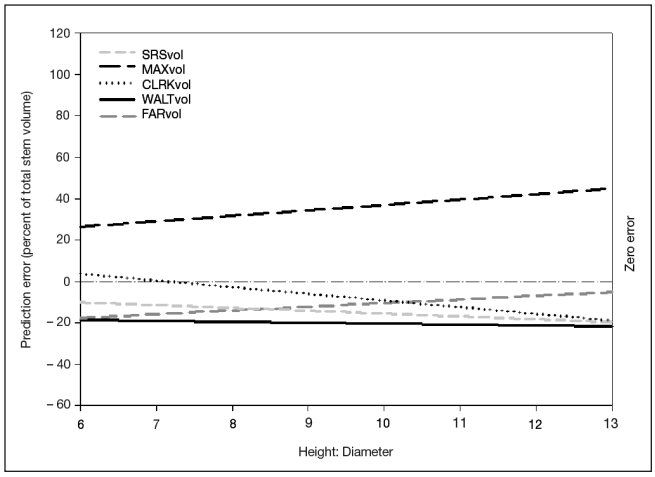


cumulative volume estimates. Even though both the SRSvol and FARvol models appeared to be less sensitive to changes in diameter, the FARvol model was superior to the SRSvol model in this respect and appeared to demonstrate increasing accuracy within the larger diameter classes of the evaluation data set (fig. 3).

Using height-diameter ratio as an analogue for stem form, the MAXvol and CLRKvol models were more influenced by changes in stem form than the WALTvol, SRSvol, and FARvol models (fig. 4). Differences in stem form are a result of stem developmental history and can vary widely in a large sample population, such as that collected by FIA. As such, additional error may propagate through volume estimates through the impact of variable stand histories. This error can be removed through the development of models that lessen, if not remove, such influence. The WALTvol model, although a consistently poor predictor, appeared resistant to the influence of changing stem form (figure 4), indicating the possibility of using stem-profile models to reduce volume estimate error.

Overall, although the FARvol model represented the "best" performing stem-profile model, the cherrybark oak volume equation currently used by the SRS FIA (SRSvol) consistently outperformed the other four models. SRSvol was the only model in which the 95-percent prediction interval completely contained the 1-to-1 line for the range of volumes represented by the evaluation data set (fig. 5). The performance of the SRS-vol model in this external validation study should give users of FIA data confidence in the volume estimates generated by the SRS FIA; however, the analysis in this study is restricted by the limited independent data available and that only one species is represented.

Figure 4.—*Regression of percent prediction error by height:diameter ratio for five volume estimation methods.*



Evaluation of Four Tested Stem-Profile Models

1. *WALTvol*

Description of Equation. This is a variation on the Max and Burkhart (1976) three-part polynomial equation that uses a percentage of crown base as the join point for the upper stem.

Parameterization. This equation is very long and complicated. Mistakes are easily made when entering the equation due to its complexity. The alpha parameter needs to be hand estimated before regression estimation and needs to be iterated and reregressed until hand-estimated and regression-estimated alpha agree. This process can be very time consuming or can appear to be impossible in some cases.

Pros. The upper join parameter (alpha) is flexible and incorporates crown ratio into the equation in a logical way. The equation produces very nice taper plots.

Cons. This is a very long and complicated model. Alpha is needed to set indicator variables before regression and needed within the model. Iterative steps need to be made until the hand-estimated and regression-estimated alpha values are equal.

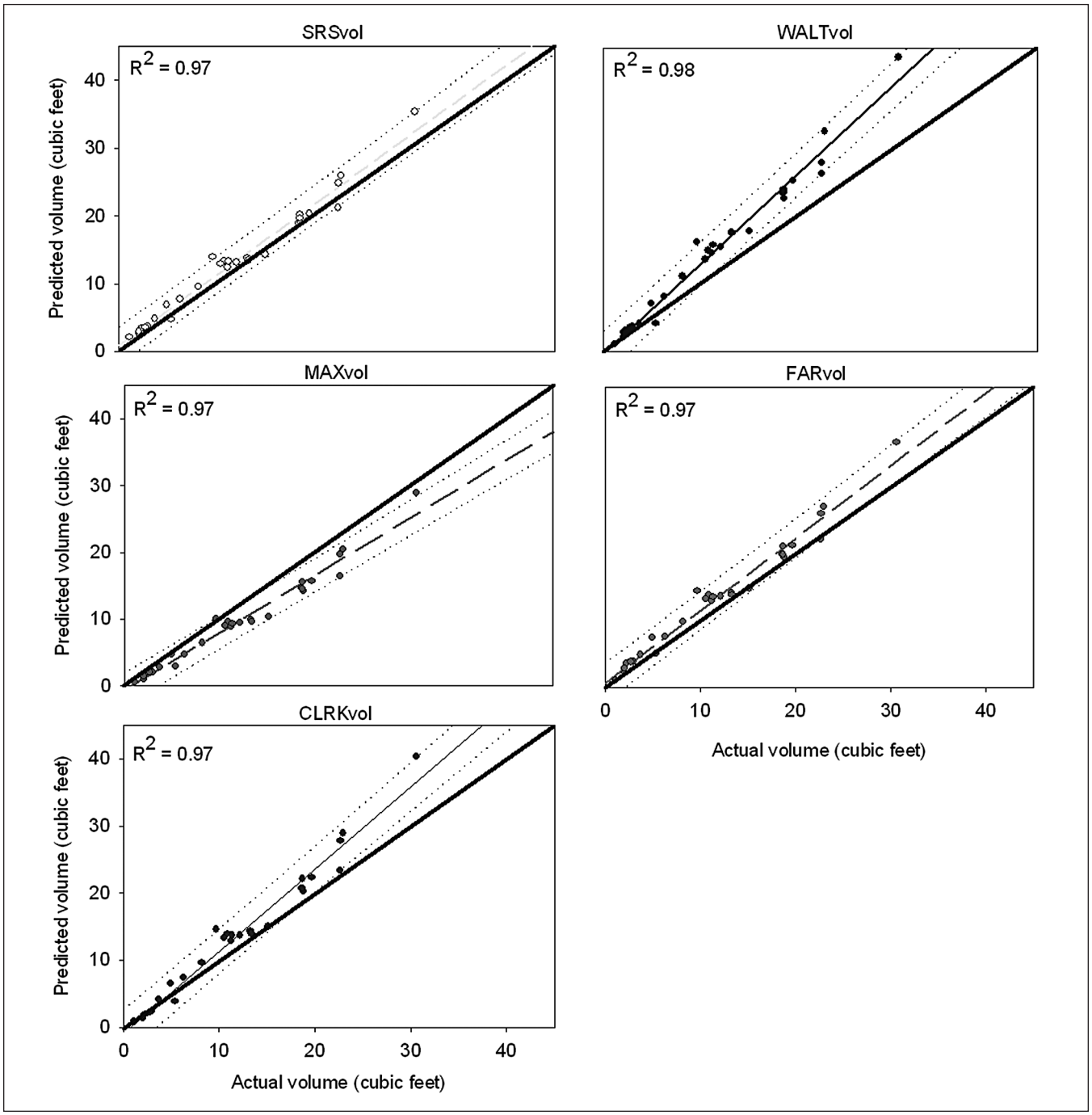
Results. Model performance was poorest of the five that we examined. The model produced the largest standard deviation (2.98) and the largest mean difference (-3.49). This model tends to overpredict volume more than any of the five tested models.

2. *MAXvol*

Description of Equation. This is a three-part equation using a quadratic-quadratic-quadratic form.

Parameterization. The equation is fairly simple and easy to code; as a result, mistakes are less likely. Like the WALTvol equation, an alpha parameter is present that is hand estimated to set an indicator variable before regression and is estimated in the regression. This hand-estimated alpha needs to be iterated until it agrees with the regression-estimated

Figure 5.—*Actual compared with predicted volume for five volume estimation methods with 95-percent prediction intervals and associated coefficient of determination for cherrybark oak (*Quercus pagoda Raf.*).*



value. Finding matching $alpha_{xx}$ can be time consuming or impossible.

Pros. The equation creates very nice realistic taper profile curves.

Cons. The indicator variable requires iteration, which is time consuming and complicated.

Results. This model predicted fourth best of the five examined models. It had the third smallest standard deviation (1.61) but the second largest mean difference (2.26). This model tends to predict volume too conservatively and produced the only overall conservative prediction.

3. ***CLRKvol***

Description of Equation. The CLRKvol model uses form-class measurement at 17.3 ft to aid in fitting the segmented-profile equation.

Parameterization. Iteration from preregression parameter estimations and outputted regression estimates can be difficult to stabilize.

Pros. Incorporating form class into the model helps to produce nice fits. Model flexibility allows for many forms of the model to be used in analysis.

Cons. The paper repeats parameter names in multiple parts of the paper, which creates some confusion. One must use existing parameters to estimate a form class if one has not been measured.

Results. The model predicted the third best of the five examined models. It had the fourth smallest standard deviation (2.38) and the third smallest mean difference (-1.76). This model tends to overpredict volume but predicts well.

4. ***FARvol***

Description of Equation. This is a two-part model, below and above breast height. Farrar and Murphy (1987) recommend separating the data into three crown ratio classes and fitting unique parameters for each class.

Parameterization. This is a simple equation that is easy to code and fit. The model predicts real diameters from real heights (not in relative space).

Pros. The model produces good estimates and is very easy to apply. In addition, it produces estimates in real space. No parameter is present that needs to be hand estimated before regression, iteration, and refit.

Cons. The equation originally applied to shortleaf pine. The current study applies the equation to hardwoods.

Results. The model predicted the second best of all the five. The standard deviation is the second smallest (1.53) and the mean difference was also the second smallest (-1.62). The small standard deviation reduces the t-test value, making this model the second closest to 0. This model tends to overpredict volume but predicted very well in the current study.

Conclusions

The five models examined in this study vary greatly in user friendliness, accuracy, and method. Each model had its strong and weak points; however, the SRSvol model and FARvol model clearly outperformed the others with the superior model being SRSvol. Although the simple form of the SRSvol model is clearly the easiest to apply, the FARvol model is uncomplicated relative to the other stem-profile models in this study and could be easily applied in large-scale estimation efforts such as that of FIA. As a result, the FARvol model form may be appropriate for further testing as interest increases and efforts proceed in developing nationally consistent FIA estimation procedures.

The FIA volume equation is the easiest to use because a taper profile equation step is not involved. This model uses simple linear regression to directly predict volume from d.b.h. and height. The four other models predict a taper profile equation as an intermediate step before volume is estimated. Volume is derived from the taper profile by inserting heights at every 10th ft into the equation and outputting diameter estimates accordingly. The 10th-ft diameter estimations are converted into “discs” and summed to yield total volume for each tree. The intermediate taper profile step is difficult to apply at first but goes very quickly after the appropriate algorithm and code have been written. The largest problem with three of the taper models examined in this study is an iterative process required to fit the model. These models require the user to estimate

a parameter related to a join point or an indicator variable that is also a parameter fit in the model. If the preregression parameter does not match the regression estimated parameter, then the preregression parameter must be changed slightly and the regression rerun. This procedure must be repeated until the preregression parameter is equal to the regression found parameter. In some cases, this procedure can take tremendous time or is impossible. This process was not discussed in any of the model citations and no advice was given on how to find fits efficiently. The FARvol equation does not have any iterative parameters, which makes it the easiest taper equation to use.

The five studied models each have unique assumptions and model forms. Indepth discussion of the model forms is beyond the scope of the article but would make for a very interesting comparison. The largest difference between the taper and FIA volume procedures is the taper profile intermediate steps. Taper profiles provide very good flexibility to make changes to volume prediction criteria, such as small-end diameter limits. Taper profiles have great potential to provide information on wood quality, wood value, and growth modeling. A need exists to develop models that provide the flexibility associated with taper profile equations with a stronger emphasis on usability and simplicity.

Although the evaluation data set used in this study is not truly representative of the population of cherrybark oaks sampled across the 13 southern States, the results do suggest that the currently used SRSvol model form is appropriate. Certain stem-profile models, such as the FARvol model, may be potentially applicable if the form of a stem-profile model is desired at the national level. As such, it appears that FIA may benefit from pursuing, through multispecies testing, alternative estimation procedures that have reliable performance across regions. Concomitantly, it may prove advantageous to the national FIA program to pursue nationally consistent volume models. As seen here, and in other studies cited in this work, stem-profile models could be a workable alternative that has the potential to improve FIA volume estimates while providing for region-specific and species-specific flexibility and temporally and geographically shifting utilization specifications. In addition, the performance of the SRSvol lends a certain degree of credibility to the SRS FIA estimation procedures.

Acknowledgments

The authors gratefully acknowledge Larry Royer and Tony Johnson for providing data, Wayne Clatterbuck and Dave Larsen for their reviews and numerous discussions that greatly improved this article, and Sonja Oswalt and other FIA personnel for additional reviews. We also thank the Forest Service Southern Research Station Center for Bottomland Hardwoods Research for funding this research and the Tennessee Division of Forestry and Mr. Thomas Beasley for granting access to their property and allowing destructive sampling of trees.

Literature Cited

Clark, A.; Souter, R.A.; Schlaegei, B.E. 1991. Stem profile equations for southern tree species. Res. Pap. SE-282. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 113 p.

Cost, N.D. 1978. Multiresource inventories—a technique for measuring volumes in standing trees. Res. Pap. SE-196. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 18 p.

Edminster, C.B.; Beeson, R.T.; Metcalf, G.B. 1980. Volume tables and point-sampling factors for ponderosa pine in the front range of Colorado. Res. Pap. RM-218. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 14 p.

Edminster, C.B.; Mowrer, H.T.; Hinds, T.E. 1982. Volume tables and point-sampling factors for aspen in Colorado. Res. Pap. RM-232. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 16 p.

Farrar, R.M.; Murphy, P.A. 1987. Taper functions for predicting product volumes in natural shortleaf pines. Res. Pap. SO-234. New Orleans, LA: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station. 9 p.

FIA Statistics Band. 2006. http://srsfia2.fs.fed.us/publicweb/statistics_band/stat_documents.htm.

Freese, F. 1960. Testing accuracy. Forest Science. 6: 139-145.

Hahn, J.T. 1984. Tree volume and biomass equations for the Lake States. Res. Pap. NC-250. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station. 10 p.

Hahn, J.T.; Hansen, M.H. 1991. Cubic and board foot volume models for the Central States. Northern Journal of Applied Forestry. 8: 47-57.

Hann, D.W.; Bare, B.B. 1978. Comprehensive tree volume equations for major species of New Mexico and Arizona: II. tables for unforked trees. Res. Pap. INT-210. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 43 p.

Huang, S.; Yang, Y.; Wang, Y. 2003. Critical look at procedures for validating growth and yield models. In: Amaro, A.; Reed, D.; Soares, P., eds. Modelling forest systems. Cambridge, MA: CABI Publishing: 271-293.

Max, T.A.; Burkhart, H.E. 1976. Segmented polynomial regression applied to taper equations. Forest Science. 22(3): 283-289.

Myers, C.A. 1964. Volume tables and point-sampling factors for lodgepole pine in Colorado and Wyoming. Res. Pap. RM-6. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 16 p.

Myers, C.A.; Edminster, C.B. 1972. Volume tables and point-sampling factors for Engelmann spruce in Colorado and Wyoming. Res. Pap. RM-95. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 23 p.

Neter, J.; Kutner, M.H.; Nachtsheim, C.J.; Wasserman, W. 1996. Applied linear regression models. Chicago, IL: Irwin Professional Publishing. 720 p.

Pilsbury, N.H.; McDonald, P.M.; Simon, V. 1995. Reliability of tanoak volume equations when applied to different areas. Western Journal of Applied Forestry. 10(2): 72-78.

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. 3 p.

Spurr, S.H. 1952. Forest inventory. New York: Ronald Press. 476 p.

Walters, D.K.; Hann, D.W. 1986. Taper equations for six conifer species in southwest Oregon. Res. Bull. 56. Corvallis, OR: Oregon State University, College of Forestry. 41 p.

Zarnoch, S.J.; Clark, A.; Souter, R.A. 2003. Comparison of past, present, and future volume estimation methods for Tennessee. Res. Note SRS-12. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 6 p.

Preliminary Results of Spatial Modeling of Selected Forest Health Variables in Georgia

Brock Stewart[1], Chris J. Cieszewski[2], and Eric L. Smith[3]

Abstract.—Variables relating to forest health monitoring, such as mortality, are difficult to predict and model. We present here the results of fitting various spatial regression models to these variables. We interpolate plot-level values compiled from the Forest Inventory and Analysis National Information Management System (FIA-NIMS) data that are related to forest health. These data included information concerning mortality, trees killed by various causes of death (indicated by the FIA-NIMS variable AGENTCD), and species richness.

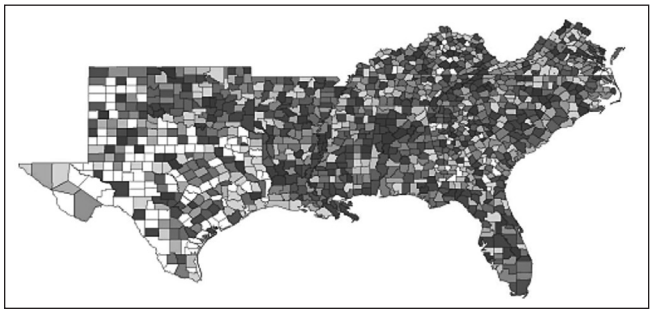
Introduction

In this study, we were interested in gathering information about forest health and mortality in the Southeast. The U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis (FIA) inventory data provide a good starting point for such analysis. The FIA data contain records of many variables, including information on stand structure with extensive coverage of large areas. We decided to use a means of displaying these estimates in a mapped context and/or testing for spatial trends that is more revealing and informative than a straightforward compilation of tabular estimates. The selection of appropriate resolution of mapping posed a challenge. While mapping State-level estimates is too course for our purposes, FIA data are intended to provide accurate information only at large scales, and county-level estimates can typically have large sampling errors. Kriging techniques have been used in abundance on FIA data.

The purpose of the current work was to investigate the fitting of various spatial models that would provide interpolation maps and significance testing. In this article, we focus predominantly on model fitting. That is, we attempt to find interpolation models with predictive value from phase 2 FIA data. Further research might include, e.g., incorporating Forest Health Monitoring (FHM) aerial surveys, which do not provide detailed information on stand structure but do provide large-scale trends of mortality and infection. Remote-sensing data such as this can provide information on large-area coverage of when and where trees are dying but cannot provide details of stand structure. On the other hand, plot-level data provide information on forest type, species, stand size, and density, etc., but are not suited well for detecting rare or sparse events. In the future, we would like to incorporate information at both scales.

In figure 1, we give an example of a FIA plot-level value mapped by county. Here, the value is the number of trees killed by insects. Our main interest was in detecting trends in mortality, and, if possible, mortality by specific causes such as insects or disease. We also examined species richness. We wish to test if plot-level quantities like these vary across the Southeast and if they vary by other covariates, such as forest type, species, age, size, etc. We also want to examine elevation, forest cover type, FIA unit, physiographic region, and ecoregion because these covariates can be determined at prediction points (i.e., where no FIA plots are present).

Figure 1.—*Relative county-level estimates of the number of trees killed by insects from newest Forest Inventory and Analysis plot-level data for each southeastern State as of August 2006.*



[1] Graduate Student, The University of Georgia, Warnell School of Forestry and Natural Resources, Athens, GA.

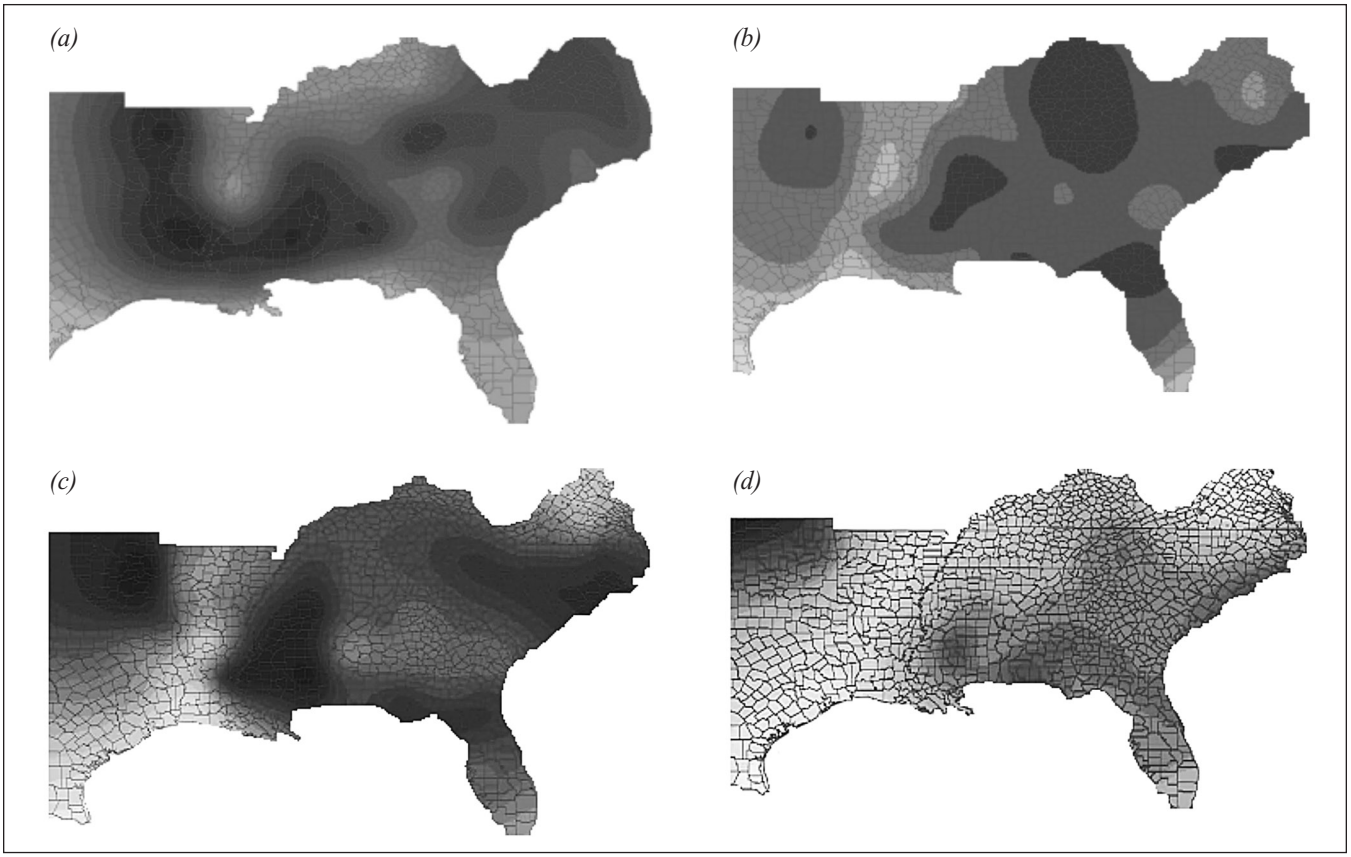
[2] Associate Professor The University of Georgia, Warnell School of Forestry and Natural Resources, Athens, GA.

[3] Quantitative Analysis Program Manager, U.S. Department of Agriculture, Forest Service, Forest Health Technology Enterprise Team, Fort Collins, CO. E-mail: elsmith@fs.fed.us.

Examples of maps produced from interpolating FIA data include those in figure 2. Although these maps were made after fitting semiparametric penalized spline models (SPSMs) (Ruppert *et al.* 2003, Shabenberger and Gotway 2004), model diagnostics for these models indicate inadequate fits. In the following text, we discuss the possible shortcomings of trying to model data like this across the Southeast. As a baseline case for spatial modeling from the FIA phase 2 data, we compare an ordinary kriging (OK) (Cressie 1993, Shabenberger and Gotway 2004) of the number of trees on forested conditions to the forest area map produced by Zhu and Evans (1994) in figure 3. We can see that the OK procedure was at least successful in detecting major, broad-level trends indicated in Zhu and Evans' (1994) map; e.g., lack of forests along the Mississippi River and southern FL. We can also notice a strong relationship between prediction error and which State the plot-level data is in. This relationship can be attributed to different sampling intensities between States and even within a State.

Several difficulties occur when trying to model data such as plot-level mortality and sources of this mortality. First of all, these data are rare and overdispersed. For example, over the whole Southeast, roughly 95 percent of the plots with a forested condition had zero trees killed by insects, with a sample variance-to-mean ratio of approximately 12.1. This is an extreme case of zero-inflated data. In situations in which the data exhibit inflation on one value (e.g., zero), transformations merely move this inflation to another value. Forest attributes, in general, tend to have large local variability. Also, in gathering the most recent FIA phase 2 data for each State in the Southeast, we had to use data for each State from different measurement years. Aside from any temporal differences that might actually exist in the data (e.g., one year with high mortality rates over the whole region), data collection methods may vary due to changes in FIA sampling procedures. In fact, data collection procedures may vary from State to State anyway, especially for more "obscure" plot values such as sources of mortality.

Figure 2.—*Trend maps from semiparametric penalized spline models predicting (a) forest area, (b) the number of trees on plots killed by insects,(c) the number of trees on plots killed by disease, and (d) the number of trees on plots killed by fire.*



To illustrate this, we provide figure 4. Here, we can see clear trends across States, which are due either to varying numbers of plots in the data, data collection procedures, or both. Even the newest data for SC did not contain information on sources of mortality (i.e., AGENTCD). Due to these difficulties, we resorted to fitting models from plots in only AL and GA. If the data were more consistent across States, we might hope to handle the situation of no data for one State by extrapolating into the State and keeping track of prediction error.

Figure 3.—*Zhu and Evans' (1994) (a) forest coverage map, (b) ordinary kriging prediction on the total number of trees on forested conditions in plots, and (c) prediction error.*

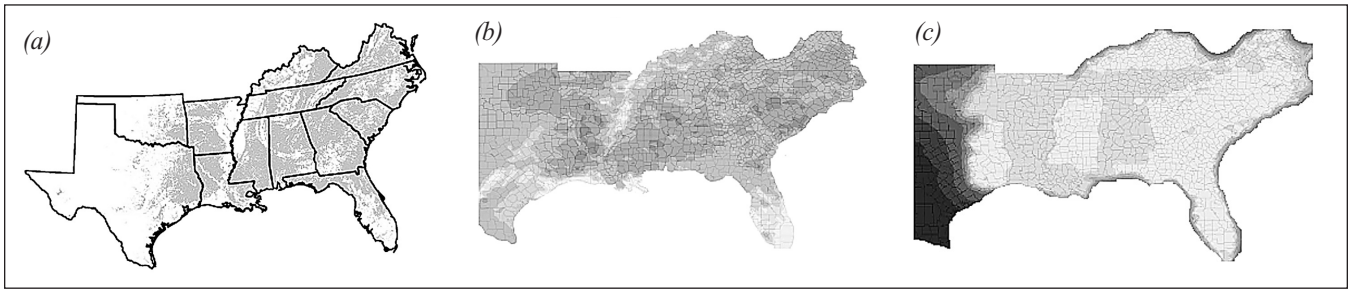
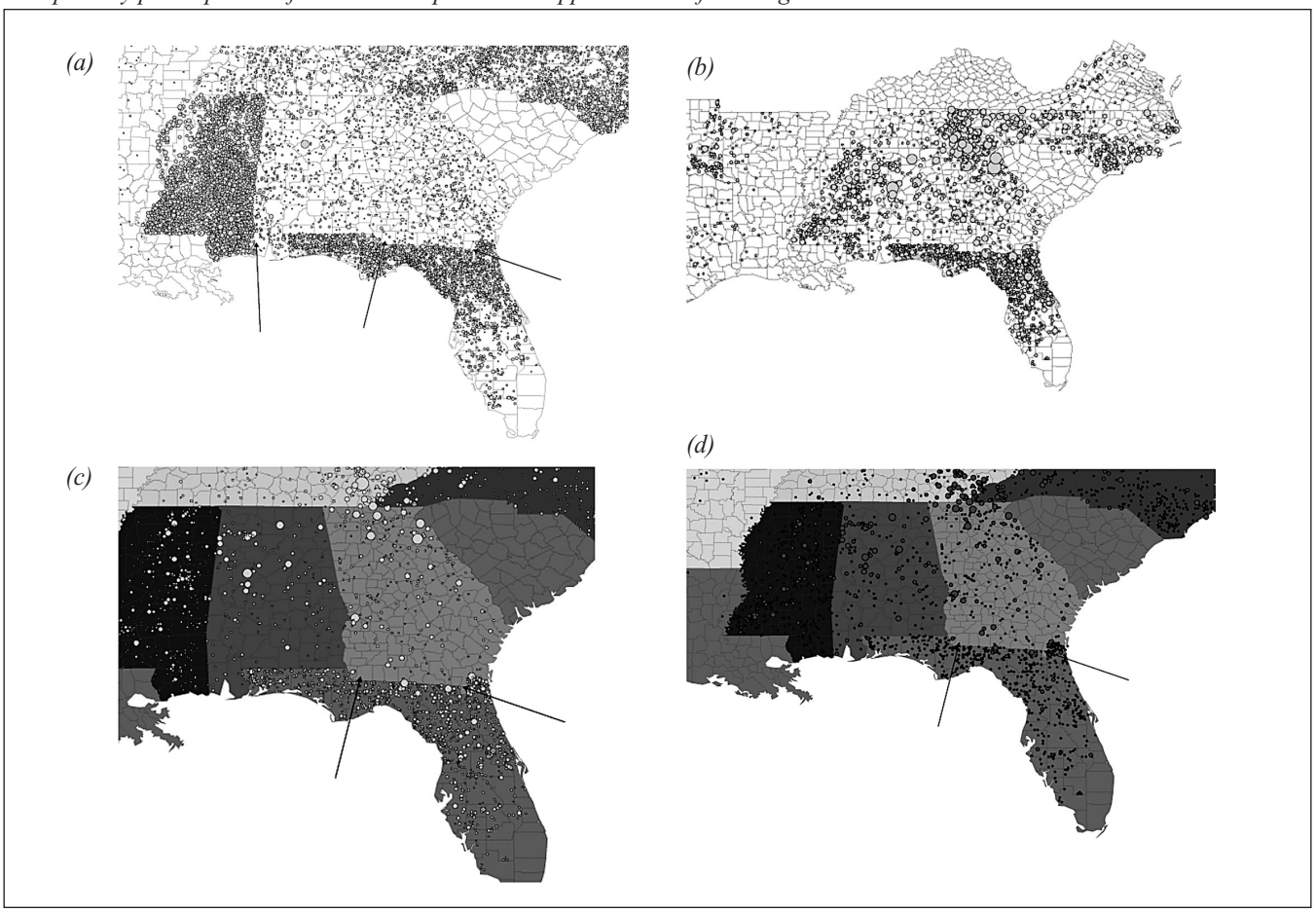


Figure 4.—*Forest Inventory and Analysis phase 2 plots with (a) symbol proportional to number of trees on plot killed by disease, (b) the number of trees on plots killed by insects, (c) a closer view of the same, and (d) the number of trees on plots killed by insects multiplied by plot expansion factor. Arrows point to an apparent trend following State boundaries.*



Materials and Methods

We fit several models to the number of trees on plots in AL and GA killed by insects. Instead of kriging, or geostatistical models, we try spatial regression models (Cressie 1993, Shabenberger and Gotway 2004). We chose these models for several reasons. A large number of FIA plots are present, even for one State, and standard Kriging techniques require inverting an n-by-n matrix, where n equals the number of plots. This may be prohibitively expensive computationally, and many authors have chosen ad hoc methods to deal with kriging on large data (e.g., over a moving window). One interesting method that we do not explore here is the fixed rank kriging method of Cressie (2006) and Johannesson and Cressie (2004). This method still uses all of the data, but necessitates inverting only smaller matrices. Also, the data we work with here exhibit extreme departures from the standard Gaussian distribution. Instead of fitting a trend to the data first and then kriging the residuals, we try direct models to the data. Although universal kriging and nonlinear kriging methods may work for these data, the problem of large data size still exists, and we do not explore them here. Hence, we chose to explore models that are low rank in that k knots are selected in the domain where k<<n. The resulting models are then spline functions connected at the knots. Theory and software is also readily available to extend these models to the generalized situation of various distributions assumed on the response.

Zero-inflated data are not rare in real-world data, and much effort has recently been applied to finding techniques for fitting models to them. Zero-inflated data, as the name implies, are data exhibiting a large number of zeros. These types of data can be found in many disciplines and often are the result of rare count data. Lambert (1992) provided techniques for modeling data with a zero-inflated Poisson (ZIP) model. Incorporating zero-inflated likelihoods in spatially explicit models is described in Agarwal *et al.* (2002), Barry and Welsh (2002), Fahrmeir and Echavarria (in press), Gschlobl and Czado (2006), Rathbun and Fei (2006), Rigby and Stasinopoulos (2005) and in general (not spatial) in Hall (2000), Lambert (1992), and Li *et al.* (1999).

We used data from GA cycle 08 and AL cycle 07. We counted the raw (i.e., not expanded) number of trees killed by insects in forested conditions (LANDCLCD=1), indicated by AGENTCD=10. Plot species richness was determined by counting the number of unique species codes (SPCD) for trees in forested conditions on each plot. We used the R Project for Statistical Computing (http://www.r-project.org) packages SemiPar and Generalized Additive Models for Location, Scale and Shape to fit the models here. The SPSMs we use through SemiPar have a smoothing parameter fit via restricted maximum likelihood. Knots were automatically selected in SemiPar via a space-filling algorithm (Ruppert *et al.* 2003) with the default number of 50 knots.

Results and Discussion

A histogram of the number of trees on plots killed by insects for only plots having at least one tree killed by insects is given in figure 5. In figures 6 and 7, we give histograms of residuals,

Figure 5.—*Histograms of (a) number and (b) proportion of trees killed by insects on forested conditions of plots in Alabama and Georgia, only for plots with at least one tree killed by insects in a forested condition.*

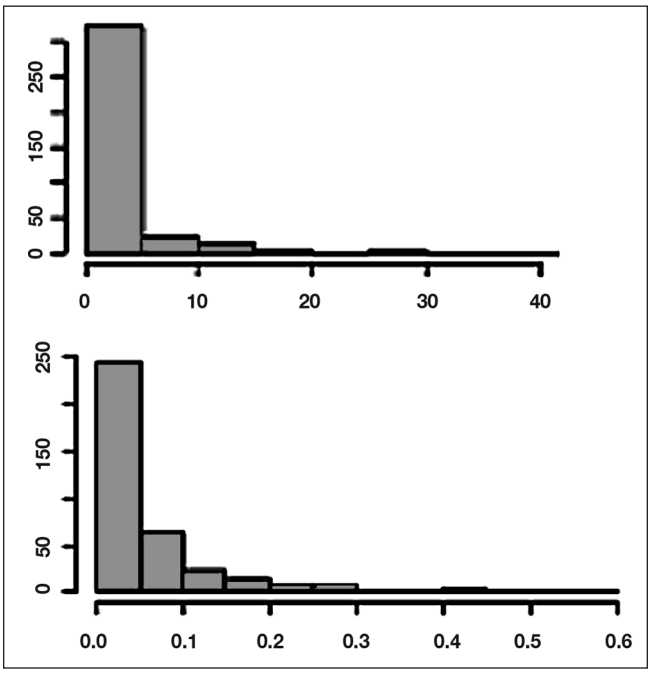


Figure 6.—*Semiparametric penalized spline model Poisson number of trees killed by insects: (a) residuals histogram,(b) Q-Q plot, and (c) histogram of fitted values.*

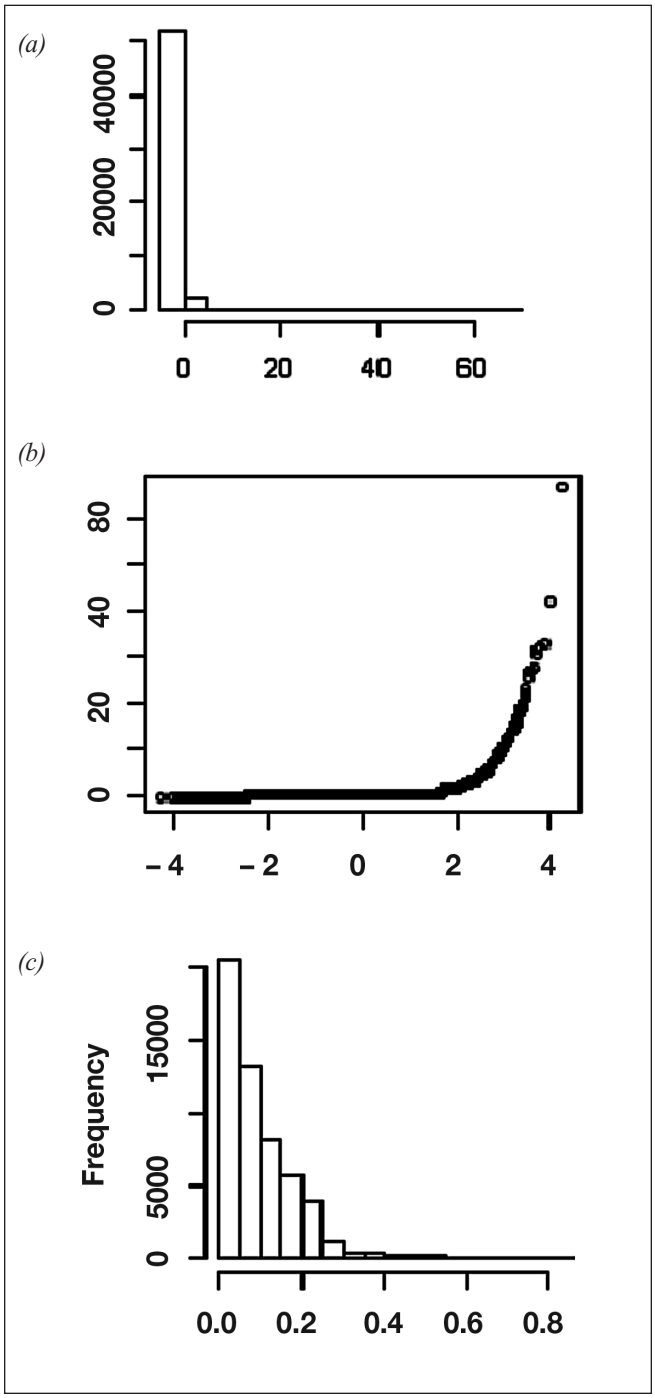
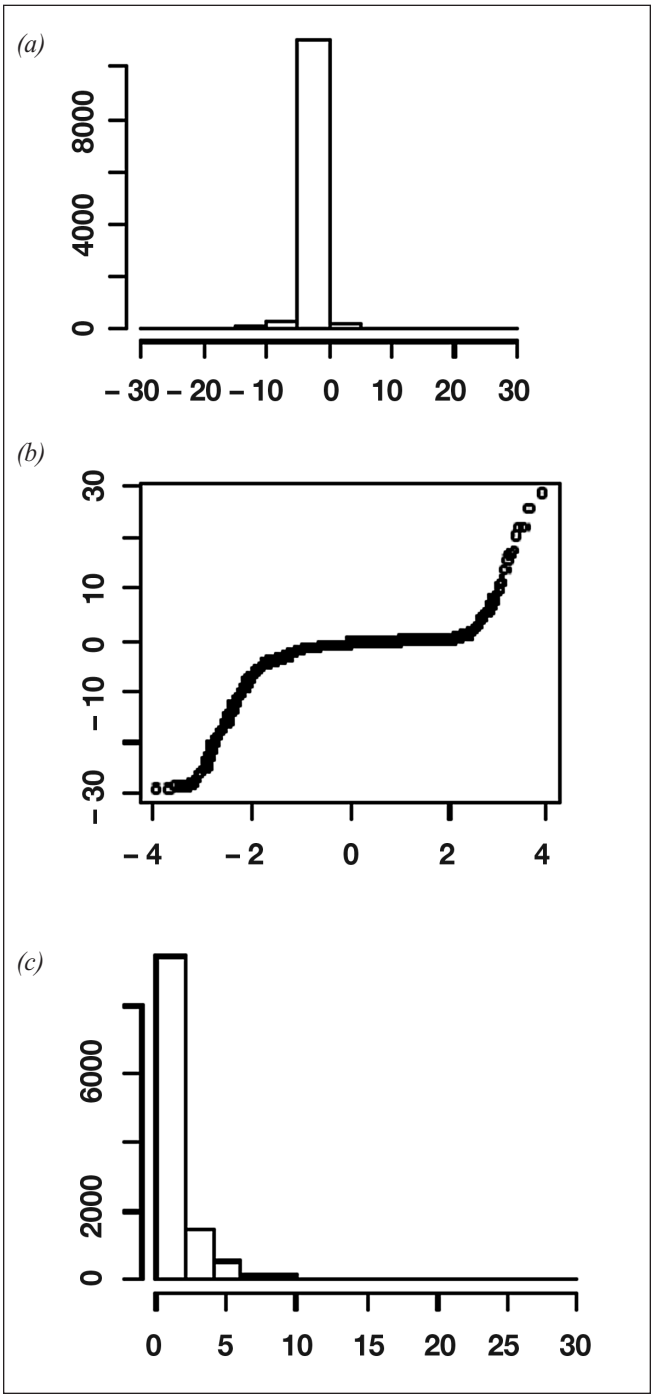


Figure 7.—*Semiparametric penalized spline model zero-inflated Poisson number of trees killed by insects: (a) residuals histogram. (b) Q-Q plot, and (c) histogram of fitted values.*



Q-Q plots of residuals, and histogram of predicted values for a SPSM fit to the number of trees killed by insects with distribution imposed on the response being Poisson and ZIP (Lambert 1992); we give these same plots for a local polynomial regression (LOESS) model fit to the number of trees killed by insects with distribution imposed on the response being ZIP and negative binomial (NB). Poisson was chosen first because these are count data. As we can see, however, the ZIP and NB performed better, which would be expected because the data are overdispersed. For the two ZIP models (figures 7 and 8), the SPSM and LOESS, it is hard to tell which one was best. The SPSM seemed to reach out to the extremes of the data better but fit worse in the middle range of the data. See figures 8 through 11.

We next fit models to species richness on phase 2 plots in AL and GA. The histogram of the number of species on plots in AL and GA is given in figure 12. These data had a sample variance-to-mean ratio of 4.1. We scaled species richness to [0,1] by the maximum number of species on plots, and we show a fitted lognormal and inverse Gaussian distribution to these species richness distribution. We included covariates of elevation, forest cover type, FIA unit, physiographic regions, and ecoregion. These covariates were chosen because they could be determined for prediction points where no FIA plots are present.

Figure 8.—*Local polynomial regression zero-inflated Poisson number of trees killed by insects: (a) residuals histogram and (b) Q-Q plot.*

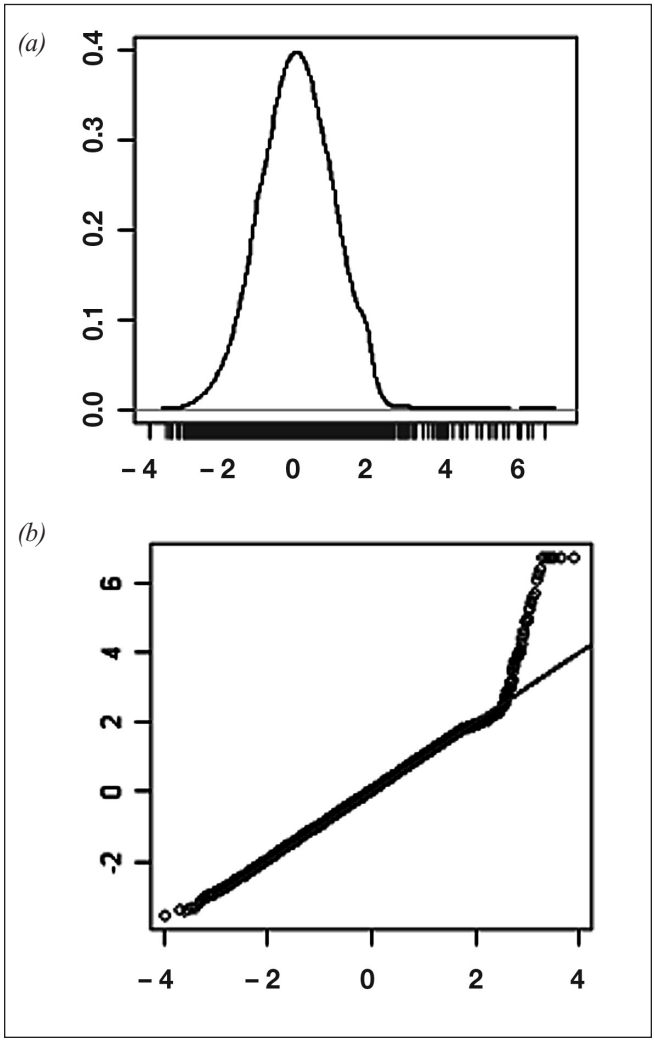


Figure 9.—*Local polynomial regression negative binomial number of trees killed by insects: (a) residuals histogram and (b)Q-Q plot.*

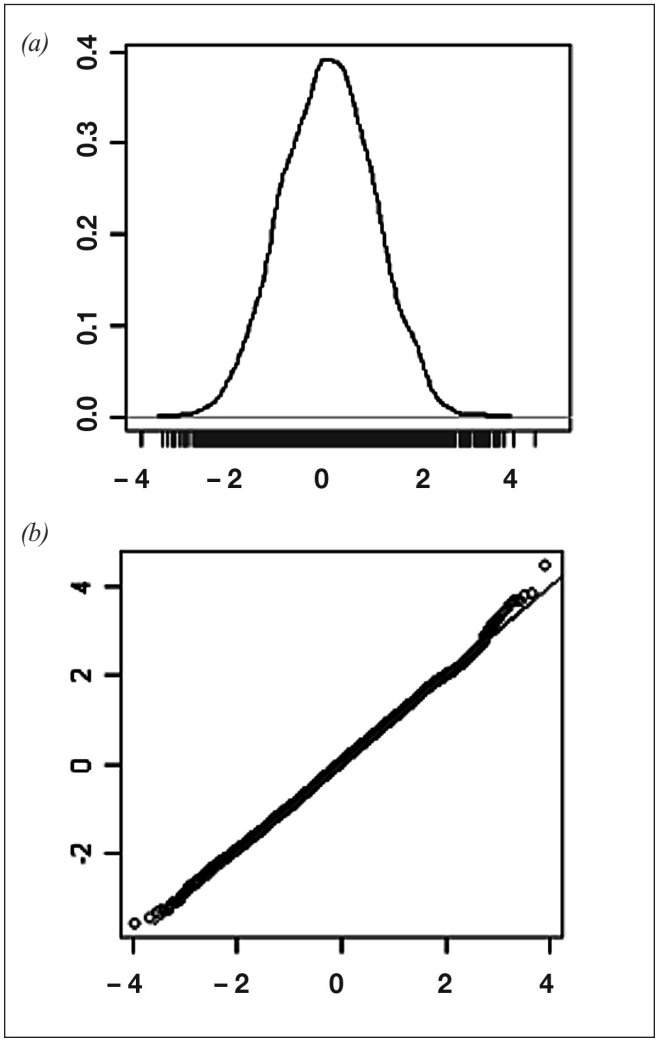


Figure 10.—*Local polynomial regression proportion of trees killed by insects: (a) residuals histogram, (b) Q-Q plot, and (c) histogram of fitted values.*

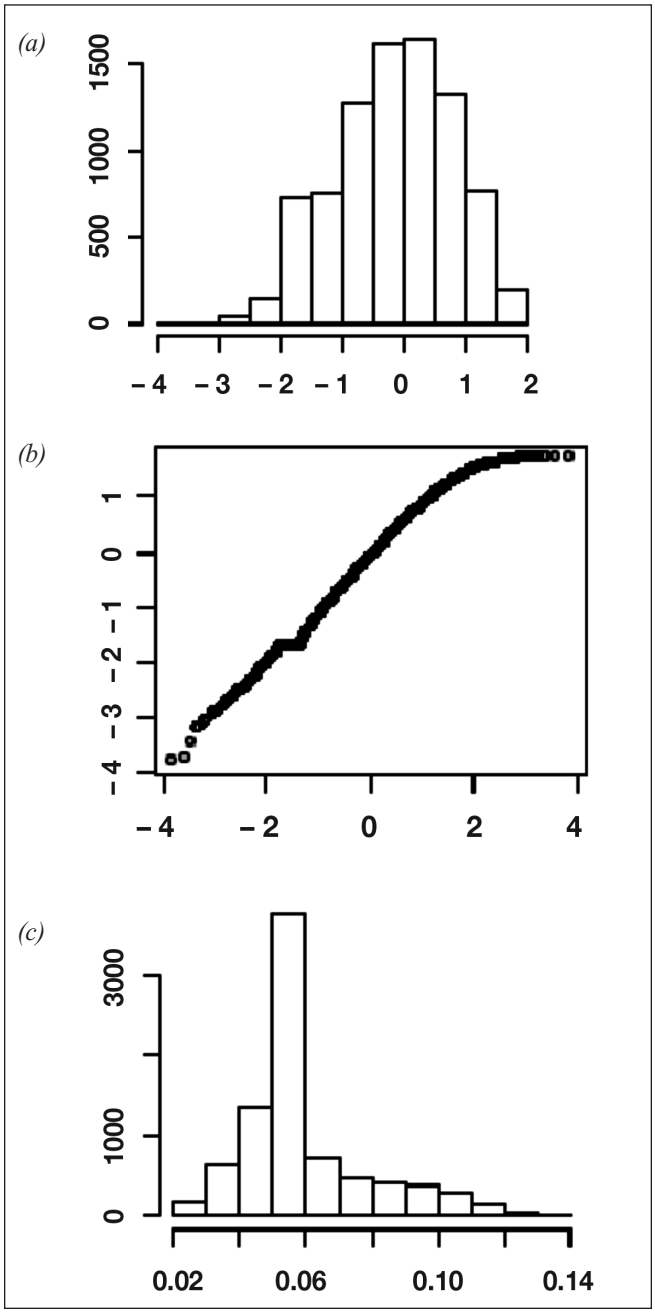


Figure 11.—*Local polynomial regression BI proportion of trees killed by insects: (a) residuals histogram, (b) Q-Q plot, and (c) histogram of fitted values.*

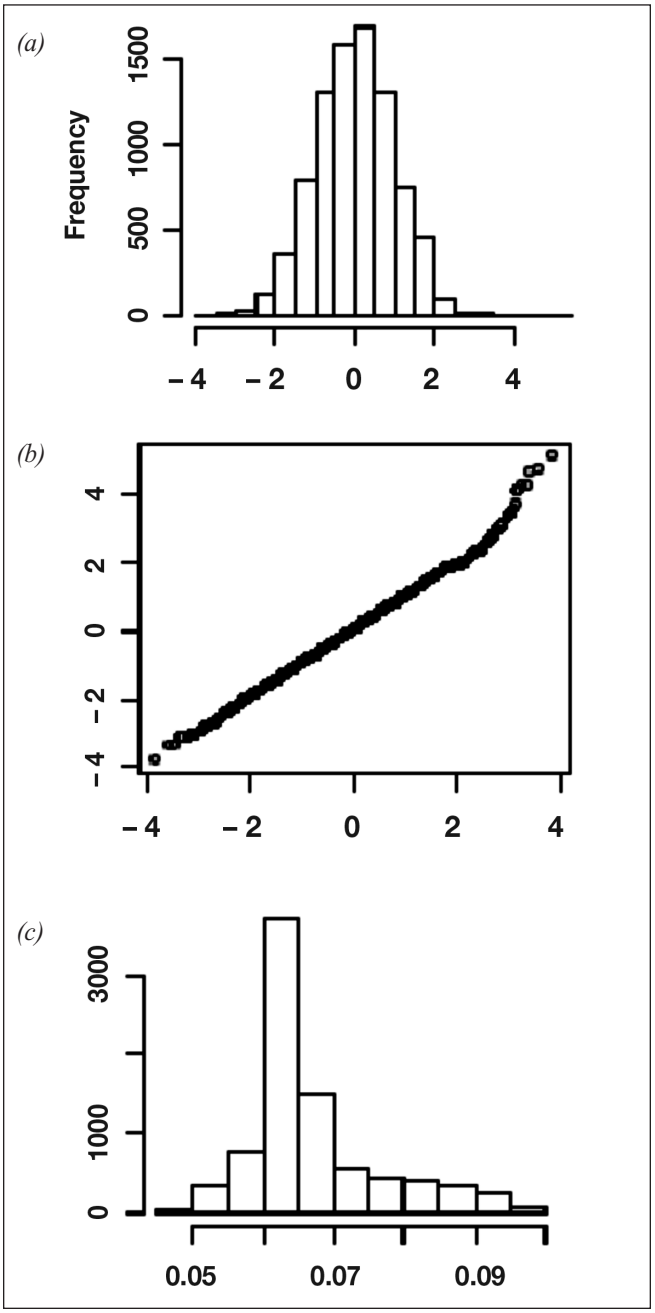
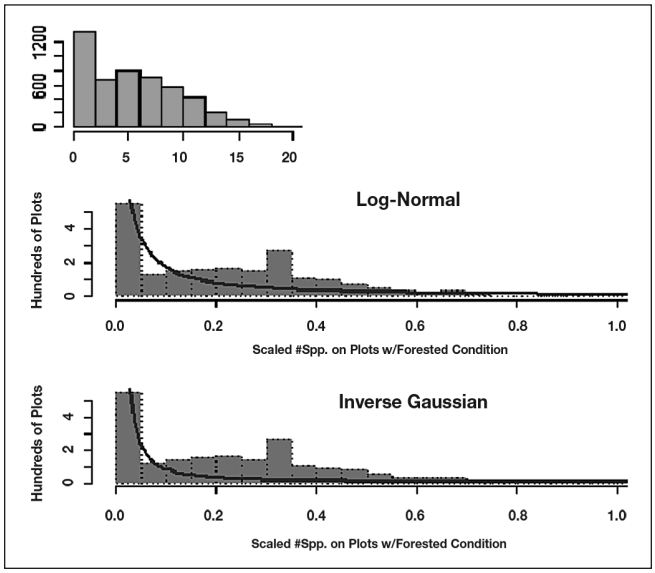


Figure 12.—*Top: histogram of species richness counts on plots in AL and GA with at least one forested condition. Middle and bottom: Two distributions fitted to the scaled species counts.*



In figures 13 and 14, we can see that models fit to the less dispersed plot-level value of scaled species richness, including covariates, have better residuals and histograms of fitted values more resembling the histogram of the measured values. Nevertheless, the models we chose so far, even though flexible, could not handle the extreme variation in the data we tried to model. We will continue with other spatial methods to test for significance. We will also include information from remote sensing, such as the FHM aerial surveys.

Figure 13.—*Local polynomial regression beta-inflated species richness, scaled to [0,1], only on plots in AL and GA with forested condition: (a) residuals histogram, (b) Q-Q plot, and (c) histogram of fitted values.*

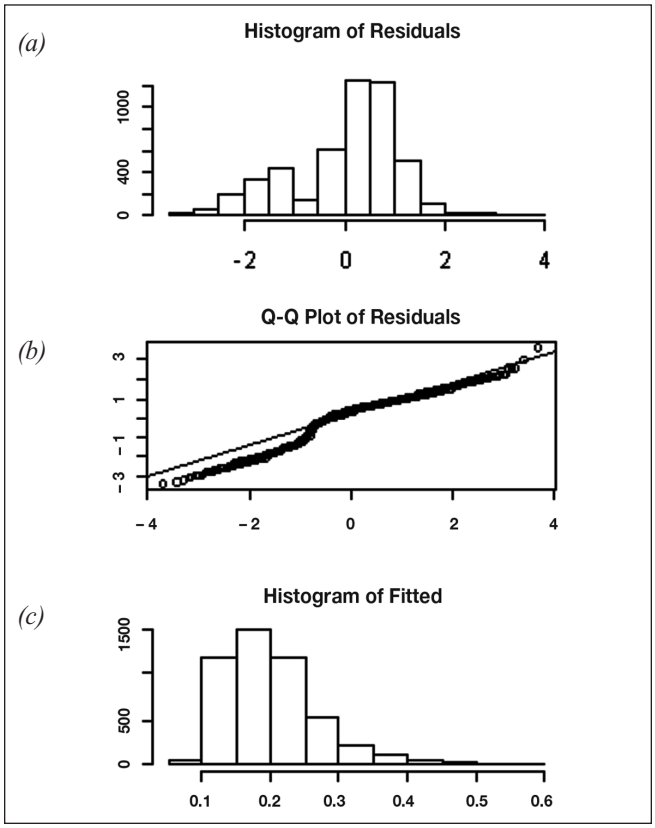
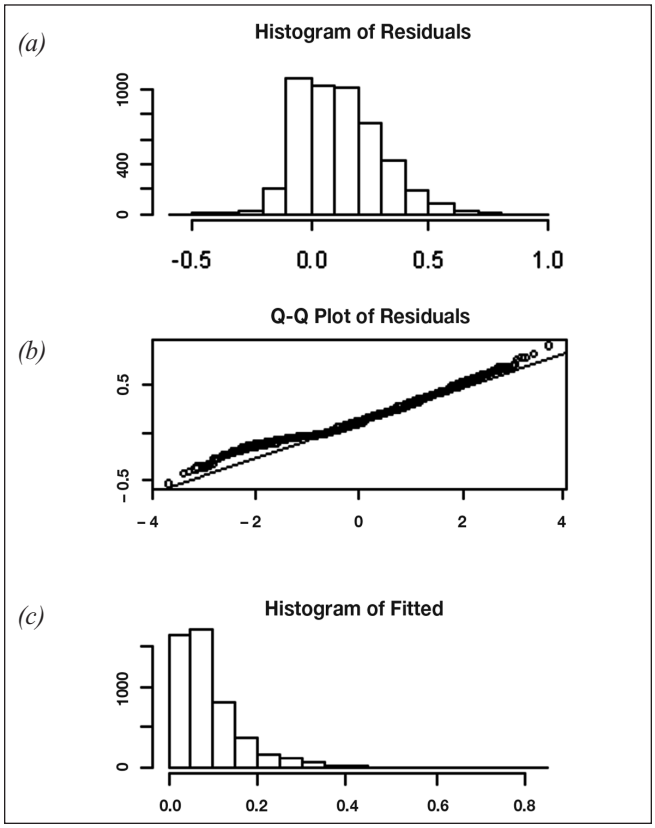


Figure 14.—*Local polynomial regression lognormal species richness, scaled to [0,1], only on plots in AL and GA with forested condition: (a) residuals histogram, (b) Q-Q plot, and (c) histogram of fitted values.*



Literature Cited

Agarwal, D.K.; Gelfand, A.E.; Citron-Pousty, S. 2002. Zero-inflated models with application to spatial count data. Environmental and Ecological Statistics. 9: 341-355.

Barry, S.C.; Welsh, A.H. 2002. Generalized additive modelling and zero inflated count data. Ecological Modelling. 157: 179-188.

Cressie, N.A. 1993. Statistics for spatial data. New York: John Wiley. 900 p.

Cressie, N. 2006. Spatial prediction for massive datasets. http://wwwrses.anu.edu.au/cadi/Whiteconference/papers/CressieMassiveData.pdf. (August).

Fahrmeir, L.; Echavarría, L.O. 2006. Structured additive regression for overdispersed and zero-inflated count data. Applied Stochastic Models Business Industry. 22: 351-369.

Gschlobl, S.; Czado, C. 2006. Modelling count data with overdispersion and spatial effects. http://www.stat.uni-muenchen.de/sfb386/papers/dsp/paper475.pdf. (August).

Hall, D.B. 2000. Zero-inflated Poisson and binomial regression with random effects: a case study. Biometrics. 56: 1030-1039.

Johannesson, G.; Cressie, N. 2004. Finding large-scale spatial trends in massive, global, environmental datasets. Environmetrics. 15: 1-44.

Lambert, D. 1992. Zero-inflated Poisson regression, with an application to defects in manufactoring. Technometrics. 34(1): 1-4.

Li, C.; Lu, J.; Park, J.; Kim, K.; Brinkley, P.A.; Peterson, J.P. 1999. Multivariate zero-inflated Poisson models and their applications. Technometrics. 41(1): 29-38.

Rathbun, S.L.; Fei, S. 2006. A spatial zero-inflated Poisson regression model for oak regeneration. Environmental and Ecological Statistics. 13: 409-426.

Rigby, R.A.; Stasinopoulos, D.M. 2005. Generalized additive models for location, scale and shape. Applied Statistics. 54(3): 507-554.

Ruppert, D.; Wand, M.P.; Carroll, R.J. 2003. Semiparametric Regression. New York: Cambridge University Press. 385 p.

Shabenberger, O.; Gotway, C.A. 2004. Statistical methods for spatial data analysis. Boca Raton, FL: Chapman & Hall. 488 p.

Zhu, Z.; Evans, D.L. 1994. U.S. forest types and predicted percent forest cover from AVHRR data. Photogrammetric Engineering & Remote Sensing. 60(5): 525-531.

Implementing the Measurement Interval Midpoint Method for Change Estimation

James A. Westfall[1], Thomas Frieswyk[2], and Douglas M. Griffith[3]

Abstract.—The adoption of nationally consistent estimation procedures for the Forest Inventory and Analysis (FIA) program mandates changes in the methods used to develop resource trend information. Particularly, it is prescribed that changes in tree status occur at the midpoint of the measurement interval to minimize potential bias. The individual-tree characteristics requiring midpoint values depend on the predictor variables needed to compute tree volume. Tree diameter change models are used to predict midpoint values for both future and past conditions. These updated diameters are used in conjunction with other information in a height model to obtain midpoint merchantable heights. These estimated diameter and height values are used to predict tree cubic-foot volume at the measurement interval midpoint. Limitations encountered in implementing this system included lack of information for some trees and inconsistencies between observed and updated values. A comparison is made between the previous method and the newly adopted technique, and effects on components of change are examined. Net change is unaffected by the new methodology.

Introduction

Since the passage of the 1998 Farm Bill that mandated an annualized approach to forest inventory in the United States, the Forest Service, U.S. Department of Agriculture's Forest Inventory and Analysis (FIA) program has been striving to develop more consistency among the regional FIA units. This move toward improved consistency resulted in the adoption of nationally consistent sampling design and estimation procedures (Bechtold and Patterson 2005). These new protocols include using stand- and tree-level attributes that reflect conditions at the midpoint of the sample plot measurement interval when estimating components of change (Scott *et al.* 2005). This approach assumes that, on average, events such as tree mortality, tree harvest, and conversions to and from forestland occur at the midpoint between inventory measurements. Under this assumption, the resulting estimates for components of change should be unbiased.

These new methods differ from those traditionally used in the northeast (NE-FIA) region, where only observed data recorded at the times the plots were measured were used to compute estimates of change components. This approach did not account for what tree volumes were at the time the change occurred, which caused bias in the estimates of change components. In this article, differences in estimates between the two methods are compared and obstacles encountered when implementing the new procedure are described along with potential solutions.

Data

The data used in this study are from NE-FIA sample plots in Maine. The 622 plots were initially measured under the annual inventory system (McRoberts 2005) in 1999 and were remeasured in 2004. These plots were partially or completely forested at either plot visit. Measurements at both times were taken on the 4-point cluster plot configuration (Bechtold and Scott 2005) in a spatially distributed sampling design (Reams *et al.* 2005). Each plot encompasses a land area of approximately 1/6 acre. Trees greater than or equal to 5.0 in diameter at breast height (d.b.h.) at either measurement were used for analysis.

[1] Research Forester, U.S. Department of Agriculture (USDA) Forest Service, Northern Research Station, Newtown Square, PA.
[2] Supervisory Forester, USDA Forest Service, Northern Research Station, Newtown Square, PA.
[3] Forester, USDA Forest Service, Northern Research Station, Newtown Square, PA.

Methods

Traditional NE-FIA

Trees were assigned to growth components based on observed history. The volume of live trees less than 5.0 in d.b.h. at initial inventory (hereafter denoted T1) but equal to or larger than 5.0 in d.b.h. at remeasurement (hereafter referenced as T2) was described as ingrowth (I). Accretion (A) was determined for trees that were measured, alive, and at least 5.0 in d.b.h. at both measurements. Removal (R) volumes were from trees at least 5.0 in d.b.h. measured at T1 but harvested before T2. Volume loss due to mortality (M) was determined from trees at least 5.0 in d.b.h. that were alive at T1 and dead at T2. Removal (R) and mortality (M) volumes were based on tree characteristics at T1. Net change (N) was defined as N = I + A – R – M. Each plot contributes an observed volume in each of the five categories (N, I, A, R, M). These observed values were used to compute population-level estimates for each individual component and for the overall net change.

Individual-tree gross volumes were computed from d.b.h. and bole height using equations of Scott (1981). Gross volume was converted to net volume using observed percent cubic-foot cull (USDA Forest Service 2004).

Measurement Interval Midpoint

Under the newly prescribed midpoint approach, the assignment of trees to components relied on predicted d.b.h. at the midpoint. Ingrowth (I) is the volume of trees at the time they grow across the minimum d.b.h. threshold (5.0 in) between T1 and T2; however, only the volume at the ingrowth threshold (5.0 in d.b.h.) was assigned to the ingrowth component. Any additional growth beyond 5.0 in d.b.h. was attributed to the accretion (A) component. Volumes from trees that were measured, alive, and at least 5.0 in d.b.h. at both measurements were assigned to the accretion (A) component (no midpoint values were needed). Removal (R) volumes were from trees at least 5.0 in d.b.h. measured at T1 but harvested before T2. For these trees, growth from T1 to the midpoint was assigned to accretion (A) and the volume at the midpoint was attributed to R. Similarly, volume loss due to mortality (M) was determined from trees at least 5.0 in d.b.h. that were alive at T1 and dead at T2. Accretion was computed for the growth on these trees from T1 to the midpoint, and the midpoint volume was assigned to the M component. The definition of overall net change and computation of gross and net volumes for individual trees were identical for both methods.

Obtaining Midpoint Values

The above description of the midpoint method demonstrates that a mechanism for providing midpoint values was needed. For this study, three variables were necessary for computing midpoint volumes— d.b.h., bole height, and cubic-foot cull percent. An additional complication was that projections from observed measurements needed to be both from future (mortality and removal trees) and past (ingrowth) perspectives. The process used in this method is outlined as follows:

1. Predict the appropriate (future or past) relative change in diameter using the equations from Westfall (2006). The estimated relative change is used to compute the midpoint d.b.h.
2. Use this midpoint d.b.h. in the height model developed by Westfall and Laustsen (2006). For this study, coefficients developed for regionwide application were used (table 1). Assume the other model predictor variables did not change from the time they were observed to the interval midpoint. This model prediction will provide a midpoint bole height. Note that sometimes inconsistencies will occur between observed and modeled values (e.g., the midpoint height is smaller than the height at T1). To eliminate these problems, the harmonic proportioning method of Sheffield and Schweitzer (2005) was implemented. This technique uses observed and predicted values at the time the tree was measured to adjust the predicted values for the midpoint.
3. Use the midpoint d.b.h. and bole height to predict gross volume at midpoint. Finally, multiply the gross volume by 1 minus cubic-foot cull percent to obtain midpoint net volume. The cubic-foot cull percent was assumed to have not changed from the most recent measured value.

Table 1.—*Estimated coefficients for NE-FIA regionwide use in height model from Westfall and Laustsen (2006)*[a].

Group[b]	n	$\hat{a}_0$	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$	$\hat{a}_5$	$\hat{a}_6$	$\hat{a}_7$
1	19807	-5.3174	101.7568	113.3938	94.1222	0.0701	0.0114	0.2937	2.6457
2	21700	-4.0879	80.9974	87.1315	72.7158	0.1037	0.0118	0.3178	2.7183
3	18719	-4.5683	79.6412	85.7464	72.4546	0.1124	0.0093	0.3557	2.5688
4	20628	-4.6716	88.5695	100.2259	82.4241	0.0797	0.0061	0.3832	2.2169
5	10850	-5.3714	97.5379	106.3982	93.0355	0.0792	0.0106	0.3753	2.7574
6	11321	-4.4011	72.6108	76.3698	69.0910	0.0856	0.0079	0.3098	2.7573
7	27235	-5.2346	92.1252	97.7856	88.4368	0.1071	0.0069	0.3267	1.0057
8	7016	-5.4536	117.1932	125.0373	109.9181	0.0934	0.0083	0.4022	1.0807
9	20527	-5.0346	96.7274	104.8101	90.9648	0.1046	0.0099	0.3591	1.1473
10	11694	-5.3508	101.5087	111.4904	94.1044	0.0866	0.0093	0.3584	1.2195
11	30489	-5.3292	89.8346	95.6066	86.4912	0.0897	0.0049	0.0780	0.1458
12	15448	-6.0106	104.1379	109.1728	98.2957	0.0746	0.0020	0.1413	0.1448
13	2494	-4.8963	108.0527	120.0344	105.3420	0.0900	0.0068	0.4163	1.1486
14	30243	-4.9946	91.5406	101.1159	86.9265	0.1023	0.0088	0.3699	0.9236
15	10833	-5.0550	97.5896	100.4159	92.9871	0.1025	0.0064	0.4685	1.0811
16	8134	-5.7404	105.8475	115.6039	96.1947	0.0972	0.0057	0.4044	1.1299
17	33815	-5.6585	100.6423	108.9352	93.9011	0.0917	0.0078	0.4060	1.1889
18	59269	-5.1022	89.5017	95.8707	86.3918	0.1076	0.0069	0.3463	1.0250

[a] $$H = \left(\hat{a}_0 D + \hat{a}_1 CC_1 + \hat{a}_2 CC_2 + \hat{a}_3 CC_3\right)\left(1 - \exp\left(-\hat{a}_4 DBH\right)\right)^{\left(\hat{a}_5 CR + \hat{a}_6 TC + \left((D/DBH)+0.01\right)^{\hat{a}_7}\right)}$$

[b] *Generally from Scott (1981). Contact an author of this paper for complete listing.*

Results and Discussion

The traditional and midpoint methods take different approaches to assigning volume to change components. Assuming the plots represent a simple random sample, estimates for each component were calculated for both methods (table 2). For ingrowth, the midpoint method produces a smaller estimate, because ingrowth volumes are always computed at the ingrowth threshold of 5.0 in d.b.h. The traditional method used the observed d.b.h. at the measurement subsequent to the crossing of the 5.0-in d.b.h. threshold, which would be a value of 5.0 or (often) higher. The accretion component under the midpoint method is notably larger when compared to the traditional approach. As noted above, the traditional method assigns all volume to ingrowth based upon the measurements at T2. The traditional method also did not account for any growth on mortality and removals—the volumes were based on observed data at T1. Thus, the difference arises from growth on ingrowth, mortality, and removals being added to the accretion component for the midpoint method. Volume of mortality and removals are also higher when using the midpoint approach. Because the midpoint values are used instead of the observed data at T1, the individual tree volumes are larger and the estimates for mortality and removals increase. When evaluating net change, the traditional and midpoint methods are identical.

Differences also occur in the precision of the estimates when comparing the traditional and midpoint methods. Table 3 shows that the standard error for ingrowth is nearly 10 percent less for

Table 2.—*Estimates of cubic-foot volume change by component on forest land in Maine for traditional NE-FIA and midpoint methods*

	Method		
Component	Traditional	Midpoint	Difference
Ingrowth	165,681,502	131,375,978	-34,305,524
Accretion	875,275,327	964,709,865	89,434,538
Mortality	357,180,873	370,311,502	13,130,629
Removals	606,569,440	648,567,825	41,998,385
Net	77,206,516	77,206,516	0

Table 3.—*Standard errors for estimates of cubic-foot volume change by component on forest land in Maine for traditional NE-FIA and midpoint methods.*

Component	Method		Percent of Traditional
	Traditional	Midpoint	
Ingrowth	12,738,701	11,514,120	90.4%
Accretion	34,781,244	34,302,255	98.6%
Mortality	21,999,892	22,715,502	103.3%
Removals	68,930,176	73,472,197	106.6%
Net	84,599,681	84,599,681	100.0%

the midpoint method. This difference is partially attributable to ingrowth volumes always being computed at the 5.0-in d.b.h. threshold under the midpoint approach, which results in less interplot variation. A related factor is that plots having ingrowth have a smaller value and the mean of the distribution is closer to zero. This factor is important because a number of plots have zero ingrowth and having these values closer to the mean reduces variance. The standard error for the accretion component is also smaller under the midpoint system. This difference is primarily due to the addition of accretion on ingrowth, mortality, and harvest trees. For instance, a plot that was entirely harvested would have zero accretion under the traditional method; however, this plot would have a nonzero accretion component using the midpoint method. This circumstance moves some plots closer to the mean of the distribution, which reduces the standard error.

The standard errors for mortality and removals components are larger when the midpoint method is used. This factor is a direct consequence of increased mortality and removal volumes that result from using the midpoint tree size instead of the tree size at T1. The means for mortality and removals increase under this scenario; however, a relatively large number of plots have zero mortality (174/622 = 28 percent) and/or zero removals (502/622 = 81 percent). The zero values for these plots are further from the mean under the midpoint method, which produces an increase in variance. The relationship between the amount of increase and number of zero-valued plots is evident by the larger increase for the removals component, which has substantially more zero-valued plots than mortality has. Because net change remains the same at the plot level, the net change standard errors for both methods are equal.

Several issues needed to be resolved when computing midpoint values for all trees. First, the values of many of the tree-level predictor variables in the d.b.h. and height models were assumed not to have changed since the last observation (i.e., crown class, crown ratio, tree class). This method is probably reasonable for short remeasurement intervals (~ 5 years); however, the validity of this approach may become more questionable as measurement intervals increase.

Another issue that arose was missing values for some predictor variables for certain trees and was most problematic for sapling trees (less than 5.0-in d.b.h.) that were alive at T1 but were either mortality or removal at T2. These trees needed midpoint values, but no data is collected for tree class (needed for height model) or cubic-foot cull (needed to estimate net volume). For this study, these attributes were examined for trees between 5.0 in and 6.0 in d.b.h. Nearly 75 percent of the trees in this diameter range were tree class code 2 (acceptable quality). Thus, for saplings missing tree class information, a 2 was assigned. Similarly, the mean cubic-foot cull percent was computed as approximately 6 percent; this value was assigned when the information was missing. Other possible solutions include (1) beginning to collect these data in the field, (2) respecifying the updating models so these variables are not needed, and (3) randomly selecting values from a distribution of valid values. In this study, nearly 6 percent of trees crossing the 5.0-in d.b.h. ingrowth threshold either died or were removed during the measurement interval. Factors affecting the proportion of trees in this category include length of measurement interval, site quality, stand age, and tree size/density relationships.

Last, an appropriate method for handling standing dead trees needed to be determined. Unlike trees that died and fell down during the measurement interval, standing dead trees are measured for d.b.h., bole height, and cubic-foot cull percent, so net volume can be computed and the use of midpoint values is not necessarily needed. Using observed data instead of modeled midpoint values, however, could create bias because it is unknown whether the tree still retains bark where the d.b.h. measurement is taken. Thus, trees that have a tendency to shed bark earlier would tend toward smaller volumes than trees of the same size that retain bark longer. To avoid this potential

bias, all mortality volumes were based on projected midpoint attributes, regardless of whether the tree was measured at T2.

Conclusion

The midpoint method is more difficult to execute due to the need to produce midpoint values for all predictor variables in the volume equation. Acquiring mechanisms to produce these values may require significant resources and time (e.g., model development). In addition, a number of practical assumptions may be needed to implement the system across a wide range of tree history patterns. Overall, the level of difficulty encountered primarily depends on how many variables needed to be updated and what method(s) and information are needed to compute the updated values.

The justification for implementation of the midpoint method is reduced bias. Clearly, the traditional method overestimated ingrowth and underestimated the other components. It could also be asserted that the traditional method standard errors for the individual components were biased, because it was shown that the midpoint method standard errors were notably different. Estimates and standard errors for overall net change (I+A-M-R), which is often the element of interest, were identical for both methods. Compared to the traditional method, the midpoint method should provide more accurate estimates of components of change for forest resource conditions.

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.

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. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station: 27-42.

McRoberts, R.E. 2005. The enhanced Forest Inventory and Analysis program. 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: 1-10.

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. 1981. Northeastern forest survey revised cubic-foot volume equations. Res. Note NE-304. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station. 3 p.

Scott, C.T.; Bechtold, W.A.; Reams, G.A.; Smith, W.D.; Westfall, J.A.; Hansen, M.H.; Moisen, G.G. 2005. Sample-based estimators used by 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. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station: 43-67.

Sheffield, R.M.; Schweitzer, C.J. 2005. Modeling missing remeasurement tree heights in forest inventory data. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C.; McWilliams, W.H.; Cieszewski, C.J., eds. Proceedings, fourth annual forest inventory and analysis symposium. Gen. Tech. Rep. NC-252. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station: 181-186.

U.S. Department of Agriculture (USDA) Forest Service. 2004. Forest inventory and analysis northeast field guide: field data collection procedures for phase 2 plots. Vol. 5, ver. 2. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station. 252 p.

Westfall, J.A. 2006. Predicting past and future diameter growth for trees in the northeastern United States. Canadian Journal of Forest Research. 36: 1551-1562.

Westfall, J.A.; Laustsen, K.M. 2006. A merchantable and total height model for tree species in Maine. Northern Journal of Applied Forestry. 23: 241-249.

Mapping Forest Inventory and Analysis Data Attributes Within the Framework of Double Sampling for Stratification Design

David C. Chojnacky[1], Randolph H. Wynne[2], and Christine E. Blinn[3]

Abstract.—Methodology is lacking to easily map Forest Inventory and Analysis (FIA) inventory statistics for all attribute variables without having to develop separate models and methods for each variable. We developed a mapping method that can directly transfer tabular data to a map on which pixels can be added any way desired to estimate carbon (or any other variable) for a particular area. The method uses a remote-sensing layer coverage to stratify forest from nonforest, a standard double sampling analysis, and a definition of spatial variables for summarizing FIA data within the double sampling framework to link to maps. Further refinements are needed but the new method appears to be accurate, flexible, and straightforward.

Introduction

Forest Inventory and Analysis (FIA) inventories provide population statistics with appropriate sampling errors for many forest attributes according to the underlying sample design. Often, these statistics are presented in tabular format for categories of interest such as geographic areas (States or counties), landownerships, or ecological classifications. FIA's statistical design was developed long before today's explosion in remote-sensing and Geographic Information System (GIS) technologies, however, and contemporary users of FIA data often want more than statistical tables and charts—they want maps.

Generally, FIA data are used in spatial analysis either by mapping selected plots to view plot scatter or to use by using map-based technologies where remote-sensing information is modeled with the aid of FIA plots. A reliable methodology is lacking to more easily map FIA inventory statistics for all attribute variables without having to develop separate models and methods for each variable (such as k-Nearest Neighbor [kNN] classification for forest area, basal area, or volume).

This article presents an overview of a new mapping method that can directly transfer tabular data to a map on which pixels can be added in any way desired to estimate carbon (or any other variable) for a particular area. We describe our approach by way of an example that maps FIA data for the entire State of Nevada.

Have Design, Will Map

FIA uses the double sampling for stratification design, sometimes called two-phase sampling. In this simple approach, a large sample of phase-1 units are selected for an inexpensive measurement and then a smaller phase-2 subsample of the phase-1 units are remeasured in more costly detail. Neyman (1938) devised the method for sampling human populations and Bickford (1952, 1959) first applied it to FIA. Bickford (1952) used aerial photographs to obtain relatively cheap estimates of forest-area strata (phase 1) that were then combined with field plots (phase 2) to obtain forest statistics.

This design has served FIA well over the years (Bechtold and Patterson 2005, Born and Barnard 1983, Chojnacky 1998), but it has not taken advantage of recent advances in remote-sensing and GIS technologies for a spatial display of results. Typically,

[1] Adjunct Faculty, Virginia Polytechnic Institute and State University, Department of Forestry, 144 Rees Place, Falls Church, VA 22046. E-mail: dchoj@cox.net.

[2] Associate Professor, Virginia Polytechnic Institute and State University, Center for Environmental Applications of Remote Sensing, Department of Forestry, Blacksburg, VA 24061. E-mail: wynne@vt.edu.

[3] Postdoctoral Research Associate, Virginia Polytechnic Institute and State University, Center for Environmental Applications of Remote Sensing, Department of Forestry, Blacksburg, VA 24061. E-mail: cblinn@vt.edu.

an FIA table or chart summarizes forest statistics for various forest parameters or categories (fig. 1). Because FIA includes a wealth of species detail for forest cover, it would be desirable, although not necessary, to have map categories to reflect some of this information. Maps are not involved in the estimation of forest cover but they provide an attractive and credible way for a nontechnical audience to see roughly where different species or forest types might occur on a map instead of needing to consult statistical tables.

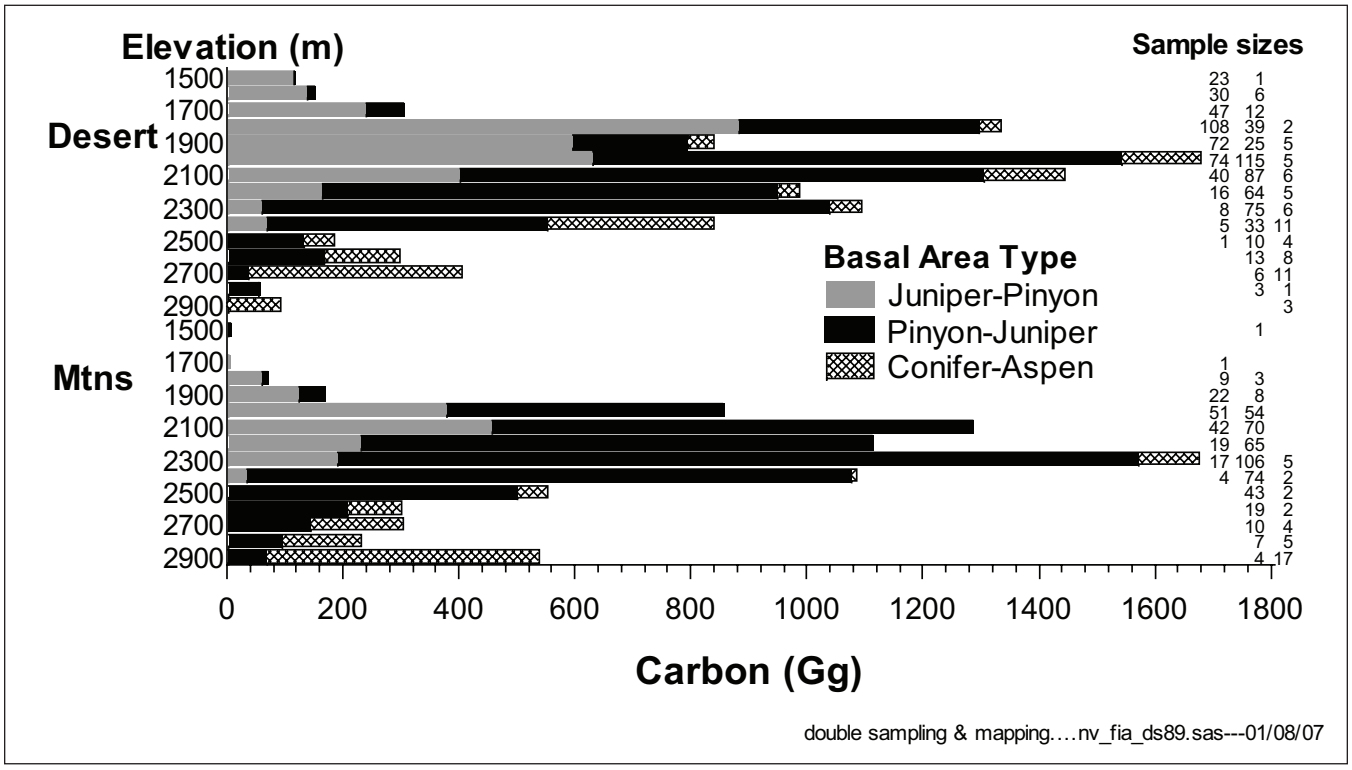
To overcome this shortcoming, researchers are using methods like kNN,, generalized linear models, and other techniques that seek to "fill in" or "expand" the nonsampled area between field plots (Franco-Lopez *et al.* 2001, Haapanen *et al.* 2004, McRoberts *et al.* 2005, Ohmann and Gregory 2002). These methods, however, ignore the double sampling design by using only the phase-2 field plots for developing models that estimate forest attributes directly from remote-sensing measurements.

We think it makes good sense to start with the full-blown double sampling design and look for ways to expand it to mapping applications. We developed a mapping concept that can directly transfer tabular data to a map on which pixels can be added in various ways, by first defining a set of variables called "category variables" that are common to both phase 1 and phase 2. The chart data can be mapped if the variables can be spatially depicted.

Several key concepts underlie our approach to mapping FIA data. First, for phase 1 of double sampling, a remote-sensing layer coverage or aerial photographs are needed for stratifying forest from nonforest.

Second, all calculations are done within the double sampling framework, which includes all the variables that FIA estimates. For our example, we estimate carbon; however, results could apply to volume, stand structure indexes, understory, deadwood,

Figure 1.—*Example of Forest Inventory and Analysis (FIA) tabular data for Nevada. Carbon amounts calculated for categories of Nevada FIA data are grouped into basal-area types, elevation classes, and two ecoregion classes ("Mtns" includes the Great Basin Mountains Section in the center of Nevada and "Desert" includes the rest of the State). Sample sizes are numbers of field plots for each category. A double sampling for stratification formula was used to summarize data.*



or any other variables that FIA estimates. Also, the double sampling design provides a framework to calculate variance estimates for each result. In other words, we have the power of a sample design backing our mapping method and do not have to start from scratch to develop statistical estimates for our maps.

Our third step is a new addition to double sampling. For this initial illustration, we link the double sampling design to maps by defining spatial variables for summarizing FIA data within the double sampling framework. We call these variables "category variables" for bridging the link between sampling design and map; the variables must be "categorical" to include groups of FIA data based on sufficient numbers of plots. For Nevada, we selected three category variables: forest cover type, Bailey's (1995) ecoregion, and 100-m elevation classes.

The Nevada Example

First, a Standard Analysis

A standard double sampling analysis was done for combining the phase 1 and 2 data. In phase 1, Moderate Resolution Imaging Spectroradiometer (MODIS) imagery (FIA program 2006b) was used for distinguishing forest from nonforest; for Nevada, 4.5 million pixels, 250 m^2 in size, covered the State. For phase 2, 14,138 FIA plots were available online (FIA program 2006a) for Nevada. Although the official survey date for Nevada was 1989, actual measurements ranged from 1979 to 1997, with 85 percent measured in 1982. Ownership of the plots included 69 percent Bureau of land Management, 13 percent private, and the rest national forest or other public ownership. The sampling design was roughly a grid that included much nonforest desert or rangeland; 88 percent of the plots were nonforest and the remainder were classified into three forest types according to species mix and dominate tree basal area.

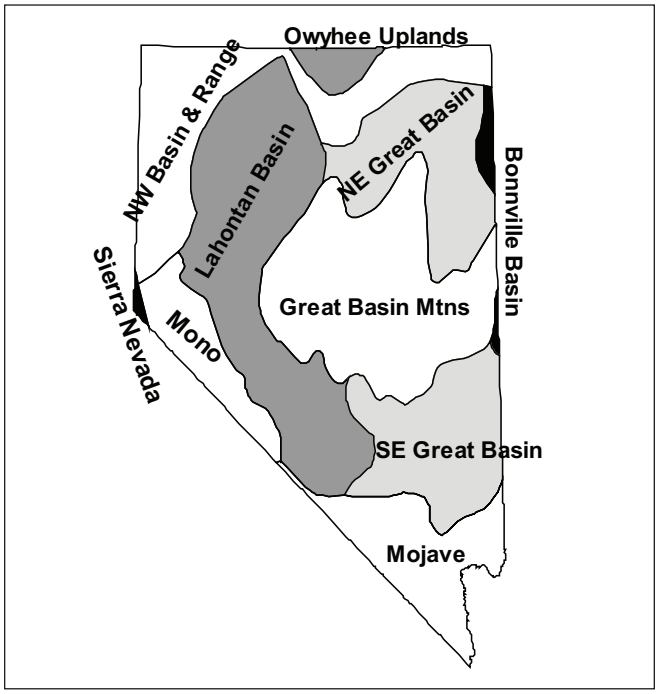
Basal area was calculated at diameter at breast height, with a conversion (Chojnacky and Rogers 1999) even though most of the woodland tree species were measured at ground line. Forest type or basal-area type then was defined in the following manner:

- Pinyon-juniper (P-J) plots: 953 included less than 67 percent juniper basal area and/or pinyon and/or mountain mahogany.
- Juniper-pinyon (J-P) plots: 589 were those in which more than 67 percent of the basal area on the plot included juniper species, could be pinyon, but was not mountain mahogany.
- Conifer-aspen (C-A) plots: 104 had basal area that was predominately ponderosa pine, fir, or other conifer species and/or aspen.
- Other plots: 2 included desert acacia species but were ignored for this study.

This classification roughly followed elevation. Predominately juniper (and perhaps some pinyon or J-P) was at lower elevations but decreased higher up, where pinyon-juniper mix (or P-J) became predominate. Mountain mahogany tended to mix with pinyon on the upper end of P-J, but we did not separate with another class. The other conifers and aspen (or C-A) occupied the highest elevations.

Of the 14,138 FIA plots, 1,645 forested plots were grouped into 34 strata defined as forest and nonforest; the 9 ecoregion sections (fig. 2); and a lower, mid-, and upper elevation zone corresponding to percentiles of Nevada elevation divided into thirds. Plots per stratum ranged from 2 to 481; the median was 13 plots.

Figure 2.—*Ecoregion sections for Nevada.*



The double sampling analysis was categorized into three basal-area types and 100-m elevation classes; a third variable corresponding to the ecoregion sections was considered but grouped into two classes to include some plots for each 100-m elevation class. The center Great Basin Mountain section (666 plots) was separated from the rest of the State in what we called Great Basin Desert (980 plots).

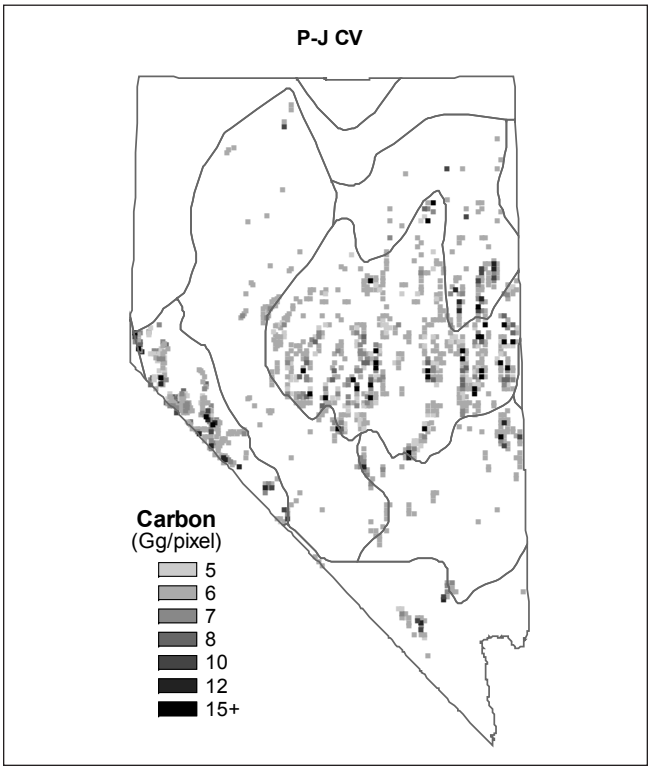
Next, Mapping the Results for Pinyon-Juniper

The double sampling results for carbon estimation (in Gg, or 10^9 g) show trends for the categories chosen (fig. 1, earlier). Although the chart provides useful information showing elevation distribution among J-P, P-J, and C-A forests, we next wanted to map the results.

The MODIS data used for phase 1 contained elevation and ecoregion information; however, they did not distinguish the forest types shown in figure 1. Therefore, we added a classification of forest into the three basal-area types by using logistic regression from work on another ongoing study. The model prediction corresponding to basal-area type with the largest probability of occurrence was assigned to each MODIS pixel for mapping. Then figure 1 data were assigned to appropriate pixels corresponding to elevation, ecoregion, and basal-area type.

Figure 3 shows carbon mapped for P-J in classes as carbon per pixel, which allows one to add the pixels in any manner to estimate carbon from the map. This feature is possible because the category means in figure 3 are simply divided by the number of pixels for that category on the map. The figure also includes coefficients of variation for the categories shown in figure 1.

Figure 3.—*Carbon for pinyon-juniper basal-area type (in fig. 1) mapped for Nevada Forest Inventory and Analysis data on left; coefficient of variation (CV) on right. Although it is difficult to illustrate map data in gray scale, the legends correspond to the range of data.*



Discussion

We have shown that one can (1) in advance of analysis, define categories common to phase 1 and phase 2; (2) estimate means and variances for the categories from double sampling; and (3) map results for category variables having spatial significance.

The model we used is preliminary but the method appears to be unbiased. It results in maps that are similar in concept to mapping FIA data by county, congressional district, or watershed but are more accurate because the phase 1 categories are used to separate forest from nonforest and further divide forest land into similar categories. Since the categorical variables are spatially defined in phase 1, the resulting maps cover a landscape with polygons defined by means, totals, or standard errors for attributes compiled from the sample design.

Defining good, robust categories is challenging, however. For example, instead of using ecoregion sections, it probably would have been more effective to use ecological provinces (the next scale up in Bailey's [1995] hierarchy) and to include neighboring States for full province analyses. Further study is needed on the definition of category variables for splitting the data because these definitions greatly affect sample sizes in splitting the data, which affects map precision.

A useful feature of the map is the ability to add "carbon per pixel" for any area of the map to estimate population totals. A

variance for this estimate is not so simple; a variance estimate should exist but we have not yet studied this.

Finally, mapping basal-area type for Nevada using MODIS data was challenging because we needed the additional logistic regression model to classify "MODIS forest" into basal-area type. This classification worked reasonably well except for a few vegetation categories with small sample sizes at upper and lower elevations where missing values resulted for mapping.

Conclusion

We have initiated a new method to combine map-based technology into the double sampling design that enables us to more easily map FIA inventory statistics for all attribute variables without having to develop separate models and methods for each variable. Because the approach builds on FIA's double sampling design, it offers all the power of the FIA database for map-based applications, which, we think, is a big advantage over some other modeling methods for expanding field plot grids.

Acknowledgments

The authors thank Ron Tymcio of the Forest Service, U.S. Department of Agriculture's Interior West Forest Inventory and Analysis Unit in Ogden, UT, who provided the MODIS data for Nevada; the Nevada Bureau of Land Management for funding studies in Nevada that allowed us to acquire data for this study; and Mary Carr of the Forest Service's Content Analysis Team Ecosystem Management Coordination Publishing Arts group for editing the manuscript.

Literature Cited

Bailey, R.G. 1995. Descriptions of the ecoregions of the United States. 2nd ed. Misc. Pub. 1391. Washington, DC: U.S. Department of Agriculture, Forest Service. 108 p.

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

Bickford, C.A. 1952. The sampling design used in the forest survey of the northeast. Journal of Forestry. 50(4): 290-293.

Bickford, C.A. 1959. A test of continuous inventory for national forest management based upon aerial photographs, double sampling, and remeasured plots. In: Proceedings, Society of American Forestry. Bethesda, MD: Society of American Forestry: 143-148.

Born, J.D.; Barnard, J.E. 1983. FINSYS-2: subsystem TABLE-2 and OUTPUT-2. Gen. Tech. Rep. NE-84. Broomall, PA: U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station. 133 p.

Chojnacky, D.C. 1998. Double sampling for stratification: a forest inventory application in the interior west. Res. Pap. RP-RMRS-7. Ogden, UT: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 15 p.

Chojnacky, D.C.; Rogers, P. 1999. Converting tree diameter measured at root collar to diameter at breast height. Western Journal of Applied Forestry. 14(1): 14-16.

Forest Inventory and Analysis (FIA) program. 2006a. Forest Inventory and Analysis national program. http://fia.fs.fed.us. (17 January 2007).

FIA program. 2006b. Interior west forest inventory and analysis, map products: interior west region. http://www.fs.fed.us/rm/ogden/map-products/intwest/iwmaps.shtml. (17 January 2007).

Franco-Lopez, H.; Ek, A.R.; Bauer, M.E. 2001. Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method. Remote Sensing of Environment. 77(3): 251-274.

Haapanen, R.; Ek, A.R.; Bauer, M.E.; Finley, A.O. 2004. Delineation of forest/nonforest land use classes using nearest neighbor methods. Remote Sensing of Environment. 89: 265-271.

McRoberts, R.E.; Holden, G.R.; Nelson, M.D.; Liknes, G.C.; Gormanson, D.D. 2005. Using satellite imagery as ancillary data for increasing the precision of estimates for the Forest Inventory and Analysis program of the USDA Forest Service. Canadian Journal of Forest Research. 35(12): 2968-2980.

Neyman, J. 1938. Contribution to the theory of sampling human populations. Journal of American Statistical Association. 33: 101-116.

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. Canadian Journal of Forest Research. 32(4): 725-741.

Precise FIA Plot Registration Using Field and Dense LIDAR Data

Demetrios Gatziolis[1]

Abstract.—Precise registration of forest inventory and analysis (FIA) plots is a prerequisite for an effective fusion of field data with ancillary spatial information, which is an approach commonly employed in the mapping of various forest parameters. Although the adoption of Global Positioning System technology has improved the precision of plot coordinates obtained during field visits, in many circumstances, the coordinate uncertainty of plot coordinates remains substantial. Because Light Detection and Ranging (LIDAR), an emerging remote sensing technology, in its discrete return form delivers three-dimensional data with centimeter precision, FIA field data registered to LIDAR data would yield exceedingly precise plot coordinates. In this study, a fully automated, three-dimensional variant of the world-view method is used to match individual trees, identified by processing dense (> 4 returns/square meter) LIDAR data, to trees tallied in 45 FIA-like plots, thus retrieving the coordinates of the plot center. Results indicate that this method yields precise plot registration in stands exhibiting local heterogeneity in forest structure, performs better in coniferous rather than deciduous stands, and is robust against the tolerances embedded in tree location and height measurements obtained in the field.

Introduction

Meaningful integration of forest inventory and analysis (FIA) plot information obtained in field visits with ancillary, spatially distributed information organized in Geographic Information System layers or imagery acquired via remote sensing requires precise plot registration. Many popular modeling approaches that use FIA information to map forest inventory parameters across the landscape (McRoberts 2006, Ohmann and Gregory 2002, Schroeder *et al.* 2006) operate on the implicit assumption that the coordinates (i.e., registration) of FIA plots are known with precision that exceeds (is smaller than) half the size of the plot or the resolution of the coarser ancillary data source (Heuvelink 1998). Using information from plots with registration that does not meet this precision standard likely reduces the fidelity of the mapping product by a magnitude that is a function of the registration error and the local spatial continuity of the mapped parameter (Goodchild *et al.* 1993).

Historically, FIA plot registration information was collected and maintained to primarily facilitate plot visitation. Plot registration was performed in the office by overlaying analog aerial photographs with maps, usually topoquads or digital raster graphics, on which the location of the plot was pin-marked. This process was laborious and prone to registration errors, especially in mountainous terrain. Attempts to substitute reference maps with satellite or high-resolution airborne multispectral imagery resulted in little or no gain in precision or efficiency; the resolution of satellite imagery was typically too coarse to allow fine tuning of the aerial photographs on the imagery and high-resolution airborne imagery was often poorly registered. The introduction and continuous upgrade of Global Positioning System (GPS) technology, used by FIA field crews since the mid-1990s, have improved plot coordinate precision but, in many instances, registration error has remained substantial, often exceeding the spatial extent of an FIA plot.

Unlike most maps and satellite or airborne imagery, data acquired by using Light Detection and Ranging (LIDAR), an emerging remote sensing technology, has no registration issues. In its discrete form, LIDAR data consist of a collection of points arranged in three dimensions. The points, or ‘returns’ as

[1] Research Forester, U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Forest Inventory and Analysis, Portland Forestry Sciences Lab, 620 SW Main Street, Suite 400, Portland, OR 97205. Phone: 503–808–2038. E-mail: dgatziolis@fs.fed.us.

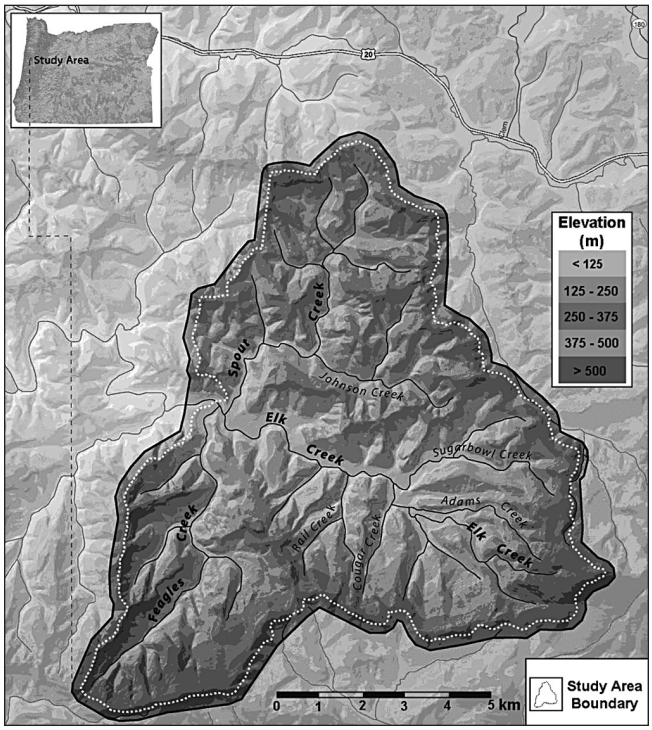
they are some times referred to, represent the locus of where narrow beams of infrared light, emitted from the airborne platform, are intercepted by the target that the beams illuminate. Because the three-dimensional location of each LIDAR point is recorded with centimeter precision, FIA field data registered to LIDAR data would yield exceedingly precise plot coordinates. This article introduces a novel, fully automated method, which uses tree data collected in a typical FIA plot and dense LIDAR data to retrieve the plot center coordinates. The method is evaluated using data from 45 FIA-like plots, established in a 9,500-hectare study area, located at the coastal forests of Western Oregon.

Methods

Study Area

The study area extends over about 9,500 hectares of the upper part of the Big Elk watershed located in Lincoln County, OR (fig. 1). More than 90 percent of the area is forestland, 47 percent of which is privately owned and under very intensive, timber-oriented management. The State of Oregon owns 1,550 hectares of land and 3,850 hectares are part of the Siuslaw National Forest where management has been limited to occasional precommercial thinning (very little of which occurred after 1984). Elevation in the study area ranges from 66 m to 1,123 m above sea level, and terrain is characterized by steep slopes that often exceed 100 percent. The forestland contains Douglas fir, bigleaf maple, and red alder, with the hardwoods dominating the buffer zones around the drainage network.

Figure 1.—*Map of the study area located at the coastal range of the State of Oregon.*



Vegetation and LIDAR Data

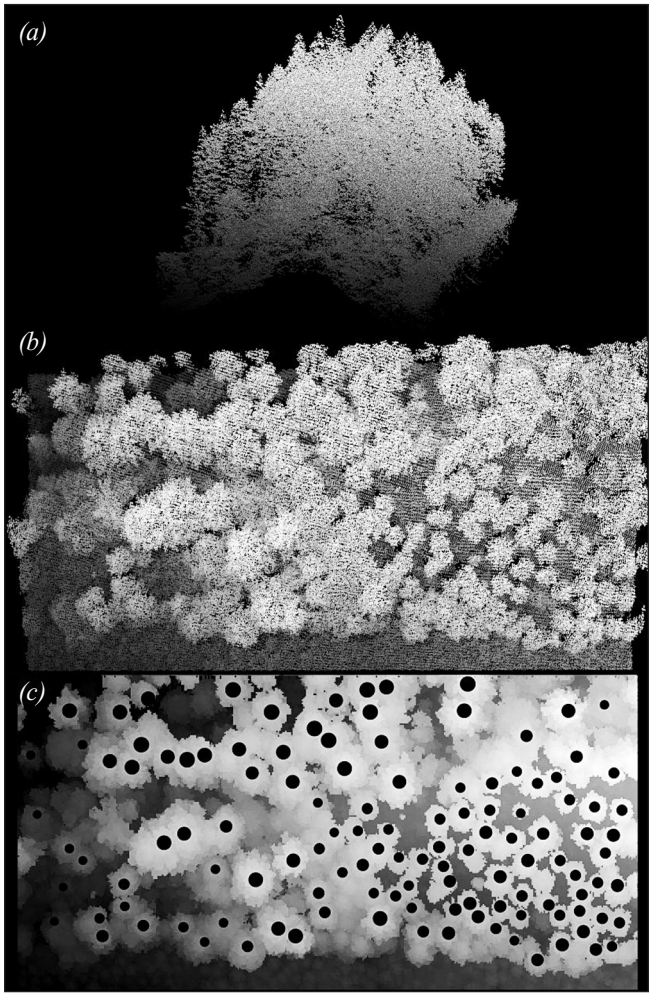
Forty-five FIA–like plots were established in the study area during the summer of 2005, stratified in classes of cover type (conifers, hardwoods, and mixed), tree size, and stand density. To ensure that in each plot an adequate number of trees would be present, the radius of the plot was expanded from the standard 7.32 m to 15 m. Plot centers were georeferenced by using a real time kinematic (RTK) GPS instrument. Compared to local benchmarks, the RTK instrument yielded a two-dimensional root mean square error of 2.6 cm. All trees with diameter at breast height (d.b.h.) exceeding 12.7 cm or of dominant or codominant status regardless of d.b.h. were tallied in each plot. Sketch maps depicting the presence, type, and height of understory vegetation (if any) were produced. LIDAR data over the study area was acquired at leaf-on conditions in July 2005 using an aircraft-mounted Optech 3100 system from an average height of 1,000 above ground along flight lines with 50 percent sidelap. The LIDAR instrument, at 71 kHz pulse rate, captured a 20^{o} scan width (10^{o} from nadir), averaged 9.81 points per square meter, and had spot spacing of 32 cm and laser footprint diameter of 33 cm.

Data Transcription to Point Patterns

Matching plot and LIDAR data required conversion to a common form. This conversion was accomplished by computing or identifying the tops of individual trees in the plot and on the LIDAR data sets. Using the distance and azimuth from the plot center and the tree height recorded by the field crew, the tree-dimensional coordinates of tops for dominant or codominant trees were computed. The origin of the coordinate system coincided with the plot center. It was assumed that all treetops were positioned vertically above corresponding tree bases, un-

less, as stated in the FIA field protocol, substantial tree leaning mandated recording of the leaning angle and computing an offset. The set of points representing individual treetops computed from field data is henceforth mentioned as the "plot point pattern." Identifying individual treetops in the LIDAR data set was a far more complex exercise. It involved processing the LIDAR data to generate ground and vegetation raster surfaces of 1 m resolution and then querying the vegetation surface to detect local maxima believed to correspond to the tops of individual tree crowns (fig. 2). The height of each identified tree was computed as the distance between the vegetation and ground surfaces at the two-dimensional location of the treetop. Gatziolis (2006) details the algorithms employed in the processing of the LIDAR data. The set of points representing individual treetops in real-world coordinates computed from the LIDAR data for a square area of approximately two hectares containing the plot is referred to as the 'stand point pattern' (fig. 3).

Figure 2.—*(a) Perspective and (b) nadir view of raw LIDAR data for one of the stands hosting a plot. (c) Raster representing the elevation of the vegetation surface derived from the LIDAR data. Darker/lighter tones correspond to lower or higher elevation. The dark circles indicate individual treetops identified by processing the vegetation surface raster. The size of the circles is proportional to tree height.*



Point Pattern Matching

Initial attempts to match plot and stand point patterns using Euclidean geometry, heuristic decision rules, and brute computing force proved inadequate and inefficient. It was realized that a pattern matching method should perform adequately in stands with few trees as well as stands with numerous trees and be robust against errors in point locations introduced during the data transcription stage. Common sources of such errors include the various tolerances in the measurement of tree location and height adopted by FIA to promote data collection efficiency, omission or commission of trees identified using LIDAR data, uncertainty in LIDAR-derived tree heights, and the ever-present discrepancies in two dimensions between treetop and base. After some additional experimentation, a modified version of the 'world-view' method (Murtagh 1992) was adopted (fig. 4).

The original world-view method retrieves the coordinates for a group of stars on an image acquired with a telescope by using reference star maps. After converting all the stars on the image to points, the method selects a star located near the centroid of the group and counts the number of stars present in each one-degree sector radiating outwards, or viewed, from the

Figure 3.—*Three-dimensional arrangement of (a) a stand point pattern derived from LIDAR data and (b) a plot point pattern derived from field data. Both patterns depict tops of individual trees. The vertical lines are included to enhance visualization. Scale is variable between and within each point pattern.*

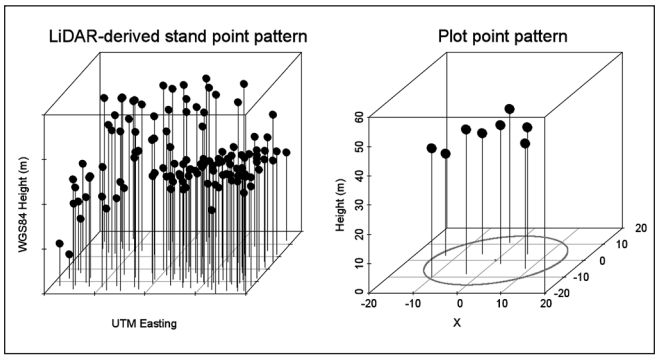
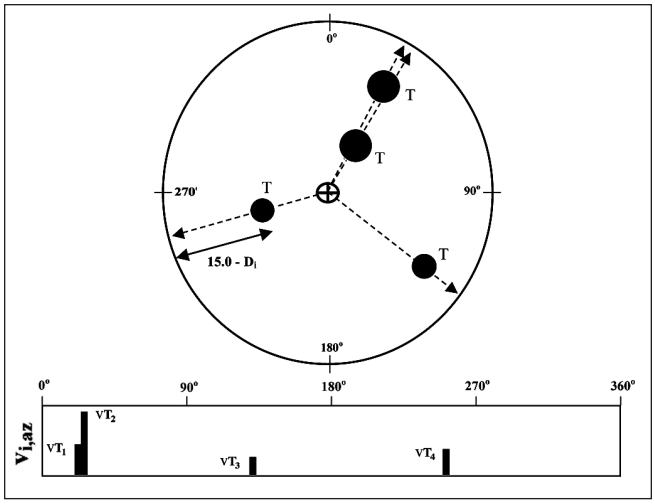


Figure 4.—*Graphical illustration of the modified world-view method for a plot point pattern. The top part features a nadir view of the plot with solid circles representing the location of four (T1 to T4) treetops. Circle size is proportional to tree height. Dash lines point at the azimuth at which a tree is viewed from the plot center. The lower part shows the values of the view vector contributed by each tree in the form of bars.*



central star. The 360-element long (view) vector of star counts is compared with the view vectors computed for all stars in the reference maps present in the general vicinity of the sky the telescope was pointing at the moment the image was acquired. The similarity of view vectors computed for the central star, once by using the image and once by using the reference maps, allows the central star to be identified in the reference maps and thus registers the star group.

Substituting the points representing stars with the points representing treetops provides a functional implementation of the world-view method. Unfortunately, the original form of the method exhibits four notable weaknesses when used in matching plot and stand point patterns. First, because it operates strictly in two dimensions, it ignores valuable tree height information. Second, it assumes that at least one tree is near the plot center. Where this is not the case, the tree labeled as 'central' would be at some distance from the plot center, and it would produce an asymmetrical plot view with azimuth-specific scope, longer for view sectors across the center of the plot and shorter in the opposite direction. Third, only the azimuth of a tree relative to the central tree affects the view vector; its distance from the central tree does not. Lastly, no provision exists on how errors of commission or omission should be handled. Apparently such errors do not exist in star reference maps, but are common where individual trees are identified by processing remotely sensed data. To neutralize these weaknesses, the original method was modified as follows: the plot view vector is now computed from the plot center instead of from the location of the tree closest to the plot center; stand views are computed for each vertex of a lattice of user-specified resolution covering the two dimensional extent of the stand; and instead of simply counting the tree(s) in a view sector, each tree in the plot contributes a value *v* to the view sector that corresponds to the tree's azimuth computed as

$$v_{i,az} = P_R - D_i + \left| (H_i - \overline{H}) \Big/ \overline{H} \right| * \mathrm{var}(H) \qquad (1)$$

where *az* represents the azimuth of tree *i* in reference to the plot center (view origin), H_i is the tree height, P_R is the plot radius, D_i is the distance of the tree from the plot center, *var(H)* is the variance, and $\overline{H}$ is the mean height of the trees in the plot. Hence, the closer a tree is to the plot center and the more its height differs from the mean height in the plot, the larger its influence on the view vector. The *var(H)* coefficient is, in essence, a scale adjustment between the distance and height components. The value of the view vector $\boldsymbol{V}$ at *az* is computed as the sum of $v_{i,az}$ for all trees (if any) present at the *az* sector of the view, or

$$V^T = \sum v_{i,1}, \sum v_{i,2,}, \ldots, \sum v_{i,360}\ , \forall i \qquad (2)$$

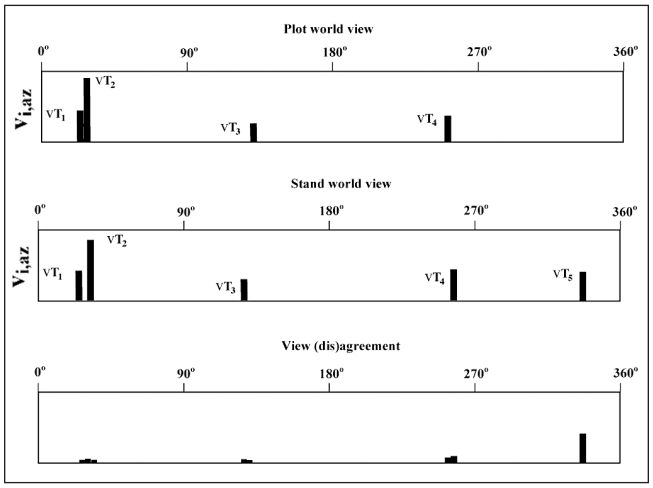
The plot world-view vector is subsequently compared to the stand-view vectors computed at each vertex of the lattice. The comparison entails computing the sum of absolute discrepancies between the azimuth-paired elements of the stand and plot vectors and producing a statistic

$$\Phi = \sum_{az=1}^{360} \left| PLOT \sum v_{i,az} - STAND \sum v_{i,az} \right|, \forall i \qquad (3)$$

that quantifies the agreement between the plot world view and each stand-view instance. The lattice vertex at which Φ is minimized is expected to represent the most likely location of the plot center within the stand. To ensure that the occasional imprecision in tree azimuths recorded in the field does not inflate Φ, the value of each one-degree plot view sector PLOT$v_{i,az}$ is subtracted for each of the five one-degree stand vector values

STAND$v_{i,]az-2,az+2]}$ centered on *az,* and the smallest absolute value is used in the computation of Φ (fig. 5). The performance of the method, in terms of its ability to retrieve the plot center, is evaluated by calculating the two-dimensional Euclidean distance (retrieval precision) between the stand location with the smallest Φ value and the plot center location recorded with the RTK instrument.

Figure 5.—*Graphical illustration of the view vectors computed for the plot (top) and the stand (middle) when viewed from the plot center. The collective height of the bars on the lower portion of the figure represents the amount of disagreement between the views.*



Results and Discussion

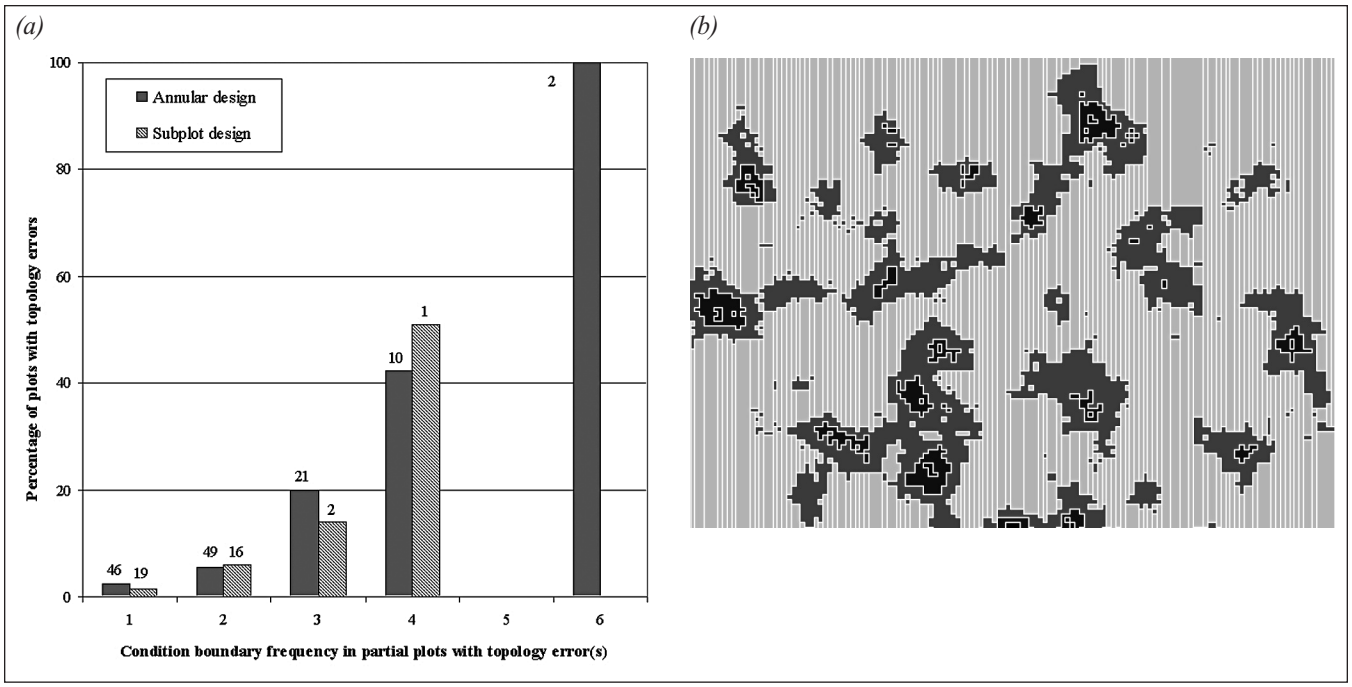
The modified world-view method applied over a stand containing a plot, ideally, will produce one group of distinctly small Φ values clustered spatially around a single stand-wide Φ minimum that corresponds to the plot center. Pending primarily on stand and plot conditions, the method is found to occasionally yield more than one clusters of local Φ minima. Rarely, it produces a nearly uniform distribution of Φ values with numerous weak local Φ minima. The latter case indicates method failure. In the study area, the modified world-view method failed for 4 of the 45 plots. For eight other plots, the retrieval precision exceeded the plot radius (table 1). For the remaining 33 plots, 73 percent of the sample total, retrieval precision was better (smaller) than the plot radius and, hence, superior to the expected registration precision currently available for FIA plots.

Plot center retrieval precision depends primarily on cover type and stand structure. Precision was higher for coniferous than deciduous stands and higher in the presence of local heterogeneity in tree height and spatial arrangement. The retrieval precision median for the 20 coniferous plots was 2.7 m. Figure 6a shows, in raster form, the spatial arrangement of Φ values computed over a coniferous stand. The portion of the raster in warmer tones depicts a single cluster of cells with small Φ values that envelope the plot center. Although a few other clusters of local minima are present, their values are too large to be considered candidate locations of the plot's center. In deciduous stands, the retrieval precision median was 6.1 m. Although the arrangement and shape of clusters in rasters of Φ values computed over stands is similar for all cover types, the value variability among local Φ minima for deciduous stands is typically lower than in coniferous stands. The lack of a pronounced Φ minimum hinders the process of determining which local minimum is the true location of the plot center among the many present, and often leads to sizeable retrieval errors (fig. 6b). Homogeneous stand conditions, regardless of cover type, cause either retrieval failure or large errors, as it is the case in plots installed in young, unthinned plantations of Douglas fir or high density red alder stands. Thinned Douglas fir plantations, even those having the operation performed on a grid or planting line, have adequate tree height variability to allow precise retrieval, a benefit attributed to the inclusion of the *var(H)* component in equation (1); without the scale adjustment between tree height and distance from the plot center provided by *var(H)*, the location of the plot center in plantations cannot be retrieved.

Table 1.—*Frequency of plots in classes of registration precision and cover type. Registration precision is the distance between the actual plot center and the location identified as plot center by the modified world-view method.*

		Cover type			
		Conifers	Mixed	Deciduous	Total
Registration precision	< 1 m	4	1	—	5
	1–5 m	12	4	2	18
	5–10 m	1	3	4	8
	10–15 m	—	1	1	2
	> 15 m	3	2	3	8
	NA	—	—	4	4
	Total	20	11	14	45

Figure 6.—*One-meter resolution raster of plot center retrieval precision* Φ *computed by using the modified world view method over (a) coniferous stand and (b) deciduous stand, each hosting a plot. Squares represent the true plot location and circles the location identified by the world-view method.*



A few other factors are collectively responsible for the discrepancy in retrieval precision between coniferous and deciduous plots. Dominant red alder and bigleaf maple trees growing near streams and creeks form nearly continuous canopies with few openings and have large, umbrella-shaped crowns composed of several branches or groups of branches each. In such conditions, even the most sophisticated algorithms that identify individual trees by processing imagery or LIDAR data are known to perform suboptimally and produce many tree omission or commission errors. Unlike deciduous species, the pyramid-shaped crowns of conifers produce comparatively fewer tree identification artifacts. Because the horizontal discrepancy between the location of the center of a tree stump, recorded by FIA, and its crown top, detected by processing LIDAR data, increases monotonically with crown size, the two dimensional distance between treetop and base for most red alder and maple trees exceeds that for Douglas fir. Hence, the plot point pattern extracted from LIDAR data approximates the plot point pattern computed from field data better for conifers than for deciduous species. Similar trends characterize the representation fidelity of the stand point pattern. Ultimately, the amount of noise present in the representation fidelity of the plot and stand point patterns determines whether the plot location can be retrieved and with what precision. Although a portion of the noise present in a plot point pattern originates in field measurement errors, its impact on retrieval precision was found to be minimal. Alternative world views, 20 for each plot, computed using intentionally introduced errors in tree height, stump distance, and azimuth, of magnitude comparable to that typically encountered in FIA data collection, produced results similar to those obtained by using the actual field measurements. This indicates that the accuracy and precision of field measurements is not a factor in the success or failure of the world-view method nor its plot retrieval precision.

Because the expected number of trees in a plot, and therefore the number of points in a plot pattern, is related linearly to the plot area, progressively smaller plot radii would produce more and more sparsely populated view vectors. Φ values computed from only a few non-zero Σv_i are affected by the stochasticity in the positional relationship between tree base and treetop much more than those computed in the presence of

numerous non-zero Σv_i's. Smaller plot radii cause a reduction in the specificity of a Φ value to a particular location within the stand. The current FIA plot design comprises four circular subplots of radius nearly half the one used in the study area plots. Although the total area between a standard FIA and a study plot is approximately the same, applying the method to each of the four subplots and identifying the plot center as the spatial average, for example the centroid, of the individually retrieved subplot centers will produce poor plot registration and a substantial increase in the number of method failures. A functional alternative merges all subplot views into a single V^T vector. Computing the merged plot view vector presents no difficulties; calculating the merged stand-view vector for each lattice vertex, however, is a computational challenge because it requires identifying the subset of the stand point pattern located within the four subplots centered on the lattice vertex processed at the time. On-going experimentation with data from a limited number of FIA plots present in the study area indicates that the alternative approach produces retrieval precisions equivalent to those mentioned in this study but with a five-fold increase in computation time. Spatial indexing of the point patterns is expected to reduce the computational intensity of the alternative method.

Conclusion

This study has demonstrated that the majority of plots established in stands scanned with LIDAR technology can be registered via the modified world-view method. The application of the method is expeditious and requires virtually no user input, and therefore no expertise in handling LIDAR data, or advanced computer resources. The occasional method failures or imprecise results occur strictly in stands that are exceedingly homogeneous. The strong spatial continuity and low stand attribute variability that characterizes such stands, however, ensures that information from plots registered poorly by the modified world view method likely will not degrade the output of models that overlay plot and other ancillary information. Registration of actual FIA plots is of fidelity equal to the one obtained by processing continuous, larger plots commonly used elsewhere, only it is computationally more intensive.

Literature Cited

Gatziolis, D. 2006. Evaluation of tree height measurement accuracy using FIA and LIDAR data. Proceedings, Society of American Foresters 2006 national convention. Bethesda, MD: Society of American Foresters. 19 p. [on CD-ROM].

Goodchild, M.F.; Parks, B.O.; Staeyert, L.T. eds. 1993. Environmental modeling with GIS. New York: Oxford University Press.

Heuvelink, G. 1998. Error propagation in environmental monitoring. London: Taylor and Francis. 127 p.

McRoberts, R.E. 2006. A model-based approach to estimating forest area. Remote Sensing of Environment. 103(1): 56-66.

Murtagh, F. 1992. A new approach to point pattern matching. Publications of the Astronomical Society of the Pacific. 104: 301-307.

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. Canadian Journal of Forest Research. 32: 725-741.

Schroeder, T.A.; Cohen, W.B.; Song, C.; Canty, M.J.; Zhiqiang, Y. 2006. Radiometric calibration of LANDSAT data for characterization of early successional forest patterns in Western Oregon. Remote Sensing of Environment. 103(1): 16-26.

Exploring the Association of the Minnesota Department of Natural Resources' Satellite-Detected Change With the Forest Inventory and Analysis System of Observed Removals and Mortality

Dale D. Gormanson[1], Timothy J. Aunan[2], Mark H. Hansen[3], and Michael Hoppus[3]

Abstract.—Since 2001, the Minnesota Department of Natural Resources (MN-DNR) has mapped forest change annually by comparison of Landsat satellite image pairs. Over the same timeframe, 1,761 U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis (FIA) plots in Minnesota have been remeasured on a 5-year cycle, providing field data on growth, removals, and mortality. This study compares estimates of change from these two sources. The FIA-based estimate of annual removals compares closely to the average MN-DNR estimate. FIA plots showing large removals were generally included in Landsat-detected change polygons, but image analysis usually failed to map as changed those plots exhibiting only partial removals or tree mortality.

Introduction

The Forestry Division of the Minnesota Department of Natural Resources (MN-DNR) employs two inventory systems. Forest managers use the map-based Cooperative Stand Assessment system, while strategic analysis relies on the Forest Service's plot-based Forest Inventory and Analysis (FIA) system. Between 1991 and 1999, the Division's Resource Assessment (RA) unit cooperated with the Forest Service's North Central and Rocky Mountain Research Stations in devising and testing the Annual Forest Inventory System (AFIS), a plan to transform the Federal FIA program in the Lake States from a periodic inventory conducted at 15-year intervals to a continuous inventory, with a proportion of plots examined every year (Hahn *et al.* 1992). The AFIS project prompted the adoption of annualized sampling nationwide by FIA. Remote sensing of forest change for the prioritization of plot visits was an integral part of the original AFIS design (Befort 2000), and RA was responsible for the design and implementation of remote sensing methods for AFIS.

Since the testing of the AFIS project, RA has continued to monitor and map changes in Minnesota forests by satellite image analysis, with the basic aim of compiling a continuous record of forest-cover disturbances on all ownerships (Aunan *et al.* 2006.). Ancillary objectives from year to year have included the detection of logging impacts on riparian zones, classification of disturbances by cause, and random targeting of individual harvest sites for field monitoring of forest practices. To date, five iterations have been conducted: 1999–2001, 2000–02, 2001–03, 2002–04, and 2003-–05. Meanwhile, FIA has been conducting annual inventories by plot remeasurement on a 5-year cycle. These plots provide ground-based observations of forest change (growth, removals, and mortality) that form the basis of FIA estimates. The purpose of the present inquiry was to compare satellite disturbance detection results against FIA's plot-based observations. Two questions were addressed:

1. Do satellite and FIA estimates of harvest acreage agree?
2. At site level, are satellite-detected changes being mapped at the locations where FIA plot data would lead us to expect them?

[1] Supervisory Forester, U.S. Department of Agriculture (USDA) Forest Service, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108.
[2] Assistant Supervisor, Minnesota Department of Natural Resources, Division of Forestry, Resource Assessment, Grand Rapids, MN 55744.
[3] Research Forester, USDA Forest Service, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108.

Data and Methods

MN-DNR's change monitoring effort has been detailed in a series of project reports (Aunan *et al.* 2004, 2006; Befort and Deegan 2002; Befort *et al.* 2003; and Deegan *et al.* 2005) from which the following synopsis of methodology is taken. Each year's project has followed the same general plan.

Imagery

The two Landsat satellites presently in service (Landsats 5 and 7) provide 30-meter 7-band multispectral Thematic Mapper (TM) images of Minnesota in five overlapping orbital paths, revisiting each path every 8 days in a sun-synchronous orbit.

To detect forest changes, a summer image from 2 years previous (Time 1) is matched against a current summer image (Time 2) at each of 19 Landsat scene locations covering Minnesota. MN-DNR purchases 10 new images from even-numbered Landsat orbital paths in even years and 9 from odd-numbered paths in odd years, thus obtaining 70 percent coverage of the State every year, a 2-year interval between image pairs, and a manageable analyst workload.

Image Preparation

Much extraneous variation must be filtered out before multispectral scanner scenes from different dates can be compared to detect particular types of vegetation change. Steps under the heading of "image preparation" are geared to ensure that detected changes represent actual alterations of ground reflectance rather than unrelated mismatches between images.

- Image preparation includes geometrically correcting and referencing images to the MN-DNR-standard NAD83 Universal Transverse Mercator extended Zone 15 projection. The Minnesota Department of Transportation statewide roads coverage serves as the accuracy standard. Original multispectral brightness values are converted to at-satellite reflectance. This radiometric calibration adjusts for differences in solar elevation, distance, and sensor differences over time between image pairs. Image preparation follows the procedures of Chandler and Markham (2003).
- Clouds and cloud shadows are detected and excluded from analysis by Normalized Difference Cloud Index techniques.
- The Gap Analysis Project vegetation map of Minnesota is used to "mask out" nonforest lands.

Change Detection Algorithm

Within cloud-free forested portions of the scenes, RA employs two straightforward image differencing algorithms for detecting vegetation changes between T1 and T2: a three-band difference using Landsat Bands 3 (visible red), 4 (near infrared), and 5 (first middle infrared)[4], and a two-band difference omitting Band 4:

Three-band = (T1 - T2, Band 3) + (T2 - T1, Band 4) + (T1 - T2, Band 5)

Two-band = (T1 -T2, Band 3) + (T1 -T2, Band 5).

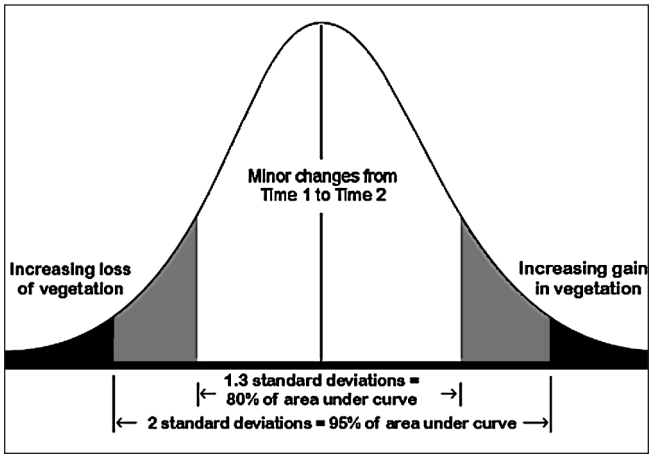
A change image is produced by differencing the values of corresponding pixels in each image for each band and then summing the results. The "change image" consists entirely of pixel-by-pixel difference scores. These scores usually display a frequency distribution that is bell shaped (i.e., nearly normal), with most values clustering around a mean of "no change" (fig. 1). Note in figure 1 that the arbitrary statistical margins depicted display the same level of "change" and "no change" no matter what has actually happened in the area analyzed. For this study, the focus was confined to the left side of the frequency distribution, representing vegetation losses.

Reconciliation and Analysis

In addition to satellite detection and mapping of forest disturbance, the MN-DNR uses an aerial photo sampling stage to identify the detected sites positively as harvests. In this application,

[4] *Three-band differencing*—Landsat Bands 3, 4, and 5 are useful for vegetation analysis. Band 3 has a wavelength of 0.63 to 0.69 μm and it has a nominal spectral location of red. Band 3 can detect chlorophyll, which aids in plant identification. Band 4 has a wavelength of 0.76 to 0.90 μm and its nominal spectral location is the near infrared. Band 4 is useful for interpreting different types of vegetation, detecting moisture in soils, and delineating water and land. Band 5 has a wavelength of 1.55 to 1.75 μm and its nominal spectral location is the mid-infrared. Band 5 can be used to detect the moisture content of various plants and soils.

Figure 1.—*Generalized distribution of picture elements in an idealized change image. For this study, interest was confined to the left side of the frequency distribution, representing vegetation losses. Courtesy Minnesota Department of Natural Resources.*



photos serve as a double sample to refine satellite-derived estimates of harvest acreage. Briefly,

- From the thousands of disturbances detected, 200 to 300 sites are randomly selected and photographed from the air.
- Disturbance acreages measured on the high-resolution aerial photos are used to adjust acreage estimates made from the coarse-resolution satellite data.
- Computer and visual analysis is used to distinguish forest harvests from other classes of forest land use change; this process includes thresholding the image margin to identify areas of high change.

MN-DNR Statewide Harvest Acreage Estimation

To estimate statewide rate of harvest, satellite-detected removals are first annualized: their acreage is divided by the years separating the two images from which they had been detected. The similarly adjusted photo-measured acreage of each double-sampled harvest site is then regressed on annualized satellite-estimated harvest acreage, and total annual harvest acreage is calculated from the regression relationship. The adjusted acreage is converted to an annual basis and expanded to include the remaining 30 percent of the State and becomes the MN-DNR statewide timber removal estimate.

FIA Plot Data

FIA field plot data include plot- and tree-level observations. Trees 1.0 to 4.9 in in diameter at breast height (d.b.h.) (4.5 ft) are measured on 1/300-acre microplots, and trees 5.0 in d.b.h. and larger are measured on four plots, each of which is 1/24 acre. Plot locations are monumented using Global Positioning System (GPS) technology. In order to facilitate tree relocation, field crews identify and map trees by polar coordinates (i.e., bearing and distance from the plot center to each tree). Crews identify which trees have died or have been removed since the previous inventory with a series of codes used to track tree history. As new trees grow into the microplots and subplots, they, too, are tracked until death. These tree histories along with measurements of tree diameter and other characteristics provide plot-based observations of the total volume that was removed and/or lost to mortality over the 5-year period between plot measurements.

FIA plot-level removal and mortality data for trees 5 in d.b.h. or greater were obtained from remeasurement plots initially observed during the first 2 years of FIA's 5-year annual inventory in Minnesota (1999–2003). In 2004, 577 forest plots initially observed in 1999 were remeasured and, in 2005, an additional 1,184 forest plots from 2000 were remeasured. Twice as many plots were observed in 2005 because of intensified field sampling made possible through the cooperation and assistance of MN-DNR.

Linking MN-DNR Harvest Polygons and FIA Plots

The mapped satellite-detected change harvest data set contained 44,964 forest harvest polygons derived from Landsat image pairs taken 2 years apart (1999–2001, 2000–02, 2001–03, 2002–04, and 2003–05). Years 1999–01 had a 5-acre minimum mapping unit. For the other periods, the detection threshold for forest removals was 2 acres. The largest polygon in the data set was approximately 250 acres. FIA plots observed with removals and FIA plots observed with mortality were linked spatially by GPS plot coordinates to the mapped satellite-detected harvest polygons they were closest to or contained in.

FIA Average Annual Removals

FIA timber removal and mortality estimates are typically provided on a volume basis following procedures explained in Bechtold *et al.* (2005). Estimates of acres harvested have not been typically reported in standard FIA reports. Such estimates require classification of conditions measured on an FIA plot as either harvested or not harvested. The MN-DNR classification is intended to identify areas that are clearcut or had major harvesting activities, not areas where a few scattered trees had been removed, such as a thinning or partial cut. It would be ideal if the Landsat Thematic Mapper could identify such partial harvests; however, these types of disturbances are beyond the capability of the sensor. To estimate annual removals, it was necessary to classify FIA conditions as clearcut based on the condition observed at T2. The estimated acres clearcut per year for the intersurvey period 1999/2000 to 2004/2005 are based on an estimate obtained from trees measured in the initial survey and cut or otherwise removed from the timberland base using the following criteria:

- A starting volume of 560 cubic ft per acre (all live trees).
- At least 75 percent of initial volume was observed cut.

Out of the 1,761 forest plots remeasured in this study, 68 plots contained conditions that met these criteria. In addition to theses 68 plots, another 138 plots had conditions that did not meet the "clearcut" rule but had some observed tree removals.

Results

FIA and MN-DNR Annual Removal Estimates Compared

The MN-DNR yearly average harvest estimate was 140,121 acres. This figure is based on estimates reported in Aunan *et al.* (2004) (112,000 acres), Aunan *et al.* (2006) (160,180 acres), Befort and Deegan (2002) (157,212 acres), Befort *et al.* (2003) (133,082 acres), and Deegan *et al.* (2005) (138,133 acres).

The FIA per-year removal estimate for the intersurvey period between 1999/2000 and 2004/2005 was 142,534 acres, with an average of 1673 cubic ft per acre cut and a range of between 532 and 6,380 cubic ft per acre cut. The sampling error on the FIA estimate is 11.8 percent, indicating no statistically significant difference between the two independent removal estimates.

MN-DNR Harvest Polygons and FIA Plot Intersections

Removals. Figure 2 shows boxplots and histograms[5] of FIA plots with observed tree removals. Out of the 1,761 plots remeasured, 206 were observed by field crews to have trees removed from the plot during the remeasurement period by harvesting, cultural operations such as timber stand improvement, land clearing, or changes in land use. The top boxplot and histogram to the immediate right were derived from plots that landed in harvest polygons (67 FIA plots).

The bottom boxplot and histogram to the immediate right were derived from plots that did not fall into harvest polygons (139 FIA plots). For the 67 plots that landed in harvest polygons, a statistically significant difference occurs in the mean plot-level volume (218 cubic ft) when compared to the mean plot-level volume of the 139 plots that do not occupy harvest polygons (129 cubic ft).

Figure 3 is a boxplot and histogram of plots with tree removals in relation to the nearest harvest polygons that conceivably might but did not contain them. Thirty of the 139 plots were located within 100 m of a harvest polygon. The average distance between mapped plots with tree removals that did not occupy satellite-detected harvest polygons was 1,238 m.

Mortality. Boxplots and histograms of FIA plots with observed tree mortality are shown in figure 4. The top boxplot and histogram to the immediate right were derived from plots that landed in harvest polygons (49 FIA plots). The bottom boxplot and histogram to the immediate right were derived from plots that did not fall into harvest polygons (837 FIA plots). Only about 5 percent of all plots observed with tree mortality land

[5] Each boxplot gives an idea of the spread (i.e., the data's symmetry and skewness at a glance). The box itself contains 50 percent of the data. The upper edge (hinge) of the box indicates the 75th percentile. The lower hinge indicates the 25th percentile. The range of the middle two quartiles is the interquartile range. The line in the box indicates the median value of the data. The ends of the vertical lines (whiskers) indicate minimum and maximum data values unless outliers are present, in which case whiskers extend to a maximum of 1.5 times the interquartile range. The histogram bins show the plot distribution frequency. The number of plots (n), mean (x bar), and standard deviation (s) for each data set are also shown.

Figure 2.—*Boxplots and histograms of Forest Inventory and Analysis removal plots intersected with Minnesota Department of Natural Resources harvest polygons. The top boxplot and histogram to the immediate right were derived from plots with observed tree removals that landed in harvest polygons. The bottom boxplot and histogram to the immediate right were derived from plots with observed tree removals that did not fall into harvest polygons.*

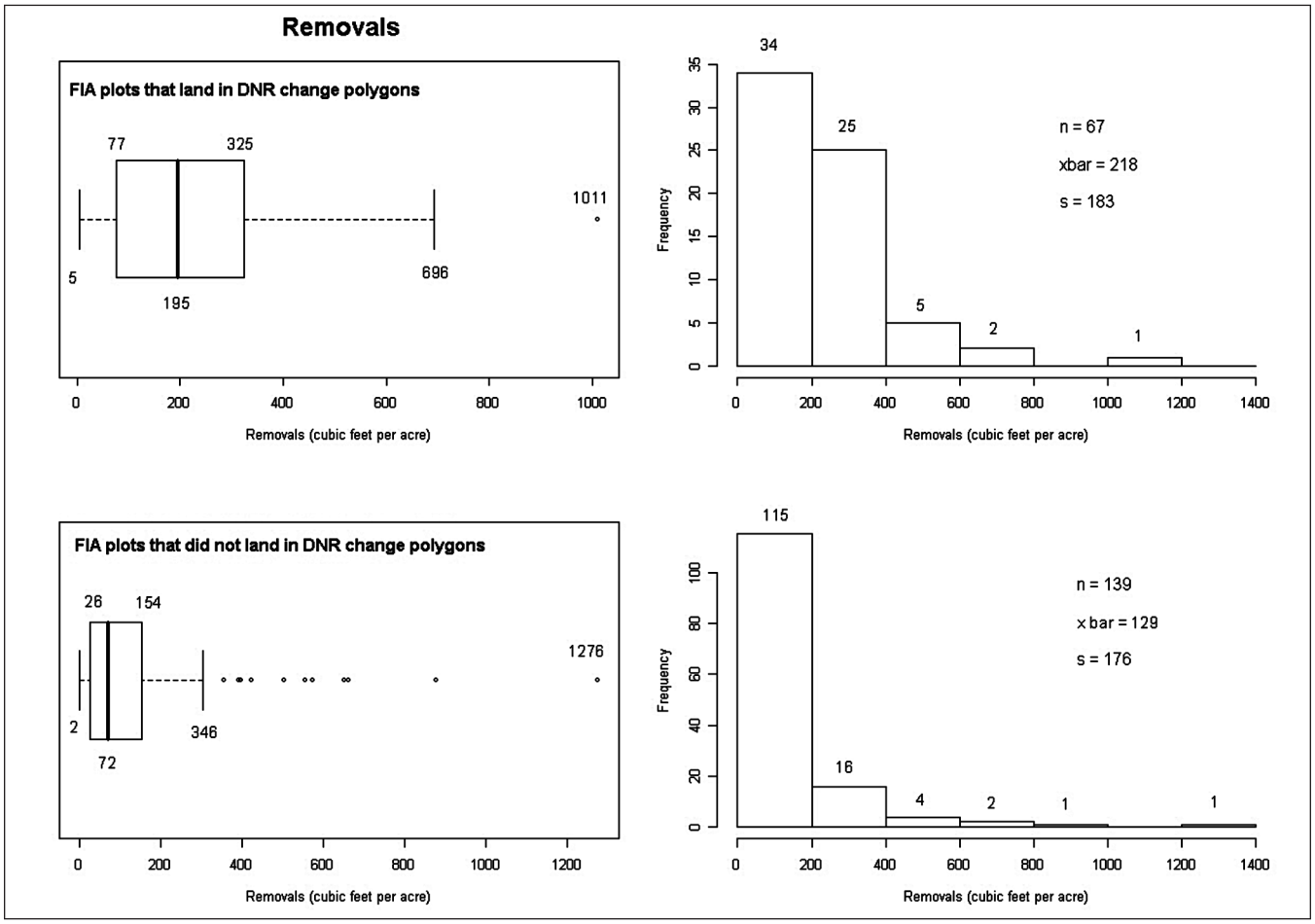


Figure 3.—*Proximity boxplot and histogram of Forest Inventory and Analysis removal plots that did not land in satellite-detected harvest polygons.*

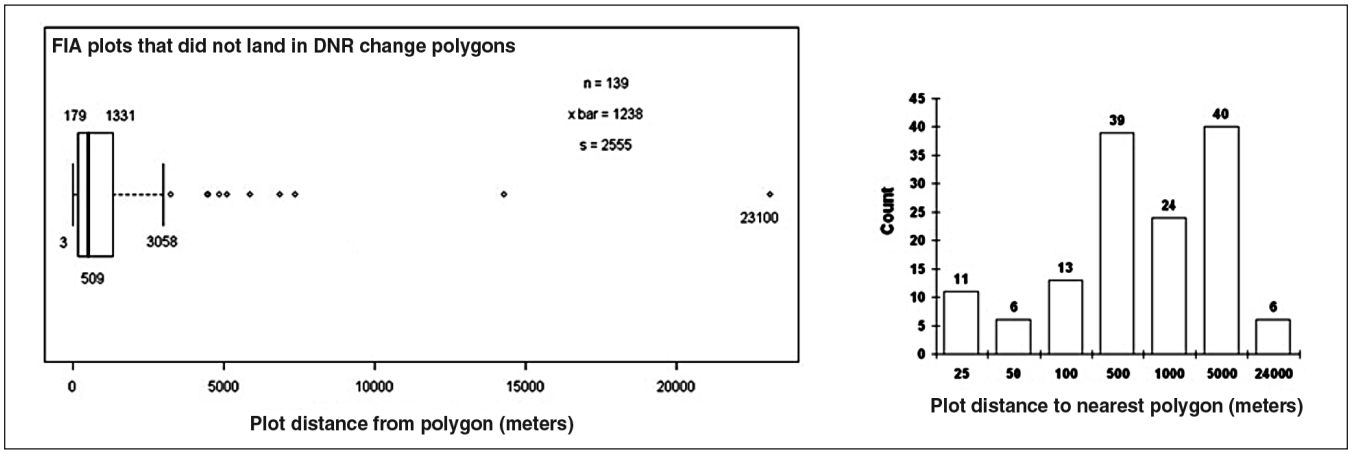
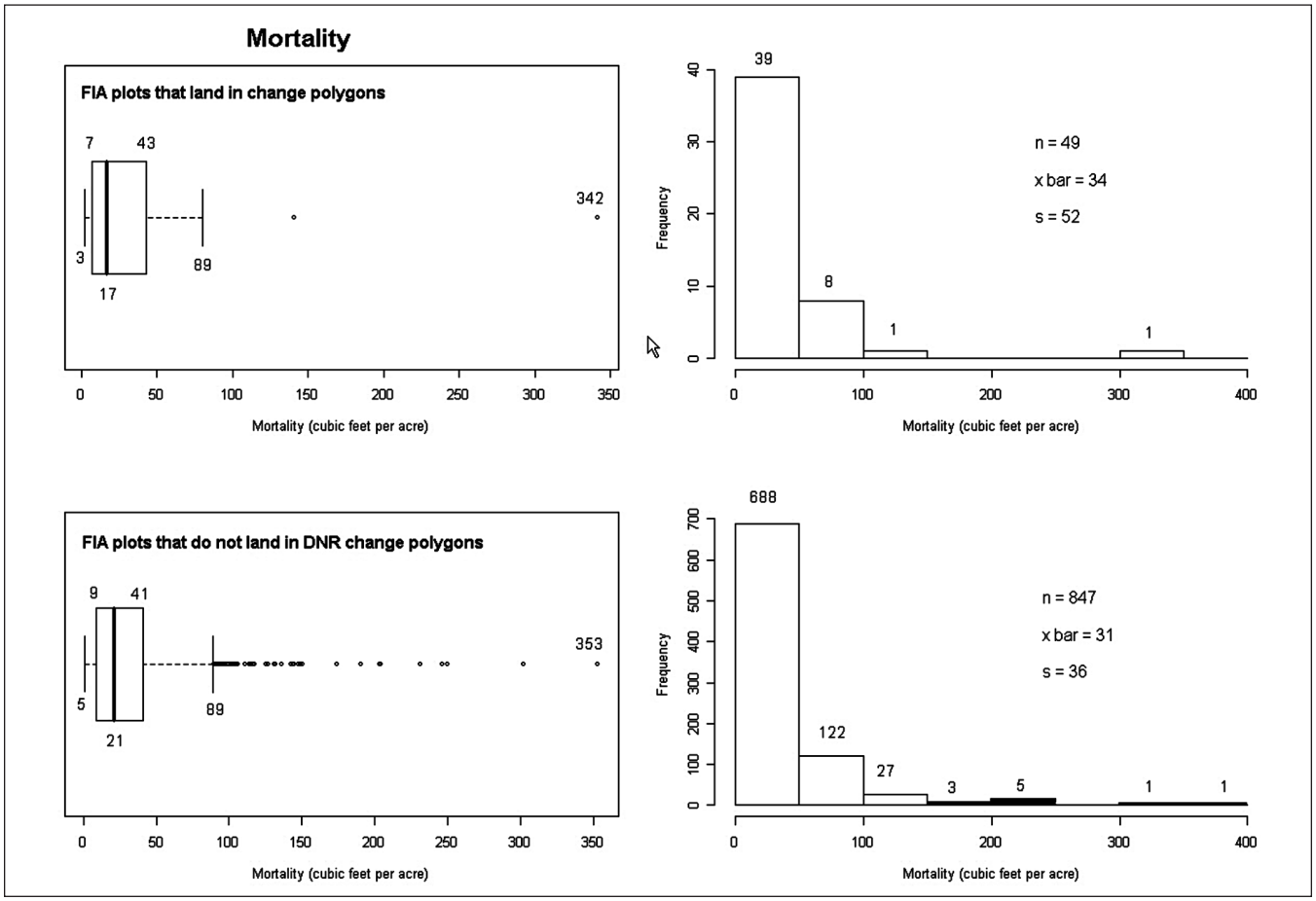


Figure 4.—*Boxplots and histograms of Forest Inventory and Analysis mortality. The top boxplot and histogram to the immediate right were derived from plots with observed tree mortality that landed in harvest polygons. The bottom boxplot and histogram to the immediate right were derived from plots with observed tree mortality that did not fall into satellite-detected harvest polygons.*



in harvest polygons. For the 49 plots that landed in harvest polygons, a statistically significant difference did not occur in the mean plot-level volume (34 cubic ft) when compared to the mean plot-level volume of the 847 plots that do not land in harvest polygons (31 cubic ft).

Discussion

The FIA-based estimate of annual removals compares closely to the average annual MN-DNR estimate. Analysis of site-specific change suggests FIA plots with observed high-volume removals tend to be located in harvest polygons. Plots observed with tree mortality tend not to land in harvest polygons satisfying the RA harvest mapping objective. Partial removals, however, are not typically mapped as change in satellite-mapped harvest polygons. Plot-polygon proximity does not seem to be a reason, but it cannot be ruled out. The following factors may help explain why FIA plots with observed partial tree removals fail to get mapped inside Landsat-detected harvest polygons:

- Temporal differences between FIA field observations and image acquisition date(s). Change detection work over time requires close correspondence between data sets representing T1 and T2. Some FIA field plots were probably observed after image acquisition and vice versa.
- Resolution differences. The FIA field plot is spread out over approximately 1 acre compared with the minimum 2-acre mapping unit of MN-DNR polygons. Befort *et al.*

(2002) notes that a 2-acre harvest minimum might be expected to press the limits of possibility in Landsat-based disturbance detection. A Landsat 5 Thematic Mapper or Landsat 7 Enhanced Thematic Mapper multispectral scene consists of a rectangular array of approximately 6,000 x 6,000 picture elements (pixels) covering an area of 180 x 180 km (110 x 110 mi). Each pixel measures 30 x 30 m in ground dimensions—about 100 x 100 ft, or roughly ¼ acre. A 2-acre harvest may thus involve only 8 of the 36 million pixels in the scene. As reported in Befort *et al.* (2002), "This is near the level of random 'noise' likely to arise in any between-date satellite image comparison through geometric misregistration, atmospheric interference, and other inexactitudes."

- Practically speaking, nothing in nature is permanent except change. Forest ecosystems are dynamic. In particular, deciduous vegetation undergoes dramatic seasonal changes. Each satellite image imparts a unique collection of illumination, atmospheric conditions, and canopy cover that exists once and never recurs exactly the same way twice. The purpose of the change algorithm is to filter out, using clues provided by reflected light, unique types of change that are discernible. In many cases, change is detected and validated, but some uncertainty at the margin threshold is inevitable (Aunan *et al.* 2006).
- No two instruments are identical and they deteriorate over time. In 2003 and 2004, RA employed a mix of data from Landsat 5, which has been in continuous service since 1984, and Landsat 7, which was launched 15 years later in 1999. Since being placed in orbit, Landsat 7 has developed a mechanical fault with the scan line corrector that affects sensor performance. Although standard normalization routines are applied, it is unlikely that any calibration can completely remove all performance differences between both sensors (Aunan *et al.* 2006).
- Satellite forest removal detection is complicated by unrelated forest stressors. For example, Miles *et al.* (2005) reports that, over the 5-year period from 1999 to 2003, millions of acres of Minnesota's northern boreal forests were defoliated in early summer by the forest tent caterpillar (*Malacosoma disstria* Hubner). Other problematic defoliators during the time period include jack pine and spruce budworm. Such defoliation produces reflectance effects that mimic partial removals. In order to avoid confusing defoliation with removals, RA makes every attempt to obtain late-summer imagery taken after refoliation, but refoliation is sometimes late and therefore the timing of image acquisition is imperfect. The imagery used throughout this period probably incorporates defoliation effects and contributes to mismatch between FIA removal plots and MN-DNR harvest polygons.
- Finally, forest practices change. Clearcutting is the prevalent logging method used in Minnesota; however, partial cuts form an increasing fraction of harvests. The satellite image differencing method detects many partial cuts, but, as noted, a threshold exists at which a change in canopy density ceases to be obvious. We cannot precisely map that limit, because in marginal cases the other dynamics pointed out may lead to sentinel-site–specific change mapping uncertainty.

Conclusions

Three conclusions can be gleaned from this study:

1. FIA and MN-DNR statewide area estimates of annual clearcutting are fundamentally the same when a threshold of 560 cubic ft per acre and 75 percent of initial volume is applied to FIA plot data.
2. FIA plots with observed partial removals and mortality are not typically contained in one of the satellite-detected harvest polygons. FIA plots have the resolution to show change that cannot be detected by Landsat satellites unless a significant change in canopy cover occurs.
3. There appears to be a relationship between Landsat satellite-detected clearcuts (polygons) and FIA plots with large observed removals.

RA combines satellite imagery and aerial photography as an effective means of estimating annual harvest in Minnesota forests. The satellite imagery provides not only an estimate of change but also a moderately high-resolution map that has value in evaluating forest loss using other high-resolution data. The satellite-based change map has great potential as an ancillary data layer for increasing the precision of FIA removal estimates.

Literature Cited

Aunan, T.J.; Deegan, G.; Befort, W.A. 2004. Satellite-based forest change monitoring: 2003 report. Resource Assessment. Grand Rapids, MN: Minnesota Department of Natural Resources, Division of Forestry. 9 p.

Aunan, T.J.; Deegan, G.; Rack, J. 2006. Minnesota forest change monitoring: 2003-2005 report. Resource Assessment. Grand Rapids, MN: Minnesota Department of Natural Resources, Division of Forestry. 9 p.

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.

Befort, W.A. 2000. Change detection, stratification and mapping for continuous forest inventory. Proceedings, 1999 section on statistics and environment, joint statistical meetings. Alexandria, VA: American Statistical Association: 7-12.

Befort, W.A.; Aunan, T.J.; Deegan, G. 2003. Satellite-based forest harvest monitoring: 2002 report. Resource Assessment. Grand Rapids, MN: Minnesota Department of Natural Resources, Division of Forestry. 12 p.

Befort, W.A.; Deegan, G. 2002. Change analysis applications in DNR resource monitoring. Resource Assessment. Grand Rapids, MN: Minnesota Department of Natural Resources, Division of Forestry. 16 p.

Chandler, G.; Markham, B. 2003. Revised Landsat 5 TM radiometric calibration procedures and post-calibration dynamic ranges. U.S. Geological Survey online document. http://landsat.usgs.gov/documents/L5TMCal2003.pdf. [Date accessed unknown].

Deegan, G.; Aunan, T.J.; Befort, W.A. 2005. Minnesota forest change monitoring: 2004 report. Resource Assessment. Grand Rapids, MN: Minnesota Department of Natural Resources, Division of Forestry. 10 p.

Hahn, J.T.; McRoberts, R.E.; Befort, W. 1992. Annual forest inventory system (AFIS): integrating data base techniques, satellite imagery, annual designed sampling and modeling. Proceedings, Conf. Integrated Forest Resource Information, Space, Time. Canberra, Australia: 314-324.

Miles, P.D.; Brand, G.J.; Mielke, M.E. 2005. Minnesota's forest resources in 2003. Resour. Bull. NC-246. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station. 36 p.

The Utility of the Cropland Data Layer for Forest Inventory and Analysis

Greg C. Liknes[1], Mark D. Nelson[2], Dale D. Gormanson[3], and Mark Hansen[2]

Abstract.—The Forest Service, U.S. Department of Agriculture's (USDA's) Northern Research Station Forest Inventory and Analysis program (NRS-FIA) uses digital land cover products derived from remotely sensed imagery, such as the National Land Cover Dataset (NLCD), for the purpose of variance reduction via postsampling stratification. The update cycle of the NLCD product is infrequent; NLCD 2001 was the first update since the release of NLCD 1992, and was not yet fully completed as of late 2006. Consequently, FIA field data collected as recently as 2005 are being poststratified with land cover data collected more than a decade before. In addition, NRS-FIA has performed its own land cover classification of remotely sensed imagery for use in the stratification of some States, a time-consuming process. Alternative sources of information need to be evaluated both to eliminate the temporal mismatch between land cover data and FIA plot information and to reduce the amount of analyst time required to perform the stratification process.

The USDA's National Agricultural Statistics Service, in conjunction with the Foreign Agricultural Service, produces the Cropland Data Layer (CDL) using satellite imagery. The product, updated yearly for multiple States, includes detailed classification of crop type and also a "Woodland" category. In this study, the CDL was compared to the NLCD 1992 data set for Wisconsin. This comparison included two components: (1) county-level, pixel-derived estimates of forest land area relative to each other and to plot-based FIA estimates and (2) variance reduction produced by each layer when used in postsampling stratification of FIA plots for estimating primary FIA attributes (forest area, number of trees, timber volume, and tree biomass). Results indicate poor agreement between CDL pixel-based area estimates and FIA plot-based estimates; however, when used for poststratification, the CDL produces similar estimates of primary FIA attributes to those of the NLCD at the State level and higher relative efficiency at the FIA survey-unit level.

Introduction

The Forest Service's Forest Inventory and Analysis (FIA) program makes estimates of the area of forest land across all 50 States and Puerto Rico. In addition, the volume, growth, removal, and health of forest resources are assessed. Estimates are achieved via data collected on a consistent, nationwide sampling framework (Bechtold and Patterson 2005). The FIA program at the Northern Research Station (NRS-FIA) uses digital land cover products derived from remotely sensed imagery, such as the National Land Cover Dataset (NLCD), for the purpose of reducing the variance of estimates via postsampling stratification. The update cycle of the NLCD product is infrequent. NLCD 2001 (Homer *et al.* 2004) was the first update since the release of NLCD 1992 (Vogelman *et al.* 2001), and was not yet fully completed as of late 2006. Consequently, FIA field data collected as recently as 2005 are being poststratified with land cover data collected more than a decade before. In addition, NRS-FIA has performed its own land cover classification of remotely sensed imagery for use in the stratification of some States, a time-consuming process. Alternative sources of information need to be evaluated both to eliminate the temporal

[1] Research Physical Scientist, U.S. Department of Agriculture (USDA) Forest Service, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108.
[2] Research Forester, USDA Forest Service, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108.
[3] Supervisory Forester, USDA Forest Service, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108.

mismatch between land cover data and FIA plot information and to reduce the amount of analyst time required to perform the stratification process.

The goal of this study was to compare another consistent land cover data set to NLCD and to evaluate the utility of a data product from another U.S. Department of Agriculture (USDA) agency for FIA. The Cropland Data Layer (CDL) (Craig 2001) was compared to the NLCD 1992 data set for Wisconsin. This comparison included two components (1) county-level, pixel-derived estimates of forest land area relative to each other and to plot-based FIA estimates and (2) variance reduction produced by each layer when used in postsampling stratification of FIA plots for estimating primary FIA attributes (forest area, number of trees, timber volume, and tree biomass).

Data

National Land Cover Dataset

The NLCD, a digital product of the Multi-Resolution Land Characterization Consortium (Loveland and Shaw 1996), is a land cover map of the conterminous United States consisting of the assignment of each 30-m x 30-m pixel to 1 of 21 land cover classes, 4 of which are assumed to be sufficiently similar to FIA's definition of forest to be considered forest for this study. The land cover classification was produced by the U.S. Geological Survey and was based on nominal 1992 Landsat Thematic Mapper (TM) satellite imagery and a variety of ancillary data. NLCD 1992 data are freely available for each of the 48 conterminous United States (Vogelman *et al.* 2001).

Cropland Data Layer

The CDL (Craig 2001, Mueller and Ozga 2002) is a digital product produced as part of a cooperative venture between two USDA agencies, the National Agricultural Statistics Service and the Foreign Agricultural Service. The product, updated yearly for multiple States, includes detailed classification of crop type and also a broad "Woodland" category, which includes woods and wooded pastures. The primary purpose of the CDL is to monitor cropland and other agricultural lands; the identification of woodlands is a side product of the classification effort. The layer used in this study is the Wisconsin 2005 CDL, which is based on multiple dates of Landsat TM imagery from that same year. As with the NLCD, CDL pixels are 30 m x 30 m.

FIA Plot Data

For this study, measurements taken during the first FIA annual inventory in Wisconsin (2000–04) were used. Plot-level data on forest area, number of trees, timber volume, and tree biomass were obtained for 12,885 plots. Of these plots, 6,478 were partially or completely forested and 6,407 were not forested. County-level land area estimates based on plot data were obtained from the FIA Mapmaker program (Miles 2001).

Methods

NLCD/CDL Processing

The 21 land cover classes of the NLCD 1992 were collapsed into forest and nonforest categories, and a raster filtering process was used to remove forest/nonforest pixel groupings smaller than 1 acre. This filtering process was done to conform to the FIA definition of forest land, which requires that groups of trees occupy a minimum of 1 acre. Three forest classes (deciduous forest, evergreen forest, and mixed forest); shrublands; woody wetlands; and lands in transition to forest were included in the aggregate forest class. Next, an edge category, 2 pixels in width, was created along both sides of the forest/nonforest interface. This procedure resulted in a total of four categories, or strata: forest, forest edge, nonforest, and nonforest edge.

CDL categories were also collapsed into forest and nonforest classes and filtered to remove groupings smaller than 1 acre. The Woodland category was the sole member of the forest stratum; all other categories became nonforest, with the exception of clouds. Cloud pixels accounted for less than 0.1 percent of the total CDL area and were recoded to either forest or nonforest class by applying a series of successively finer majority filters to assign unclassified cloud pixels the majority land cover value surrounding them. As with the NLCD, 2-pixel edge strata were created, resulting in a total of four strata. No CDL data were available for small portions of Wisconsin (e.g., a few islands in Lake Michigan beyond the Door County peninsula), and therefore the NLCD was clipped to the smaller extent of

the CDL to create equivalent extents for each stratification layer.

Pixel-Based Area Estimation

Estimates of forest land area were compared between FIA and CDL, FIA and NLCD, and CDL and NLCD data. County estimates were compared using percent differences and area-weighted root mean square deviation (RMSD). Percent differences for each pair were defined by equations 1, 2, and 3.

$$\frac{\left(CDL - FIA\right)}{FIA} \quad (1)$$

$$\frac{\left(NLCD - FIA\right)}{FIA} \quad (2)$$

$$\frac{\left(CDL - NLCD\right)}{NLCD} \quad (3)$$

where CDL and NLCD are the pixel-based estimates of forest area and FIA is the plot-based estimate of forest area. The RMSD was modified from Häme *et al.* (2001), equation (4).

$$RMSD_{rv} = \sqrt{\sum_i \frac{a_i}{A}\left(\hat{p}_{ir} - \hat{p}_{iv}\right)^2} \quad (4)$$

where a_i is the area of the ith county, $A = \sum a_i$ is the total Wisconsin area, and $\hat{p}_{ir}$ and $\hat{p}_{iv}$ denote the estimated proportion of forest land area in the ith county obtained from FIA, CDL, or NLCD (r) and FIA, CDL, or NLCD (v), respectively, for each of three pair-wise comparisons of estimates.

Stratified Estimation

The purpose of stratified estimation is to reduce the variance of sample-based estimates without increasing sample size. For this study, each FIA plot was assigned to one of four strata for both the NLCD and CDL corresponding to the pixel in which the center of the plot was located. Stratum proportions were estimated by counting the number of pixels in each stratum and dividing by the total number of pixels in the study area; these proportions are used as stratum weights when calculating stratified estimates of means and variances. In addition, relative efficiency (RE) was calculated by determining the variance of an estimate under the assumption of simple random sampling and dividing by the variance using stratified estimation. All estimation followed the standard FIA estimation procedures presented by Scott *et al.* (2005). Estimates, variances, standard errors, and RE for forest area, number of trees, timber volume, and tree biomass were calculated for the entire State as well as for FIA survey units (fig. 1).

Figure 1.—*FIA survey units for the State of Wisconsin.*



Results

Pixel-Based Area Estimation

Estimates of Wisconsin forest land area were 16.0, 13.0, and 16.9 million acres for FIA, CDL, and NLCD, respectively. On a per-county basis, estimates differed by -63 to 74 percent for FIA and CDL, -48 to 289 percent for FIA and NLCD, and -75 to 54 percent for CDL and NLCD (fig. 2). Wisconsin county comparisons of the three forest land area estimates resulted in area-weighted RMSD values of 8.2 percent for FIA compared with CDL, 4.1 percent for FIA compared with NLCD, and 9.6 percent for CDL compared with NLCD.

Stratified Estimation

Comparing FIA's plot-based estimate using NLCD and CDL for poststratification, the results for number of trees, forest

area, biomass, and volume are all comparable and within one standard error of one another (fig. 3). Comparing the RE produced by the two different stratification layers, results are consistent across all four variables and all five FIA survey units (table 1). Specifically, CDL resulted in higher RE for number of trees, forest area, biomass, and volume in all units, with one exception. The RE for the forest area estimate was substantially higher when using NLCD compared with CDL in unit 2 (fig. 4).

Figure 2.—*Percent difference in estimates of forest land area between Forest Inventory and Analysis and Cropland Data Layer, Wisconsin. Each dot represents a city with population greater than 50,000.*

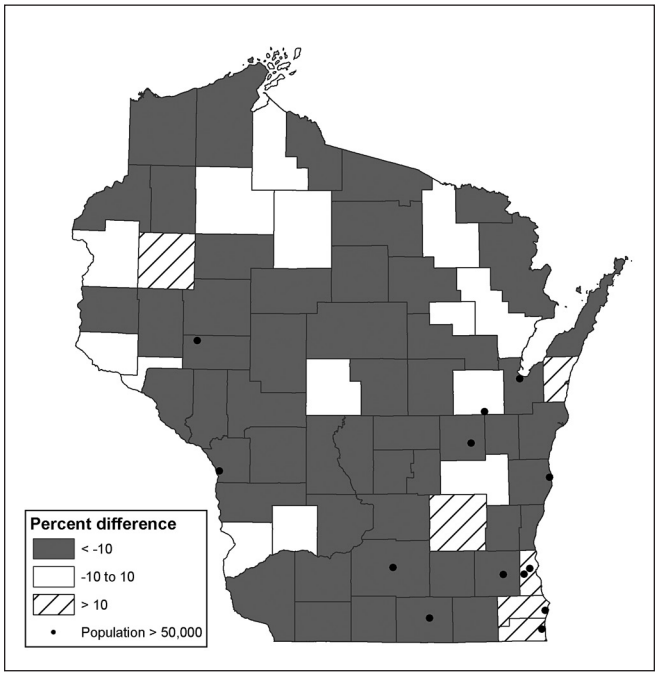


Table 1.—*The relative efficiency of stratified estimates of number of trees, forest area, biomass, and volume in Wisconsin using the 2005 CDL and the 1992 NLCD. The first number in each pair (**bold**) corresponds to the estimate generated using the CDL and the second number corresponds to the estimate generated using the NLCD.*

FIA survey unit (WI)	Relative Efficiency (**CDL**/NLCD)			
	Number of trees	Forest area	Biomass	Volume
1	**1.28**/1.22	**2.46**/2.40	**1.45**/1.36	**1.33**/1.28
2	**1.23**/1.21	**1.59**/2.17	**1.42**/1.34	**1.30**/1.27
3	**1.52**/1.36	**2.31**/2.14	**1.62**/1.44	**1.53**/1.38
4	**1.60**/1.53	**2.48**/2.30	**1.77**/1.68	**1.65**/1.55
5	**1.41**/1.30	**1.96**/1.66	**1.67**/1.47	**1.67**/1.47

CDL = Cropland Data Layer. FIA = Forest Inventory and Analysis. NLCD = National Land Cover Dataset. WI = Wisconsin.

Figure 3.—*Stratified estimates of the number of trees, forest area, biomass, and volume for Wisconsin using both the National Land Cover Dataset and Cropland Data Layer for stratum assignment of Forest Inventory and Analysis plots and weighting of strata.*

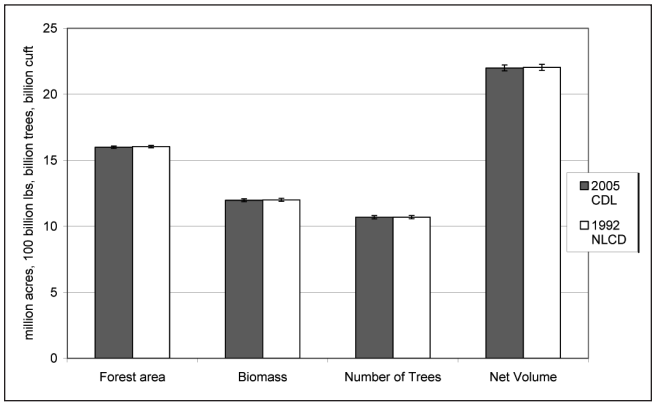
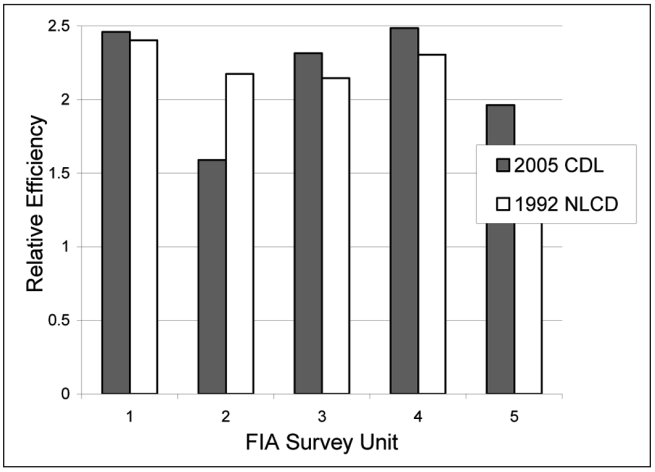


Figure 4.—*The relative efficiency of the stratified estimate of forest area for all Forest Inventory and Analysis survey units in Wisconsin using both the National Land Cover Dataset and Cropland Data Layer stratification layers.*



Conclusions

Compared with the FIA estimate of Wisconsin forest land area, the CDL estimate was 18.7 percent smaller and the NLCD estimate was 5.2 percent larger. Differences in county estimates were much larger, but area-weighted RMSD was only 8.2 percent for FIA compared with CDL and only 4.1 percent for FIA compared with NLCD. The greatest differences appeared in counties with sparse forest cover, especially those counties containing urban tree cover that does not meet FIA's definition of forest land use. Additional geospatial processing steps could separate urban tree cover from other forest land, accounting for these definitional differences.

Although, when compared to NLCD, CDL's pixel-based estimate of forest area is substantially different from the FIA estimate, each land cover data set results in statistically equivalent Wisconsin estimates of the four FIA variables when used as a stratifying layer in the sample-based estimation process. This observation lends support to the notion that FIA's estimation procedures result in an unbiased estimate, even if the stratification layer used is biased. Furthermore, because the CDL outperformed NLCD with respect to RE for the four variables, perhaps we can conclude that agreement in pixel-based forest area with sample-based estimates is not a controlling factor in reducing variance.

Why did the CDL outperform NLCD with respect to variance reduction in most cases? Because of its focus on agricultural lands, the CDL represents nonforest land with good accuracy, resulting in an accurate forest/nonforest interface. We speculate the edge classes created along this interface are more likely to contain highly variable FIA plots compared with the NLCD-derived edge strata. Confining highly variable plots to edge strata, which occupy relatively little area, results in a net reduction in the variance of estimates of primary FIA variables.

In summary, CDL appears to be a viable option for FIA to use in stratified estimation, with several cautions. First, comparisons were only made for data in a single State. Additional work should be done to determine the consistency of CDL across States. Also, because CDL is updated each year, clouds would likely be an issue, and a more robust procedure than was used in this study would have to be implemented to fill this data gap. Finally, if the FIA program is aiming for national consistency, CDL presents a problem in that data is not available for every State. An intradepartmental partnership with the goal of expanding CDL to all States would be necessary to conform to an FIA objective of national consistency.

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.

Craig, M.E. 2001. The NASS cropland data layer program. In: Proceedings, third international conference on geospatial information in agriculture and forestry. Ann Arbor, MI: Veridian. [On CD-ROM].

Häme, T.; Stenberg, P.; Andersson, K.; Rauste, Y.; Kennedy, P.; Folving, S.; Sarkeala, J. 2001. AVHRR-based forest proportion map of the Pan-European area. Remote Sensing of Environment. 77: 76-91.

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 & Remote Sensing. 70: 829-840.

Loveland, T.L.; Shaw, D.M. 1996. Multiresolution land characterization: building collaborative partnerships. In: Scott, J.M.; Tear, T.; Davis, F., eds. Gap analysis: a landscape approach to biodiversity planning. Proceedings, American Society for Photogrammetry and Remote Sensing/Gap Analysis Program symposium. Moscow, ID: National Biological Service: 83-89.

Miles, P.D. 2001. Forest inventory mapmaker users guide. Gen. Tech. Rep. NC-221. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station. http://www.ncrs2.fs.fed.us/4801/fiadb/available_data.html. (6 November 2006).

Mueller, R.; Ozga, M. 2002. Creating a cropland data layer for an entire state. Proceedings, 2002 American congress on surveying and mapping/American Society for Photogrammetry and Remote Sensing conference and technology exhibition. Bethesda, MD: American Society for Photogrammetry and Remote Sensing. [On CD-ROM].

Scott, C.T.; Bechtold, W.A.; Reams, G.A.; Smith, W.D.; Westfall, J.A.; Hansen, M.H.; Moisen, G.G. 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. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station: 43-67.

Vogelman, J.E.; Howard, S.M.; Yang, L.; Laron, C.R.; Wylie, B.K.; Van Driel, N. 2001. Completion of the 1990's National Land Cover Data Set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. Photogrammetric Engineering & Remote Sensing. 67: 650-662.

Landscape Scale Mapping of Forest Inventory Data by Nearest Neighbor Classification

Andrew Lister[1]

Abstract.—One of the goals of the Forest Service, U.S. Department of Agriculture's Forest Inventory and Analysis (FIA) program is large-area mapping. FIA scientists have tried many methods in the past, including geostatistical methods, linear modeling, nonlinear modeling, and simple choropleth and dot maps. Mapping methods that require individual model-based maps to be produced are time and labor intensive. FIA needs a method that will enable efficient production of large numbers of landscape-scale maps for State reports. The current study presents a case study for the State of Ohio of a nearest neighbor classification method that uses multivariate similarity as a criterion for attaching FIA plot data to pixels with unknown forest attributes to make continuous maps of any FIA attribute. The goal of the study was to devise a landscape scale mapping method that could be easily implemented at a national scale.

Introduction

Landscape scale maps of forest attributes have been of interest in the United States for decades. Sargent (1884) produced what appear to be the first relatively detailed, national maps of forestry data, which included timber volume and species group distribution. During the early to mid-20th century, some national-scale and many finer scale forest vegetation mapping efforts were undertaken by Federal, State, and academic researchers (e.g., Braun 1950; EPA 1994; Little 1971, 1981; Shantz and Zon 1924). Most of these data sets were created by either manually or semimanually digitizing vegetation polygons from photos or from field reconnaissance. During the 1980s and early 1990s, however, advanced computers, such as Geographic Information Systems (GIS) and satellite imagery, became more widely available, leading to more sophisticated vegetation mapping efforts (e.g., Zhu 1994, Zhu and Evans 1994). During the 1990s and early 2000s, remote sensing technology, spatial modeling procedures, and statistical software led to further advances in mapping.

Satellite imagery distribution systems, along with the integration of statistical methods for image classification and spatial modeling with GIS, have led to numerous applications of the use of ground inventory data for mapping forest vegetation, as described by Fassnacht *et al.* (2006) and Andersen (1998). The Forest Service's Forest Inventory and Analysis (FIA) program has used its inventory plot data in conjunction with remotely sensed data for mapping for many years (e.g., Frescino *et al*. 2001, Lister *et al*. 2000, McRoberts *et al.* 2002, Moisen and Edwards 1999). Many of these and other techniques, such as linear modeling methods, are not suitable for production-level mapping because a separate model, and its associated overhead (disk storage, processing time, etc.), is generated for each map. A goal of the FIA program is to develop a production-level mapping procedure that efficiently uses staff, computing resources, and time in order to meet its mapping goals (USDA 1998).

Generally, FIA's mapping goals involve creating accurate maps depicting the spatial distribution and levels of forest resources across the landscape. These maps are often included in publications, on Web sites, and in presentations and are meant to support other data that show the quantity, distribution, and health of the Nation's forests. FIA currently uses maps produced for State reports more as graphics and less as GIS data sets. Interest is growing, however, in using FIA maps as geospatial data sets. For example, FIA-based maps were used by the Forest Service's Forest Health Protection program as ancillary inputs to create forest pest risk maps (Downing n.d.).

[1] Research Forester, U.S. Department of Agriculture, Forest Service, Northern Research Station, Northern Monitoring Program/Forest Inventory and Analysis Unit, 11 Campus Boulevard, Suite 200, Newtown Square, PA 19073. E-mail: alister@fs.fed.us.

The advent of a steady stream of imagery data from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor (Justice *et al.* 1998) has led to increased use of these data for land cover classification. For example, FIA data have recently been used in conjunction with MODIS data and other GIS layers in the creation of two national-scale maps: dry aboveground biomass (Moisen n.d.) and forest type group (Ruefenacht n.d.). These maps were made using classification and regression trees (Breiman *et al.* 1984) and required large amounts of time and effort to produce.

A more efficient alternative for FIA's future national mapping needs is an approach based on supervised classification. In supervised classification, representative data (here, MODIS imagery and other GIS predictor data linked spatially with the FIA plot information) are used as a reference set. Pixels with unknown values for FIA attributes are populated with "nearest neighbor" reference data based on multivariate similarity between vectors of predictor data at these known locations and those at unknown locations. FIA data have been used in this way before, but generally over smaller areas, with Landsat data or for fewer attributes (e.g., McRoberts *et al.* 2002). In the current study, I present a case study of a nearest-neighbor, supervised classification method that uses FIA data, MODIS imagery, and GIS layers to create landscape maps of the State of Ohio. My goals are to produce maps that will be used for the FIA report that describes Ohio's forests and to present a framework for a methodology that can be applied to other large-area mapping problems.

Methods

Data from 691 homogeneous (single-condition) forested FIA plots collected in Ohio (fig. 1a) between 2001 and 2006 were used in the study. I assume here that the nominal date of the plot information is 2001 and that only marginal changes occurred between the date of the imagery used in the study (2001) and the date the last plot was measured (2006). The distribution of plots in the study area is based on a hexagonal tessellation with one FIA plot randomly located within each 6,000 acre (2,428 ha) hexagon. Each FIA plot consists of four circular 48-ft (14.6-m) diameter subplots, with one subplot located in the center and three equidistant subplots distributed symmetrically around and located 120 ft (31.6 m) from the center subplot. The subplots occupy 0.17 acre (0.07 ha), and the subplot array can be subtended by a circle of 1.5 acre (0.6 ha) in area. FIA attributes summarized to the plot level include basal area per acre of red maple (BARM) found on the plots, cubic foot gross volume (CFGV) of trees greater than 5 in (12.7 cm) diameter at breast height, occurrence of mixed upland hardwoods forest type (MUH), and total dry aboveground biomass (TDRYBIO). For details on the FIA plot design, sample layout, and statistical analytical methods, see Bechtold and Patterson (2006).

The predictor data used were contained in a multilayered Erdas IMAGINE image and consisted of 271 250-m resolution layers, including multidate and monthly composites and derived indices of imagery from the MODIS satellite borne sensor (Justice *et al.* 1998), several rasterized summaries of the STATSGO soils database compiled by the Natural Resources Conservation Service (1994), summaries of the landcover classes found in the U.S. Geological Survey National Land Cover Dataset (NLCD) database (Vogelmann *et al.* 2001), mean monthly and annual temperature and precipitation from the PRISM climate database (Daly *et al.* 2004), a rasterized grid representing distance to streams (USGS 1999), and various derivatives of the National Elevation Dataset (Gesch *et al.* 2002). Complete details of the steps used to prepare the data and data derivatives are on file at the Forest Service's Northern Research Station (11 Campus Blvd, Ste. 200, Newtown Square, PA 19073). These data sets were precompiled, mosaicked, and clipped to U.S. Geological Survey NLCD 2001 mapping zones (which are similar to ecoregions) (Homer and Gallant 2001) for the United States, and portions of the zones that intersect Ohio (zones 62, 53, 52, 51, and 47) were mosaicked to create a predictor data set.

The feature selection method I used was meant to find a subset of the predictor data set that would be effective at discriminating plots that are ecologically different from one another. The assumption of doing this is that plots have a unique ecological signature in feature space that can be used as a reference data set for labeling unknown locations in feature space. In order to select an effective subset of the predictor data set, a three-step process was used. The steps of this process were to (1) classify the plots based on the species composition data (independent of

the GIS predictor data), (2) rank the predictor attributes based on their ability to discriminate this species composition class, and then (3) choose a subset of predictors based on this ranking. First, WEKA[2] data mining software was used to classify each plot into one of 10 species composition classes (hereafter referred to as "forest types") using k-means cluster analysis (Witten and Frank 2005), with the total basal area of each species forming the axes used for clustering. In other words, forest types were created by forming 10 classes, each of which contained plots with similar species composition. I arbitrarily chose 10 classes because exploratory analyses showed that choosing around 10 yielded the most even distribution of points across the clusters.

Next, WEKA's implementation of the RELIEF attribute selector (Kira and Rendell 1992) was used to rank each GIS predictor variable. RELIEF worked by creating a usefulness index for each predictor attribute by randomly choosing a large number of instances from the data set and calculating for each selected instance the difference between the GIS predictor variable's value for the closest instance of the same "forest type" and that of the closest member of a different forest type. In the current study, if instances that were located close together along the axis defined by a predictor attribute are of the same forest type, RELIEF considered the attribute useful for discriminating between plots with different ecological characteristics and ranked it higher.

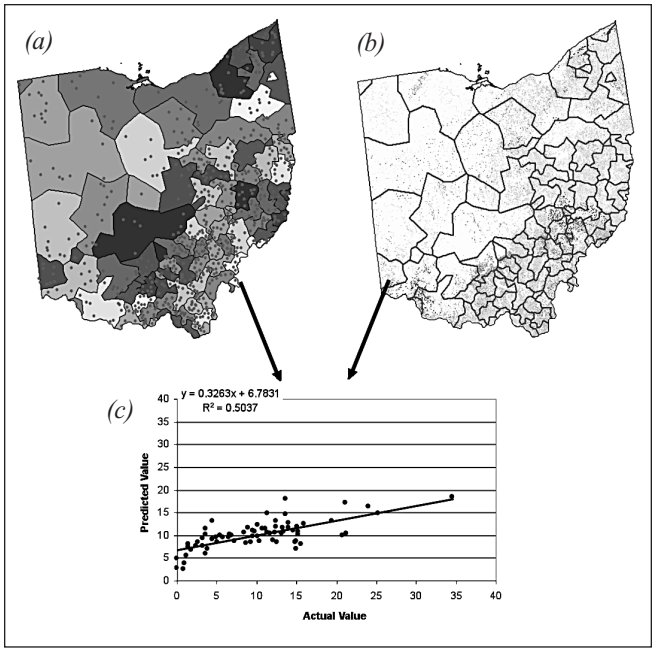
Finally, using the results of the attribute ranking, an arbitrary assessment of Pearson correlation matrices (to eliminate collinear variables with similar ranks), and my judgment from past use of similar data, I subjectively chose a set of 13 predictors that was useful for discriminating species composition class and was minimally correlated (correlations are generally less than 0.25). The 13 variables chosen in this manner were, in order of usefulness, MODIS band 5 (May 9, 2001), MODIS Enhanced Vegetation Index (EVI) (August 14, 2001), NLCD percent woody wetland, MODIS band 3 (November 17, 2001), distance to streams, the count of the variety of aspect values derived from a digital elevation model (an index of topographic roughness), MODIS band 7 (April 7, 2001), soil pH, soil texture, X coordinate, rock volume, Y coordinate, and minimum temperature in November.

Leica Geosystems' IMAGINE[2] image processing software was used to standardize the data set to the same measurement scale (0-1) and to extract values for each of the 13 standardized predictor layers where the FIA plots used in the analysis were located. IMAGINE's minimum (Euclidean) distance classifier was used to impute FIA plot information from the set of known pixels to unknown pixels based on multivariate similarity. The classification procedure gave every pixel in the study area the plot identifier (id) value of the FIA plot that is most similar to it with respect to values of the 13 predictors. A simple lookup table was then used to link pixel values in the image to tabular plot level summaries of FIA attributes based on this plot id value (as described in Lister (2005)). For this study, the plot id map was recoded to create maps of several attributes: BARM, CFGV, MUH and TDRYBIO.

Quality assurance (QA) was performed by a novel method of grouping plots into contiguous clusters, with each cluster containing exactly 10 plots (fig. 1a). This grouping was done by bit interleaving of the x and y coordinates of the plots in a manner similar to that described by Faloutsos and Rong (1991). Bit interleaving can be used to order plots based on proximity in two dimensions by drawing a line with fractal properties through the study area so that each plot is visited exactly once. The fractal line has the property of folding in upon itself in an orderly manner so that, in general, groups of points next to each other on the line are close in space. By partitioning this line into clusters that contain exactly 10 plots each, the procedure tessellates the study area based on density of forested plots. I chose 10 plots for each cluster arbitrarily, based on previous work that found this number to be a reasonable trade off between having a sufficiently small spatial cluster as the analysis unit and a sufficiently large group of plots to characterize that spatial cluster. If the spatial clusters were too big, the analysis became less meaningful, but if not enough plots were in each cluster, the variance of the cluster estimates made the analysis suspect.

By creating a raster representation of the study area and assigning each pixel the cluster id of the closest plot to that pixel, it was possible to summarize both the FIA plot data (fig. 1a) and the mapped estimates (fig. 1b) by cluster and construct a set of simple scatterplots (fig. 1c) that depict the relationship between the actual values (the average plot value) and the

Figure 1.—*(a) Contiguous clusters containing 10 plots each are created. Regions are built around each cluster using a GIS, and cluster-level summaries of the FIA plot data are calculated. (b) Map-based estimates are summarized for each region. (c) Scatterplots are created along with simple linear regression line diagnostics to characterize the actual vs. predicted relationship.*



estimates (the average pixel value). These relationships were then characterized by the parameters of the simple linear regression line (slope, intercept, and R^2) that describes the relationship between the set of actual and predicted values. The regression line method I used is a descriptive technique meant simply as a QA tool—I did not attempt to make inferences about the significance of the parameter estimates. The goal of the QA procedure was to provide a tool for data consumers to assess the relationship between actual and predicted values in a spatially explicit manner.

Results and Discussion

Many of the attributes that were selected by the RELIEF method are factors that would tend to influence vegetation composition in a landscape dominated by the effects of past glaciations, like Ohio. For example, soil pH, texture, and rock volume would be affected by past glacial activity, as would stream density and wetland occurrence. The MODIS-related imagery variables probably appeared higher in the usefulness ranking because of differences in phenology of the different species composition groups—certain species assemblages reflect light differently at different times of the year. The goal of the attribute selection approach was not strictly to extract biologically meaningful predictor variables in a quantitative way. Rather, it was to use a combination of RELIEF, correlation tests, and user opinion as a guide in attribute selection. It is noteworthy that the chosen variables (which were ranked higher by RELIEF) probably have at least some functional relationships with factors that affect plant growth in Ohio.

The maps of the FIA attributes selected are shown in figure 2 along with magnified areas to show examples of the finer scale variability of the estimates. Ohio is nearly 70 percent nonforest, so I used a nonforest mask (the NLCD 1992 data) to mask out nonforest areas and water, which accounts for some of the observed patterns in the maps. Nonforest masking allows the map to retain certain landscape features that FIA doesn't measure (e.g., the occurrence of rivers) while imputing FIA attributes to forested areas of the State.

In general, the southwest part of the State has more biomass and volume than other parts of the State. The landscape maps produced clearly show these patterns of the FIA attributes across the landscape, and the example areas that are shown at higher resolution show some of the finer scale pattern that the modeling procedure produces. In general, however, the interpretation of these maps is best made at the landscape level. Pixel-scale interpretation is possible, although inadvisable because it is nearly impossible to find QA reference data that correspond well with MODIS pixels. The FIA plots cover 0.067 ha, or approximately 1/100 of a 250-m MODIS pixel—thus making plot-pixel accuracy statements nearly meaningless.

On the other hand, the QA results within the zones depicted in figure 1 can serve to inform the user about the relative utility of the maps at a given geographic scale. Figures 3a, 3b, 3c, and 3d show the relationships (and corresponding diagnostic statistics) between sets of actual and predicted levels of the FIA attributes. The r^2 values ranged from 0.4 (for the biomass-related

Figure 2.—*Maps of total dry aboveground biomass (TDRYBIO), occurrence of mixed hardwoods forest type (MUH), cubic foot gross volume (CFGV) of trees, and basal area per acre of red maple (BARM), with small areas shown at higher resolution to show local detail.*

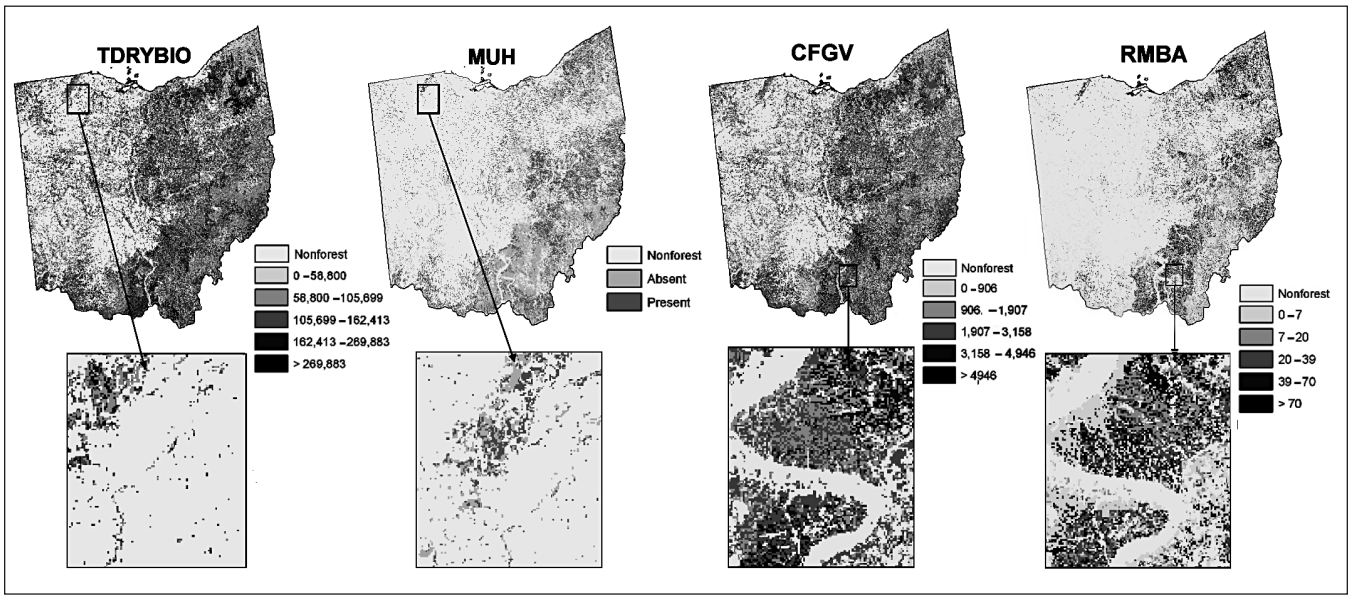
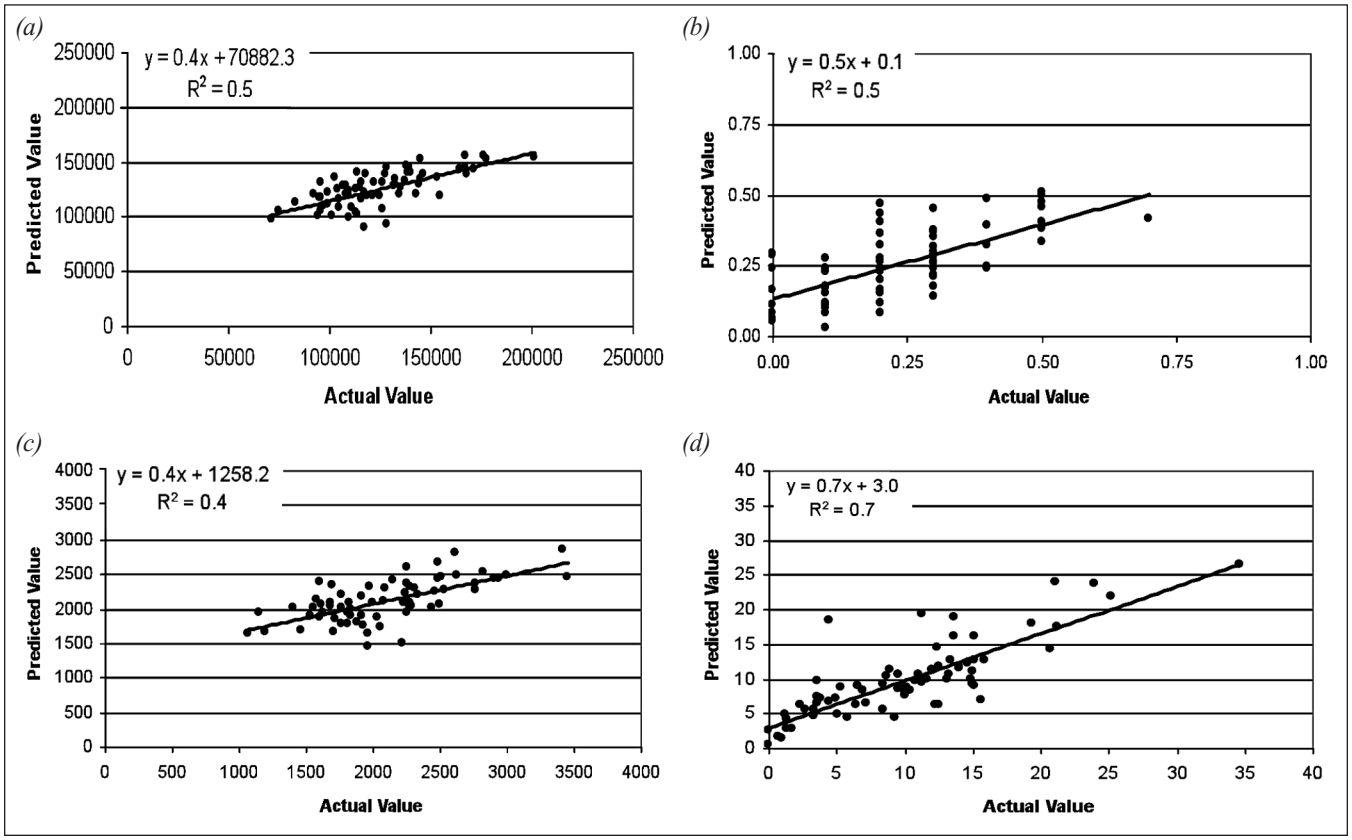


Figure 3.—*Actual vs. predicted scatterplots and associated simple linear regression diagnostics of (a) total dry aboveground biomass, (b) mixed upland hardwoods forest type, (c) cubic foot gross volume, and (d) basal area per acre of red maple.*



variables) to 0.7 (for the basal of red maple). The slope and intercepts of the simple linear regression lines that describe the relationships between actual and predicted values tend to indicate an overprediction of low actual values and an underprediction of high values. I have encountered this phenomenon in many other multivariate modeling methods and believe that it occurs because weak univariate and multivariate relationships between an FIA plot and the MODIS pixel on which it sits tend to increase the occurrence of misclassification. In most FIA data sets, a random assignment of an FIA plot to an unknown pixel (which is the extreme of what occurs in misclassification) will always yield an estimate closer to the mean of the value for all FIA plots involved than it will a value near the extremes. This principle leads to the observed pattern of truncation of the variance of the set of estimates.

The main reason I chose the novel QA method I used was to guarantee that an equal number of FIA plots would be in each QA zone (fig. 1a). In that manner, the confidence I can put in each QA point is equal with respect to the FIA plots, which generally show the largest amount of variability. Had I chosen another approach using a regular tessellation of the study area to produce QA zones, large areas of the State would not have been assessed because cells in mostly nonforest areas would not have at least 10 FIA plots and would not be used as valid QA polygons. By grouping plots using the bit interleaving method, not only am I able to perform the QA method within contiguous geographic regions but I also have increased the interpretability of the results of the analysis. Interpreting these QA results is predicated on recognizing that each QA zone represents a roughly equal amount of *forest* land, not total land.

The utility of landscape maps such as these is tempered by the truncation of the variance I observe in the QA results. Were the goal of my study to create an accurate map of a single FIA attribute, I would have optimized my choice of predictors and modeling technique. For the purposes of FIA's State reporting, however, a method that produces a single map (the plot id map) and uses a lookup table to create a map of any FIA attribute that can be associated with a plot is clearly desirable. The landscape maps that I have produced not only show the distribution of the attributes of interest across the landscape but also retain logical consistency. Each map produced retains the entire plot record for each pixel; thus, e.g., a situation where the basal area of red maple at a given location is predicted to be higher than that of the total basal area of all species cannot occur. The disadvantage of this technique, however, is that any individual map that was not produced using optimal methods or predictor data is not optimized for accuracy. New imputation techniques, like those being implemented by Wilson (2006), use advanced data reduction methods and attribute weighting, which could mitigate this problem. These advanced methods show great promise for national implementation, and future work will be in support of this goal.

Literature Cited

Andersen, G.L. 1998. Classification and estimation of forest and vegetation variables in optical high resolution satellites: a review of methodologies. Interim report IR-98-085. Laxenburg, Austria: International Institute for Applied Systems Analysis. 25 p.

Bechtold, W.A.; Patterson, P.L. 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.

Braun, E.L. 1950. Deciduous forests of eastern North America. Philadelphia, PA: Blaskiston Company. 596 p.

Breiman, L.; Friedman, J.; Olshen, R.A.; Stone, C.J. 1984. Classification and regression trees. Pacific Grove, CA: Wadsworth & Brooks/Cole Advanced Books & Software.

Daly, C.; Gibson, W.P.; Doggett, M.; Smith, J.; Taylor, G. 2004. Up-to-date monthly climate maps for the conterminous United States. In: Proceedings, 14th American Meteorological Society conference on applied climatology, 84th American Meteorological Society annual meeting combined preprints. Pap. P5.13. Seattle, WA: American Meteorological Society. http://ams.confex.com/ams/pdfpapers/71444.pdf.

Downing, M. [N.d.]. Invasive pest risk maps. On file with: M. Downing, Forest Health Protection/Forest Health Technology Enterprise Team, 2150 Centre Avenue, Building A, Suite 331, Information Technology, Fort Collins, CO 80526–1891. http://www.fs.fed.us/foresthealth/technology/invasives_sirexnoctilio_riskmaps.shtml.

Faloutsos, C.; Rong, Y. 1991. DOT: a spatial access method using fractals. In: Proceedings of the International Conference on Data Engineering (ICDE). Kobe, Japan: Institute of Electrical and Electronics Engineers: 152-159.

Fassnacht, K.S.; Cohen, W.B.; Spies, T.A. 2006. Key issues in making and using satellite-based maps in ecology: a primer. Forest Ecology and Management. 222: 167-181.

Frescino, T.S.; Edwards, T.C., Jr.; Moisen, G.G. 2001. Modeling spatially explicit forest structural attributes using generalized additive models. Journal of Vegetation Science. 12: 15-26.

Gesch, D.; Oimoen, M.; Greenlee, S.; Nelson, C.; Steuck, M.; Tyler, D. 2002. The national elevation dataset. Photogrammetric Engineering & Remote Sensing. 68(1): 5-32.

Homer, C.; Gallant, A. 2001. Partitioning the conterminous United States in mapping zones for Landsat TM land cover mapping. Sioux Falls, SD: U.S. Department of the Interior, U.S. Geological Survey, Eros Data Center. 7 p.

Justice, C.; Vermote, E.; Townshend, J.R.G. *et al.* 1998. The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. IEEE Transactions on Geoscience and Remote Sensing. 36(4): 1228-1249.

Kira, K.; Rendell, L.A. 1992. A practical approach to feature selection. Proceedings, ninth international conference on machine learning. Aberdeen, Scotland: Morgan Kaufmann: 249-256.

Lister, A.J. 2005. Creation of an n-dimensional spatial database instead of a map. In: Marsden, M.; Downing, M.; Riffe, M., comps. Workshop proceedings: quantitative techniques for deriving national-scale data. FHTET-05-12. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Forest Health Technology Enterprise Team: 14-22.

Lister, A.J.; Riemann, R.; Hoppus, M. 2000. A nonparametric geostatistical approach for estimating species importance. In: Reams, G.A.; McRoberts, R.E.; VanDeusen, P.C., eds. Proceedings, second annual forest inventory and analysis symposium. USDA GTR SRS-47. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station: 52-62.

Little, E.L., Jr. 1971. Atlas of United States trees. Misc. Publ. 1146. Washington, DC: U.S. Department of Agriculture. 9 p. Vol. 1.

Little, E.L., Jr. 1981. Atlas of United States trees. Supplement. Misc. Publ. 1410. Washington, DC: U.S. Department of Agriculture. 31 p. Vol. 5.

McRoberts, R.E.; Nelson, M.D.; Wendt, D.G. 2002. Stratified estimation of forest area using satellite imagery, inventory data, and the *k*-Nearest Neighbors technique. Remote Sensing of Environment. 82(2): 457-468.

Moisen, G. [N.d.]. Forest biomass of the conterminous United States. Manuscript in preparation. On file with: G. Moisen, U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis Unit, Rocky Mountain Research Station, 507 25th Street, Ogden, UT 84401.

Moisen, G.G.; Edwards, T.C., Jr. 1999. Use of generalized linear models and digital data in a forest inventory of northern Utah. Journal of Agricultural, Biological, and Environmental Statistics. 4: 372-390.

Natural Resources Conservation Service (NRCS). 1994. State Soil Geographic (STATSGO) soil database: data use information. Misc. Publ. 1492. Washington, DC: National Soil Survey Center. 88 p.

Ruefenacht, B. [N.d.]. Forest type groups of the conterminous United States. Manuscript in preparation. On file with: B. Ruefenacht, U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center, Salt Lake City, UT 84119. http://svinetfc4.fs.fed.us/rastergateway/forest_type/.

Sargent, S. 1884. Report on the forests of North America (exclusive of Mexico). Washington, DC: U.S. Department of the Interior, Government Printing Office.

Shantz, H.L.; Zon, R. 1924. Atlas of American agriculture. Washington, DC: U.S. Department of Agriculture, Government Printing Office.

U.S. Department of Agriculture (USDA). 1998. A Forest Service Forest Inventory and Analysis (FIA) strategic plan for forest inventory and monitoring. Washington, DC: U.S. Department of Agriculture, Forest Service. 48 p.

U.S. Environmental Protection Agency (EPA). 1994. GIRAS Landuse/Landcover data for the conterminous United States by quadrangles at scale 1:250,000. Washington, DC: U.S. Environmental Protection Agency, Office of Information Resources Management.

U.S. Geological Survey (USGS). 1999. The National Hydrography Dataset. Fact Sheet 106-99. Reston, VA: U.S. Geological Survey. 2 p.

Vogelmann, J.E.; Howard, S.M.; Yang, L. *et al.* 2001. Completion of the 1990s National Land Cover Data Set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. Photogrammetric Engineering and Remote Sensing. 67(6): 650-662.

Wilson, B.T. 2006. Northern Research Station FIA program's atlas of FIA maps. Unpublished document. On file with: B.T. Wilson, Northern Research Station, Forest Inventory and Analysis program, St. Paul, MN 55108.

Witten, I.H.; Eibe, F. 2005. Data mining: practical machine learning tools and techniques. 2nd ed. San Francisco: Morgan Kaufmann. 525 p.

Zhu, Z. 1994. Forest density mapping in the lower 48 States: a regression procedure. Res. Pap. SO-280. New Orleans: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station. 11 p.

Zhu, Z.; Evans, D.L. 1994. U.S. forest types and predicted percent forest cover from AVHRR data. Photogrammetric Engineering and Remote Sensing. 60(5): 525-531.

Estimation of a Cover-Type Change Matrix From Error-Prone Data

Steen Magnussen[1]

Abstract.—Coregistration and classification errors seriously compromise per-pixel estimates of land cover change. A more robust estimation of change is proposed in which adjacent pixels are grouped into 3×3 clusters and treated as a unit of observation. A complete change matrix is recovered in a two-step process. The diagonal elements of a change matrix are recovered from estimates of the temporal correlations of cover-type frequencies and an estimate of the odds-ratio of no change. Off-diagonal elements are recovered from least-squares solutions to a set of constrained linear equations. The proposed method produced less biased estimates on three of five sites when the average coregistration error was in excess of 0.3 to 0.7 pixels and on four of five sites if classification accuracy is below 0.9.

Introduction

Land cover change estimation from remotely sensed data is riddled with a unique set of problems due to registration errors when two images are coregistered (Coppin and Bauer 1996, Coppin *et al.* 2004) and classification errors (Congalton 2001, Pontius Jr. and Lippitt 2006). A registered change in a unit can therefore misrepresent the actual change event. Although techniques for reducing the bias due to classification errors are readily available (Czaplewski 2003, Stehman and Czaplewski 1998), their efficiency depends critically on an accurate estimate of a confusion matrix for all change classes; a reality that is rarely met.

This study proposes a new method for estimating a $K\times K$ change matrix for K cover types from clusters of pooled pixels instead of individual pixels. Change estimated from clusters of spatially adjoining units is assumed to be less sensitive to the problems stated previously than a per-pixel estimation. Clusters must be large enough to mitigate the effect of the aforementioned problems, yet small enough to reduce the loss of information that occurs when pixels are pooled. A cluster of 3×3 pixels is chosen as a compromise.

Estimates of no change on the main diagonal of a change matrix $k, k=1,\ldots,K$ are recovered from temporal correlation coefficients of cluster-level pixel counts and the odds ratio of no change (Magnussen 2004). Off-diagonal elements are recovered by least-squares solutions (CLS) to a set of linear constraints.

The performance of the proposed method is assessed for a 4×4 change matrix with data from five sites. Coregistration and classification errors are simulated across of a range of specifications. For a detailed version of this study, see Magnussen (2007).

Material and Methods

Recovery of the change matrix begins with a complete tessellation of the coregistered and the classified time 1 (t_1) and time 2 (t_2) images into M size 9 clusters with pixels in a 3×3 array. Only cluster-level counts of the number of pixels in each class at t_1 and t_2 will be used for the recovery. Specifically, the t_1 data are $\mathbf{n}_1^{(i)} = n_{1+}^{(i)},\ldots,n_{K+}^{(i)}, i=1,\ldots,M$, where $n_{k+}^{(i)}=\sum_{k'} n_{kk'}^{(i)}$; i.e., the kth row sum of $\mathbf{n}_i$, is the number of class k units in the ith cluster at time t_1; likewise, t_2 data are $\mathbf{n}_2^{(i)} = n_{+1}^{(i)},\ldots,n_{+K}^{(i)}, i=1,\ldots,M$, where $n_{+k}^{(i)}=\sum_{k'} n_{k'k}^{(i)}$. The state of the population at t_1 and t_2 is captured by the vectors $\mathbf{n}_1 = n_{1+},n_{2+},\ldots n_{K+}$ and $\mathbf{n}_2 = n_{+1},n_{+2},\ldots n_{+K}$. The total number of pixels is $\sum_k n_{k+} = \sum_k n_{+k} = n_{++}$.

[1] Research Scientist, Natural Resources Canada, Canadian Forest Service, 506 West Burnside Road, Victoria BC, V8Z 1M5, Canada. E-mail: steen.magnussen@nrcan.gc.ca.

Recovery of the Main Diagonal of the Change Matrix

A procedure for the recovery of the main diagonal in the change matrix has been detailed by Magnussen (2004). First, Pearson's correlation coefficients ρ_{kk} of $n_{k+}^{(i)}, n_{+k}^{(i)}, i = 1,\ldots,M, k = 1,\ldots,K$ are computed. From these coefficients, an estimate $\tilde{n}_{kk}$ is obtained, as outlined by Murtaugh and Phillips (1998). A second estimate $\breve{n}_{kk}$ is obtained by finding an odds ratio of no change for class k ($\equiv n_{kk}\sum_{k\neq k'} n_{k'k'} \times (n_{k'} - n_{k'k'})(n_{+k'} - n_{k'k'})$) (Fleiss 1981) that maximizes the likelihood of the observed counts $\mathbf{n}_t^{(i)}, i = 1,...,M, t = 1,2$ (Magnussen 2004). A linear combination of these two estimates is used as the final estimate—specifically, $\hat{n}_{kk} = \frac{2}{3}\tilde{n}_{kk} + \frac{1}{3}\breve{n}_{kk}$.

Recovery of Off-Diagonal Elements in the Change Matrix

Recovered off-diagonal elements $n_{kk'|k\neq k'}$ are CLS to a set of constrained equations. From (n_{k+}, n_{+k}) and the previous estimates of $\hat{n}_{kk}$, we can formulate a trivial set of $K \times K$ constraints on the row and column sums of the elements in the change matrix. If we were to recover the off-diagonal elements in a 3×3 change matrix, we could formulate a rank five set of constraints for the six off-diagonal elements. The ratio of unknowns (6) to the rank of the linear constraints (5) is the maximum possible. Hence, if we could reduce the recovery problem to a 3×3 change matrix, the CLS recovery would be the best possible. For a $K \times K$ change matrix, we can create $Bin(K,3)$ 3×3 change matrices and find a set of 14 linear constraints on their off-diagonal elements with a maximum possible rank of 9 and then find the CLS solution by least squares (Magnussen 2007). All CLS estimates satisfy: $\hat{n}_{kk'}^{CLS} \geq 0$, $\sum_{k,k'} \hat{n}_{kk'}^{CLS} = n_{++}$, and $n_{kk'} \supset \text{Integer}, k,k' = \{1,\ldots,K=4\}$.

Performance Assessment

In five case studies with four land-cover classes (K = 4), the recovered change matrix $\hat{\mathbf{n}}_{CLS}$ is compared to the actual matrix $\hat{\mathbf{n}}$ obtained by a direct counting (DIR) of pixel-level change. To facilitate a Monte Carlo simulation of errors of coregistration and classification (see the next section), the assessment is carried out with data from 200 replications of simple random sampling of m = 200 3×3 clusters with and without errors.

Monte Carlo Simulation of Coregistration and Classification Errors

Eleven levels of average coregistration errors in t_2 data were simulated at the cluster level. With a probability of P_0, the "true" data in a 3×3 cluster had no coregistration error ($P_0 = 0, 0.1,..., 0.9, 1.0$). With a probability of $P_1 = \frac{2}{3}(P_0)$, the location of a cluster was either shifted one column to the left (right) or one row up (down). With a probability of $P_2 = \frac{1}{3}(P_0)$, the location of a cluster was shifted one column to the left (right) and one row up (down). Registration errors in each of the 200 replications of a random sample of 200 clusters were distributed at random across clusters.

Classification errors at t_1 and t_2 were assumed to be equal and independent. The following symmetric 4×4 confusion matrix $\mathbf{P}$ was used to simulate multinomial classification errors:

$$\mathbf{P} = \begin{bmatrix} & True=1 & True=2 & True=3 & True=4 \\ Clsf.=1 & p_{k|k} & \frac{17}{24}(1-p_{k|k}) & \frac{5}{24}(1-p_{k|k}) & \frac{2}{24}(1-p_{k|k}) \\ Clsf.=2 & \frac{17}{24}(1-p_{k|k}) & p_{k|k} & \frac{2}{24}(1-p_{k|k}) & \frac{5}{24}(1-p_{k|k}) \\ Clsf.=3 & \frac{5}{24}(1-p_{k|k}) & \frac{2}{24}(1-p_{k|k}) & p_{k|k} & \frac{17}{24}(1-p_{k|k}) \\ Clsf.=4 & \frac{2}{24}(1-p_{k|k}) & \frac{5}{24}(1-p_{k|k}) & \frac{17}{24}(1-p_{k|k}) & p_{k|k} \end{bmatrix} \quad (1)$$

The value of $p_{k|k}$, the classification accuracy, was varied from 0.5, 0.6,…, 0.9 with the value fixed during one simulation.

Examples

Remotely sensed data from five large forested areas (BC, HI, IT, NB, SE), representing different regional landscapes with contrasting cover-type composition and rates of presumed change, are used for demonstrating the performance of the proposed alternative estimator of a change matrix when data are potentially error-prone. The data domains vary in size from 109 km^2 (BC) to 188 km^2 (IT). Details are in Magnussen (2004). Population sizes in pixels of approximately 30 × 30 m were 121,104 (BC), 129,600 (HI), 208, 675 (IT), 181,068 (NB), and 166,464 (SE). An example of a population change matrix for IT is in table 1.

Table 1.—*True change matrix for IT.*

Class	1	2	3	4	All (t_1)
1	30771 (14.7)	89 (0.0)	5117 (2.5)	4262 (2.0)	40239 (19.3)
2	2 (0.0)	1455 (0.7)	80 (0.0)	12 (0.0)	1549 (0.7)
3	3582 (1.7)	3321 (1.6)	64977 (31.1)	1876 (0.9)	73756 (35.3)
4	630 (0.3)	30 (0.0)	364 (0.2)	92107 (44.1)	93131 (44.6)
All(t_2)	34985 (16.8)	4895 (2.4)	70538 (33.8)	98257 (47.1)	208675 (100.0)

Notes: Table entries are number of pixels. Numbers in parentheses are percentages of the total. Percentages may not add to 100 due to rounding.

Results

Only results pertaining to IT are given. Diagonal elements of $\hat{\mathbf{n}}_{CLS}$ estimates were generally estimated with less bias than off-diagonal elements; a trend that is especially clear on a relative scale where the bias of diagonal estimates was one-half to one-tenth of the bias in off-diagonal change estimates. Relative bias of estimated diagonal elements was in the range of 5 to 10 percent. Scatter-plots in figure 1 give additional insight to CLS performance. For the diagonal elements, the relationship to DIR was typically linear with $0.97 \geq \hat{R}^2_{adj} \geq 0.91$, yet with a persistent bias. For the off-diagonal elements, the scatter plots suggest a linear relationship but also a persistent bias.

Coregistration errors in t_2 data introduce bias in $\hat{\mathbf{n}}$ and increased the bias in $\hat{\mathbf{n}}_{CLS}$. As the coregistration error increases, however, the bias increased three to four times faster in $\hat{\mathbf{n}}$ than in $\hat{\mathbf{n}}_{CLS}$ (fig. 2). As a result, when the average coregistration error reaches 0.3, one can expect less bias in the diagonal elements in $\hat{\mathbf{n}}_{CLS}$ than in $\hat{\mathbf{n}}$. A similar situation arises for the off-diagonal elements when the average coregistration error exceeds 0.4.

Classification errors also generate a serious bias in $\hat{\mathbf{n}}$ and $\hat{\mathbf{n}}_{CLS}$ (fig. 3). As expected, diagonal elements are most sensitive to classification errors because a classification error generates a change event where none occurred. The rate at which bias increased as classification accuracy decreased was about 1.5 times higher in $\hat{\mathbf{n}}$ than in $\hat{\mathbf{n}}_{CLS}$. The critical accuracy level below which $\hat{\mathbf{n}}_{CLS}$ would be less biased than $\hat{\mathbf{n}}$ is around 0.9.

Figure 1.—*Scatter plot of $\hat{n}^{CLS}_{kk'}$ versus $\hat{n}_{kk'}$ in 200 replicated samples of size 200 in IT.*

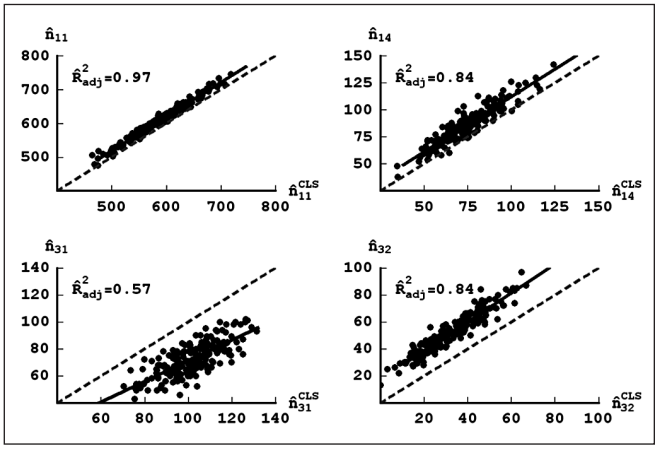


Figure 2.—*Mean absolute bias (MAD) of elements of $\hat{\mathbf{n}}_{CLS}$ and direct counts $\hat{\mathbf{n}}$ (DIR) plotted against average coregistration error (unit: pixel of 30 × 30 m).*

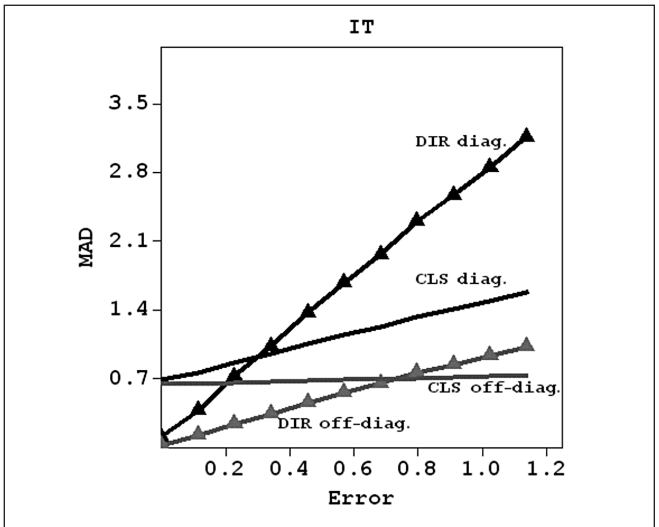
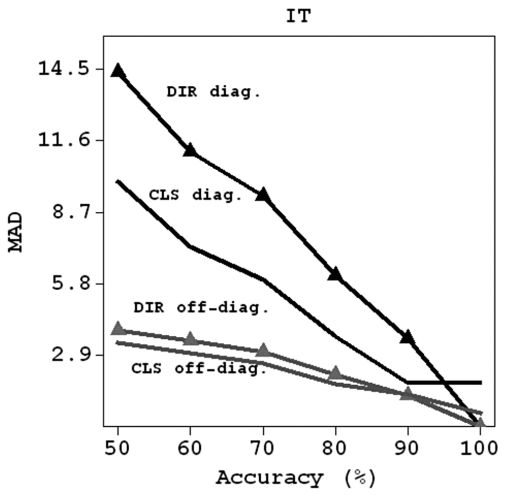


Figure 3.—*Mean absolute bias (MAD) of elements of* $\hat{\mathbf{n}}_{CLS}$ *and direct counts* $\hat{\mathbf{n}}$ *(DIR) plotted against classification accuracy.*



Black: diagonal elements. Triangles: recovered.
Gray: off-diagonal elements. Full line: direct counts.

Discussion

A change matrix recovered from cluster-level counts of units per class at two points in time is, in absence of registration and classification errors, less accurate than a change matrix obtained by direct counts of change events. The proposed recovery method is an option when unit-level change is seriously compromised by errors. The choice of a 3×3 cluster was a compromise between conflicting goals. A larger cluster would be more robust against errors but would incur additional loss of information and increases in computational complexity. A smaller cluster, however, would be less robust against registration and classification errors without offering any significant computational advantages.

The Monte Carlo simulations confirmed the sensitivity of unit-level change estimates derived directly from remotely sensed data to errors of coregistration and classification (Bruzzone and Cossu 2003, Lunetta and Elvidge 1999). Average coregistration errors in the range of 0.3 to 1.1 units (pixels) are not uncommon. Classification accuracies of 0.7–0.9 units are commonly reported for forest cover-type maps derived from Landsat Enhanced Thematic Mapper+ (Foody 2002, Holmgren and Thureson 1998). Thus, if no better approach to mitigate the bias is available, the proposed recovery should be pursued when classification accuracy and registration errors have the potential to seriously compromise the results.

Acknowledgments

Data for the IT site were kindly provided by Dr. P.M. Corona of the Università degli Studi della Tuscia in Italy.

Literature Cited

Bruzzone, L.; Cossu, R. 2003. An adaptive approach to reducing noise effects in unsupervised change detection. Geoscience and Remote Sensing. 41(11): 2455-2465.

Congalton, R.G. 2001. Accuracy assessment and validation of remotely sensed and other spatial information. International Journal of Wildland Fire. 10: 321-328.

Coppin, P.R.; Bauer, M.E. 1996. Digital change detection in forest ecosystems with remote sensing imagery. Remote Sensing of Environment. 13: 207-234.

Coppin, P.R.; Jonckheere, I.; Nackaerts, K.; Muys, B.; Lambin, E.F. 2004. Digital change detection methods in ecosystem monitoring: a review. International Journal of Remote Sensing. 25(9): 1565-1596.

Czaplewski, R.L. 2003. Accuracy assessment of maps of forest conditions: statistical designs and methodological considerations. In: Wulder, M.A.; Franklin, S.E., eds. Remote sensing of forest environments: concepts and case studies. Boston: Kluwer Academic Publishers: 115-140.

Fleiss, J.L. 1981. Statistical methods for rates and proportions. New York: John Wiley. 313 p.

Foody, G.M. 2002. Status of land cover classification accuracy assessment. Remote Sensing of Environment. 80: 185-201.

Holmgren, P.; Thureson, T. 1998. Satellite remote sensing for forestry planning—a review. Scandinavian Journal of Forest Research. 13: 90-110.

Lunetta, R.S.; Elvidge, C.D. 1999. Remote sensing change detection: environmental monitoring methods and applications. London: Taylor & Francis. 318 p.

Magnussen, S. 2004. Prediction of 2 x 2 tables of change from repeat cluster sampling of marginal counts. Canadian Journal of Forest Research. 34(8): 1703-1713.

Magnussen, S. 2007. A method for estimation of a land cover change matrices from error-prone unit-level observations. Canadian Journal of Forest Research. 37(12): 1505-1517.

Murtaugh, P.A.; Phillips, D.L. 1998. Temporal correlation of classification in remote sensing. Journal of Agricultural, Biological and Environmental Statistics. 3(1): 99-110.

Pontius, R.G., Jr.; Lippitt, C.D. 2006. Can error explain map differences over time? Cartography and Geographic Information Science. 33(2): 159-171.

Stehman, S.V.; Czaplewski, R.L. 1998. Design and analysis for thematic map accuracy assessment: fundamental principles. Remote Sensing of Environment. 64: 331-344.

Bridging the Gap Between Strategic and Management Forest Inventories

Ronald E. McRoberts[1]

Abstract.—Strategic forest inventory programs collect information for a large number of variables on a relatively sparse array of field plots. Data from these inventories are used to produce estimates for large areas such as states and provinces, regions, or countries. The purpose of management forest inventories is to guide management decisions for small areas such as stands. Management inventories collect data for a smaller number of variables using greater sampling intensities and produce estimates for small areas. Increasingly, management inventories are coming to be regarded as prohibitively expensive. A relatively inexpensive alternative is to use a combination of strategic inventory data and remotely sensed data such as satellite imagery to construct maps of forest attributes and then aggregate pixel predictions to obtain estimates of stand-level means. The study emphases included using forest inventory plot data, Landsat Thematic Mapper satellite imagery, and a modification of the k-Nearest Neighbors technique to construct stem density and basal area maps from which estimates of stand-level means were calculated. The results indicate that although unbiased and precise estimates of stand-level means for below canopy attributes may be difficult to obtain, stands may nevertheless be ranked quite accurately with respect to a combination of stem density and basal area. Such rankings inform management decisions by identifying stands that require immediate attention.

Introduction

The national forest inventories (NFI) of the countries of western and central Europe and North America may be characterized as strategic inventories. In the United States, the NFI is conducted by the Forest Inventory and Analysis (FIA) program of the Forest Service, U.S. Department of Agriculture. Strategic forest inventories typically feature observation or measurement of an array of plots distributed across entire countries or regions of countries according to a predetermined sampling design. Sampling intensities range from one plot per a few hundred hectares to one plot per 10,000s of hectares. Strategic inventory programs generally observe or measure a large number of variables for each plot and produce estimates for large areas such as states or provinces, regions, and countries. The objectives of management inventories, characterized as stand examinations in the United States, are to produce sufficiently accurate stand-level estimates to guide management decisions. Thus, management inventories focus on fewer variables for smaller stand-size areas but feature greater sampling intensities, perhaps as intense as multiple plots per hectare. Despite fewer variables, the required sampling intensities for management inventories result in considerable costs. In fact, in many countries management inventories are increasingly regarded as prohibitively expensive.

The emerging alternatives to ground plot-based management inventories require intensive use of fine resolution remotely sensed data (e.g., Flewelling 2009). The proposed uses of these data are to delineate individual tree crowns, either by border to border or for sample locations, and to produce accurate estimates of tree attributes such as crown dimensions, species, and height. These tree attribute estimates are then used as input to statistical models that are calibrated using ground plot data and used to predict tree diameters. Finally, the estimates of species, height, and diameter are then used to estimate individual tree volumes and biomasses, which are aggregated to produce

[1] Mathematical Statistician, U.S. Department of Agriculture, Forest Service, Northern Research Station, Forest Inventory and Analysis, 1992 Folwell Avenue, St. Paul, MN 55108. E-mail: rmcroberts@fs.fed.us.

estimates of stand-level means. These proposed alternative approaches are costly because fine resolution remotely sensed data and the ground plot data for calibrating the models must be acquired. In addition, these approaches are still mostly in research and development stages and have not yet been demonstrated to be operationally feasible.

Alternatives exploiting existing, accessible ground data and inexpensive satellite imagery have not been investigated. Possible approaches include using the k-Nearest Neighbors (k-NN) technique in combination with NFI data and 30-m x 30-m resolution Landsat satellite imagery. The k-NN technique is a nonparametric, multivariate approach to imputing values of variables observed or measured on sample units to units that are similar in a covariate space but without observations. The k-NN approach with inventory data and Landsat imagery has been shown to produce representative maps of forest attributes (McRoberts *et al.* 2002, Tomppo and Halme 2004). In addition, areal estimates of mean forest attributes obtained by averaging k-NN pixel predictions have been shown to be comparable to sample-based estimates for 10 km radius circular areas (McRoberts *et al.* 2007). These results suggest that aggregations of predictions over all stand pixels may produce estimates of stand-level means that are sufficiently accurate either to mitigate the necessity of conducting costly, labor-intensive, ground plot-based management inventories or to create efficiencies by informing decisions as to which stands require management inventories. The objective of the study was to estimate stand-level means of stem density (SD) (trees/hectare) and basal area (BA) (m^2/hectare) using the k-NN technique with FIA plot data and Landsat satellite imagery. Successfully achieving the objective would constitute at least a partial bridging of the gap between strategic and management inventories.

Data

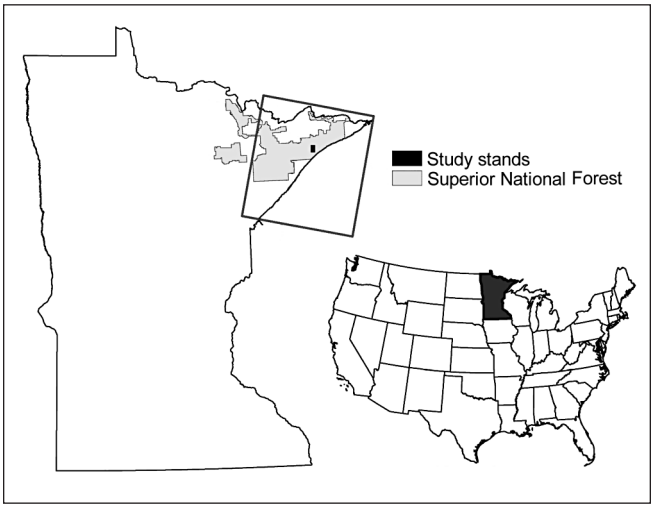
Satellite Image Data

The northern Minnesota study area was located wholly within the portion of the Superior National Forest in the path 26, row 27, Landsat scene (fig. 1). Three dates of Landsat Thematic Mapper (TM) imagery were obtained for this scene: November 1999 (fall), May 2003 (spring), and June 2003 (summer). Preliminary analyses indicated that the Normalized Difference Vegetation Index (NDVI) and the tassel cap (TC) transformations (brightness, greenness, and wetness) (Crist and Cicone 1984, Kauth and Thomas 1976) were superior to both the raw spectral band data and principal component transformations with respect to predicting forest attributes. Thus, NDVI and the three TC transformations for each of the three image dates were the 12 satellite image-based covariates.

FIA Plot Data

The FIA program has established field plot centers in permanent locations across the United States using a sampling design that is assumed to produce a random, equal probability sample (Bechtold and Patterson 2005, McRoberts *et al.* 2005). The plot array has been divided into five nonoverlapping, interpenetrating panels, and measurement of all plots in one panel is completed before measurement of plots in the next panel is initiated. Panels are selected on a 5-year rotating basis in the study area. Over a complete measurement cycle, the FIA sampling intensity is approximately one plot per 2,400 ha (6,000 acres). The State of Minnesota provides additional funding to double the sampling intensity to approximately one plot per 1,200 ha (3,000 acres). Each plot consists of four 7.31-m (24-ft) radius circular subplots. The subplots are configured as a central subplot and three peripheral subplots with centers located at 36.58 m (120 ft) and azimuths of 0, 120, and 240

Figure 1.—*Study area defined by the path 26, row 27 Landsat scene.*



degrees from the center of the central subplot. The path 26, row 27, Landsat scene included 1,086 FIA plots or 4,344 subplots measured between 1999 and 2003. For this study, observations and measurements on these plots were restricted to trees with diameter at breast height (d.b.h.) of 12.5 cm (5 in) or greater. Plot-level variables such as forest type, basal area per unit area, volume per unit area, and stem density per unit area may be calculated from the observed and measured variables.

In general, locations of forested or previously forested plots were determined using global positioning system receivers, while locations of nonforested plots were determined using aerial imagery and digitization methods. The spatial configuration of the FIA subplots with centers separated by 36.58 m and the 30-m x 30-m spatial resolution of the Landsat TM/ETM+ imagery permits individual subplots to be associated with individual image pixels. The subplot area of 167.87 m^2 (1,810 ft^2) is approximately 19 percent of the 900 m^2 (9,687 ft^2) pixel area, and subplot observations and measurements were assumed to characterize entire pixels.

Validation Data

Data from an independent study were used to assess the accuracy of the k-NN estimates of stand-level SD and BA means. Boundaries for 13 stands in the Superior National Forest study area, ranging in size from approximately 5 ha to approximately 27 ha, were delineated using segmentation techniques based on low altitude photography (table 1). In each stand, 6 to 15 temporary, variable radius plots were established using basal area factor 10; sampling intensities ranged from slightly more than one plot per hectare to more than two plots per hectare. Species were observed and d.b.h. was measured for all trees with d.b.h. greater than or equal to 12.5 cm (5 in).

Methods

k-NN Technique

The k-NN technique is an intuitive, nonparametric approach to either univariate or multivariate prediction based on the similarity in a feature space between the unit for which a prediction is desired and units for which observations are available. The set of 4,344 TM pixels for which corresponding subplot observations were available was designated the k-NN reference set, and the set of pixels with centers in the 13 stands was designated the k-NN target set. A basic implementation of the k-NN technique was used. The similarity measure in feature space was Euclidean distance,

$$d_{ij} = \sqrt{\sum_{l=1}^{L} \left(x_{il} - x_{jl} \right)^2}, \quad (1)$$

Table 1.—*Validation data.*

Stand	Area (ha)	No. plots	Sampling intensity (plots/ha)	SD		BA	
				Mean (trees/ha)	SE (trees/ha)	Mean (m^2/ha)	SE (m^2/ha)
1	18.5	8	2.3	122.3	52.4	46.3	17.6
2	7.1	7	1.0	160.1	30.2	25.7	4.8
3	5.3	6	0.9	222.9	51.4	108.3	8.7
4	10.4	12	8.6	137.2	31.2	70.8	8.5
5	10.2	9	1.1	189.4	32.1	113.3	14.1
6	5.9	7	0.8	257.6	98.8	104.3	24.6
7	4.8	6	0.8	256.7	25.3	41.7	4.0
8	17.1	8	2.1	269.4	39.0	117.5	25.5
9	27.4	15	1.8	171.0	24.1	117.3	14.0
10	16.9	9	1.9	197.9	47.2	68.9	9.5
11	10.5	11	1.0	240.1	42.6	89.1	7.4
12	9.1	9	1.0	594.8	42.6	128.9	7.4
13	10.9	9	1.2	242.0	42.3	65.6	9.7

BA = basal area. SD = stem density. SE = standard error.

where d_{ij} is the distance between the i^{th} target pixel and the j^{th} reference pixel, and l indexes the L spectral transformations, X, that define the feature space. The k-NN prediction for the m^{th} variable for the i^{th} target pixel is

$$\hat{y}_{i,m} = \frac{1}{k}\sum_{j=1}^{k} y_{j,m} \ , \qquad (2)$$

where j indexes the k nearest neighbor pixels in the reference set closest to the target pixel, m indexes the variables, and $y_{j,m}$ is the observation of the m^{th} variable for the subplot associated with the j^{th} nearest neighbor reference pixel.

Selection of Feature Space Variables

As noted by McRoberts *et al.* (2002), the inclusion in feature space of covariates unrelated to the dependent variables may be detrimental to the overall accuracy of predictions. Identification of the optimal set of feature space variables typically entails use of a bootstrapping technique in which the observations from a subset of the reference set are used to calculate predictions for the other subset. Accuracy is assessed by comparing the observations and predictions for the second subset. The leave-one-out technique is a special case of bootstrapping in which a prediction is calculated for each element of the reference set using the remaining reference set observations (Lachenbruch and Mickey 1986). Accuracy is assessed for continuous variables by comparing the observations and predictions for all reference set elements using root mean square error (RMSE) or a similar measure. When the reference set is large and the number of candidate feature space covariates is also large, an exhaustive evaluation of all covariate combinations can be extremely time-consuming.

An alternative to an exhaustive search of all covariate combinations is based on the genetic algorithm technique described by Tomppo (2004). Although this algorithm does not guarantee the optimal selection of covariates, it greatly reduces the selection time. A modification of this technique was used as follows:

1. Construct a random combination of covariates by independently selecting a random number, X, from a uniform (0, 1) distribution for each candidate feature space variable; if X > 0.5, include the covariate as a feature space variable in the combination; construct 100 such combinations.
2. Calculate RMSE for each of the 100 combinations using the leave-one-out technique.
3. Order the combinations by RMSE.
4. Select the 10 combinations with smallest RMSEs and construct 100 new combinations by pairwise crossing each combination with each other combination as follows:
    a. If a variable was included in both combinations of a pair, include it in the new combination.
    b. If a variable was not included in either combination of a pair, exclude it from the new combination.
    c. If a variable was included in one but not both combinations of a pair, randomly select a number, X, from a uniform (0, 1) distribution and include the variable in the new combination if X > 0.5.
    d. Following steps 4a through c, introduce random changes by randomly selecting a variable and reversing its inclusion or exclusion for the new combination; these changes are introduced only in new combinations constructed from pairs in which the first combination has greater RMSE from step 2 than the second combination.
5. Return to step 2.

Repeat steps 2 through 5 until the top 10 combinations determined in step 3 stabilize. Note that a pairwise crossing of a combination with itself in step 4 introduces no change and ensures that each of the top 10 combinations from the previous iteration enter into the next iteration.

The modified genetic algorithm was used to select covariate combinations for the feature space for SD and BA separately and in combination. When used to select combinations for SD and BA separately, the corresponding RMSE, denoted $RMSE_{min}$, was noted for each variable. The genetic algorithm was then used to determine a covariate combination for the two variables together. In step 2, RMSE was calculated separately for each variable for each combination. In step 3, for each variable, the ratio of the step 2 RMSE and its corresponding $RMSE_{min}$ were calculated and the average of the two ratios was used to order the combinations. Thus, the two variables, SD and BA, are weighted equally.

Stand-Level Estimation

Two approaches to k-NN stand-level estimation were investigated. The first approach entailed calculating estimates of SD and BA means for each stand by simply averaging the k-NN pixel predictions for all pixels with centers in the stand. For each pixel, the k-NN prediction for the m^{th} variable was calculated using (2), and for each stand the estimates of SD and BA means were calculated as

$$\overline{Y}_{kNN,m} = \frac{1}{N}\sum_{i=1}^{N}\hat{y}_{i,m} \ , \qquad (3)$$

where m denotes the variable, $\hat{y}_{i,m}$ is as defined in (2), i indexes pixels with centers in the stand, and N is the number of pixels in the stand.

The second approach exploits two assumptions that may contribute to improving k-NN pixel predictions. The first assumption is that stands exhibit internal homogeneity with respect to tree-level attributes such as species, diameter, and height, and stand-level attributes such as forest type, SD, and BA. The second assumption relates to the distribution of forest attributes for nearest neighbor reference pixels for a stand's target pixels. The relationship between forest attributes and the spectral values of satellite imagery is usually many-to-one in the sense that pixels with multiple forest conditions may be associated with similar sets of spectral values. Despite this many-to-one relationship, the second assumption is that most reference pixels selected as nearest neighbors for a target pixel have forest conditions similar to the actual conditions associated with that target pixel. These two assumptions may be used to modify the k-NN technique in four steps:

1. For each target pixel in a stand, select $m > k$ nearest neighbor reference pixels where k is the number intended to be used to calculate the k-NN predictions.
2. Determine the distribution of forest attributes for all m nearest neighbor reference pixels for all stand target pixels.
3. For each target pixel, create two suborderings of nearest neighbor reference pixels, one for pixels in the central portion of the step 2 distribution, and one for pixels in the tails of the distribution.
4. For each target pixel, modify the ordering of the nearest neighbor reference pixels by appending the pixels in the tails of the step 2 distribution to the end of ordering of pixels in the central portion of the distribution.

An example illustrates the modification. Suppose the stand-level SD distribution for all nearest neighbor reference pixels constructed in step 2 is as depicted in figure 2a. For a selected target pixel, neighbors 1, 2, and 4 of the original ordering of the $m = 5$ neighbors are in the central portion of the distribution, and neighbors 3 and 5 are in the tails of the distribution (fig. 2b). Thus, neighbors 3 and 5 are appended to the end of the ordering of neighbors 1, 2, and 4 to produce a modified ordering. For $k = 3$, the k-NN prediction for this pixel would then be based on original neighbors 1, 2, and 4 rather than on original neighbors 1, 2, and 3. Pixel-level predictions are calculated using (2) with the modified ordering of the nearest neighbors. Stand-level estimates are calculated using (3).

Figure 2a.—*Distribution of stem density for nearest neighbor pixels.*

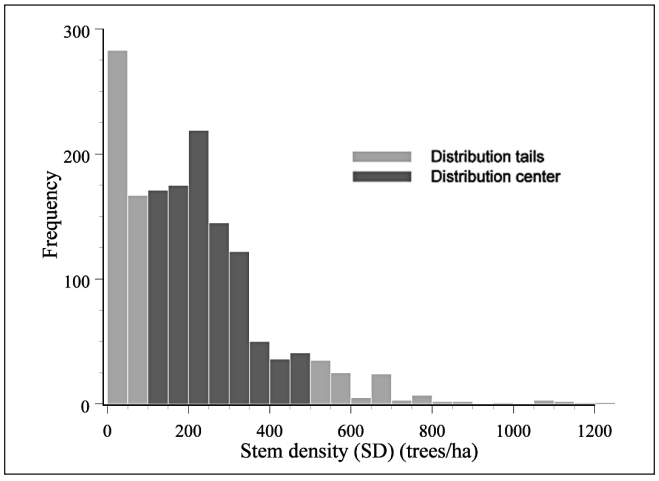
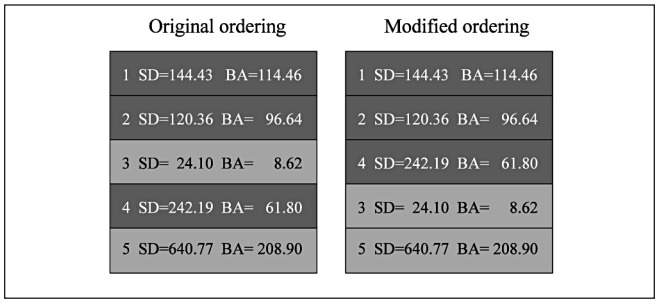


Figure 2b.—*Original and modified orderings of nearest neighbor reference pixels for a single target pixel.*



Validation Estimates

For each stand, the validation data were used to estimate stand-levels means as

$$\overline{Y}_{val,m} = \frac{1}{n}\sum_{i=1}^{n} y_{i,m}\ , \qquad (4)$$

where *val* indicates validation estimates, *m* designates the variable, *n* is the number of plots in the stand, and *i* indexes plots. The standard error of the mean was calculated as

$$SE\left(\bar{Y}_{val,m}\right) = \sqrt{\frac{\sum_{i=1}^{n}\left(y_{i,m} - \bar{Y}_{val,m}\right)^2}{n\left(n-1\right)}}. \qquad (5)$$

Comparisons

For each variable, estimates of the stand-level validation means were graphed against the corresponding original and modified k-NN estimates. Vertical lines extending 2-standard errors in both directions from the validation estimates were also graphed against the corresponding original and modified k-NN estimates. Under the hypothesis of no difference between the validation and k-NN stand-level estimates, the 2-standard error vertical lines should intersect the 1:1 line. As a method for comparing the validation and k-NN estimates, this graphic does not accommodate the uncertainty in the k-NN estimates and, therefore, may erroneously indicate statistically significant differences when none exists. Because the intent is to demonstrate that the k-NN estimates are not statistically significantly different than the validation estimates, this comparison may be characterized as conservative.

The preferred alternative to obtaining estimates of all stand-level means from ground plot observations is to produce remote sensing-based estimates that are sufficiently unbiased and precise to serve as the basis for stand management decisions. Even if the estimates are not sufficiently unbiased and precise for this purpose, however, they still may be useful for discriminating among stands with high likelihoods of requiring immediate treatment decisions or management inventories and those for which decisions may be delayed. If so, then the costly, labor-intensive, ground plot-based management inventories could be efficiently focused on the former class of stands.

A second approach to comparing the validation and k-NN estimates of the stand-level means was based on their utility for ordering stands with respect to the likelihood of requiring immediate management decisions. The assumption underlying this approach is that stands with a combination of large SD and large BA are more likely to require immediate decisions. For each variable, the 13 validation estimates of stand-level means and the two sets of 13 k-NN estimates were all divided by the largest validation estimate over the 13 stands. This calculation standardizes the estimates of the SD and BA means and permits them to be combined in a manner in which they are equally weighted. For each stand, the Euclidean distance from SD = 0 and BA = 0 of the standardized validation and k-NN estimates was calculated as

$$d = \sqrt{\overline{BA}_{std}^2 + \overline{SD}_{std}^2}\ , \qquad (6)$$

where the subscript *std* indicates standardized estimates. Stands with greater distances from SD = 0 and BA = 0 are interpreted as having greater combinations of SD and BA and, therefore, being in greater need of immediate management decisions. Orderings of stands based on this measure using the validation and the two sets of k-NN stand-level estimates were compared.

Results

For this study, the covariates selected for feature space using the genetic algorithm stabilized in five to eight iterations. Although some improvement in RMSE was achieved by selecting a subset of the 12 covariates for SD and BA separately and in combination, the improvement was not substantial (table 2). The five best combinations based on SD separately and in combination with BA had similar RMSEs; similarly, the five best combinations for BA separately and in combination with SD also had similar RMSEs. The RMSEs corresponding to the best covariate combinations selected when considering SD and BA simultaneously were nearly the same as for the RMSEs corresponding to combinations selected considering SD and BA separately. Both consistencies and inconsistencies were evident in the selection of covariates for feature space. Among the 12 spectral transformation covariates, summer brightness, spring

Table 2.—*Root mean square error.*

Variable combination rank	Separate		Combined		
	SD	BA	SD	BA	Proportion
1	182.1047	51.1821	182.3704	51.1935	1.0008
2	182.2339	51.2398	182.6407	51.2513	1.0021
3	182.6833	51.2444	182.2340	51.3858	1.0023
4	182.7614	51.2507	182.1047	51.4481	1.0026
5	182.8328	51.2438	182.8729	51.2538	1.0028
Best BA combination			184.0244	51.1821	1.0053
Best SD combination			182.1047	51.4481	1.0026
All 12 variables	184.3443	51.9538	184.3443	51.9538	1.0137

BA = basal area. SD = stem density.

and fall greenness, fall wetness, and summer and fall NDVI were selected in at least four of the five best combinations for SD, BA, and SD and BA simultaneously; spring NDVI was seldom selected for any of the five best combinations.

For the modified k-NN technique, the best results were obtained when using the distribution of SD based on the m = 5 nearest neighbors, when defining the central portion of the distribution to be between the 12th and 96th percentiles, and when using k = 3. Graphs of the validation estimates of the stand-level SD and BA means against the corresponding k-NN estimates for the 13 stands showed slightly better results when using the modified k-NN approach. For SD, 2-standard error confidence intervals for 12 of 13 stands cross the 1:1 line when using the modified k-NN approach, while 11 of 13 confidence intervals cross the line when using the original k-NN technique (fig. 3a). For BA, 2-standard error confidence intervals for 9 of 13 stands cross the 1:1 line when using the modified k-NN approach while 8 of 13 confidence intervals cross the line when using the original k-NN approach (fig. 3b).

The orderings of the stands obtained using the original and modified k-NN techniques were similar (table 3), although the modified technique produced slightly superior results relative to the ordering based on the validation data. An example from table 3 illustrates this analysis. The four highest ranked stands based on the largest estimates of SD and BA means from the validation data were stands 12, 8, 9, and 5; the four highest

Figure 3a.—*Validation versus modified k-Nearest Neighbor estimates of stand-level stem density means; each point represents a different stand; vertical lines are 95-percent confidence intervals.*

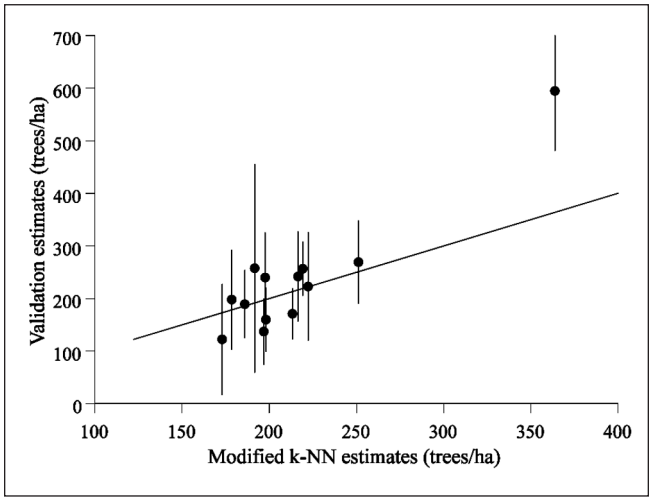


Figure 3b.—*Validation versus modified k-Nearest Neighbor estimates of stand-level basal area means; each point represents a different stand; vertical lines are 95-percent confidence intervals.*

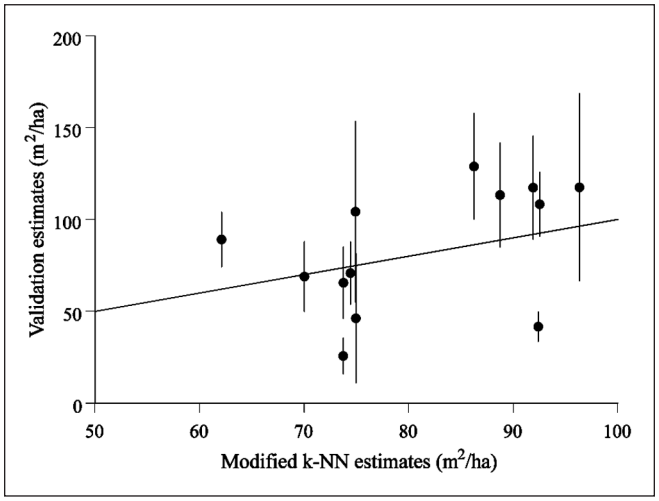


Table 3.—*Stand ranking statistics.*

Stand rank	Stand number (basis for ranking)			Cumulative proportion correctly ranked using k-NN estimates	
	Validation estimates	k-NN estimates			
		Original	Modified	Original	Modified
1	12	12	12	1.00	1.00
2	8	8	8	1.00	1.00
3	9	13	3	0.67	0.67
4	5	7	7	0.50	0.75
5	3	9	9	0.60	0.80
6	6	3	5	0.67	0.83
7	11	11	13	0.86	0.71
8	13	2	4	0.75	0.75
9	10	10	6	0.78	0.78
10	4	4	2	0.90	0.80
11	7	6	1	0.91	0.91
12	1	5	10	0.92	0.92
13	2	1	11	1.00	1.00

k-NN = k-Nearest Neighbor.

ranked stands using the original k-NN technique were stands 12, 8, 13, and 7; and the four highest ranked stands using the modified k-NN technique were stands 12, 8, 3, and 7. Thus, compared to the four highest ranked stands using the validation data, the modified k-NN technique correctly selected three stands for a proportion of 0.75, while the original k-NN technique correctly selected two stands for a proportion of 0.50.

Conclusions

Three conclusions may be drawn from this study. First, for these data, the genetic algorithm did not select covariates for the feature space that substantially decreased RMSEs. The similarity in RMSEs for the best covariate combinations when considering SD and BA separately and when considering them simultaneously can probably be at least partially attributed to this result. Second, the estimates of stand-level SD means obtained using the modified k-NN technique were nearly always within 95-percent confidence intervals of SD means obtained from the validation data. The estimates of stand-level BA means were less similar to the validation means, although the proportion within the confidence intervals was still great than 0.70. The somewhat lower proportion for BA can perhaps be attributed to two factors: (1) BA is essentially a below canopy attribute while optical sensors such as Landsat primarily respond to sunlight reflected from the forest canopy, and (2) the validation data were obtained from variable radius plots, while the data used to train the satellite imagery were obtained from FIA fixed radius plots. Because variable radius plots use BA as a factor for selecting trees to measure while fixed radius plots do not, estimates of BA obtained using these two different sampling approaches may not be entirely compatible. Third, the similarity in estimates of the stand-level SD and BA means obtained using the validation data and the means obtained using the TM imagery, the FIA plot data, and the k-NN technique suggest possibilities for effectively bridging the gap between strategic and management inventories. Even if the k-NN estimates of stand-level means are not deemed suitable to constitute the basis for stand management decisions, the high proportions of correctly ranked stands suggest that the k-NN estimates of means may be used to inform decisions as to which stands require immediate management inventories.

Future research should be conducted to extend this study in several directions. First, additional validation data should be obtained to expand the study beyond 13 stands. Second, additional studies should be conducted in other forest types. Third, validation data should be obtained using fixed radius plots of the size and configuration of FIA subplots.

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.

Crist, E.P.; Cicone, R.C. 1984. Application of the tasseled cap concept to simulated Thematic Mapper data. Photogrammetric Engineering and Remote Sensing. 50: 343-352.

Flewelling, J. 2009. Forest inventory predictions from individual tree crowns: regression modeling within a sample framework. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.A.; McWilliams, W.H., eds. Proceedings, eighth annual forest inventory and analysis symposium: 203-208.

Kauth, R.J.; Thomas, G.S. 1976. The tasseled cap—a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. Proceedings, symposium on machine processing of remotely sensed data. West Lafayette, IN: Purdue University: 41-51.

Lachenbruch, P.A.; Mickey, M.R. 1986. Estimation of error rates in discriminant analysis. Technometrics. 10: 1-11.

McRoberts, R.E.; Bechtold, W.A.; Patterson, P.L.; Scott, C.T.; Reams, G.A. 2005. The enhanced Forest Inventory and Analysis program of the USDA Forest Service: historical perspective and announcement of statistical documentation. Journal of Forestry. 103(6): 304-308.

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

McRoberts, R.E.; Tomppo, E.O.; Finley, A.O.; Heikkinen, J. 2007. Estimating areal means and variances of forest attributes using the k-Nearest Neighbors technique and satellite imagery. Remote Sensing of Environment. 111: 466-480.

Tomppo, E.; Halme, M. 2004. Using coarse scale forest variables as ancillary information and weighting of information in k-NN estimation: a genetic algorithm. Remote Sensing of Environment. 92: 1-20.

Estimating Uncertainty in Map Intersections

Ronald E. McRoberts[1], Mark A. Hatfield[2], and Susan J. Crocker[3]

Abstract.—Traditionally, natural resource managers have asked the question "How much?" and have received sample-based estimates of resource totals or means. Increasingly, however, the same managers are now asking the additional question "Where?" and are expecting spatially explicit answers in the form of maps. Recent development of natural resource databases, access to satellite imagery, development of image classification techniques, and availability of geographic information systems has facilitated construction and analysis of the required maps. Unfortunately, methods for estimating the uncertainty associated with map-based analyses are generally not known, particularly when the analyses require maps to be combined. Motivated by the threat of the emerald ash borer in southeastern Michigan, the number of ash trees at risk was estimated by intersecting a forest/non-forest map and an ash tree distribution map. The primary objectives of the study were to quantify the uncertainty of the estimate and to partition the uncertainty by source. An important conclusion of the study is that spatial correlation—an often ignored component of uncertainty analyses—made the greatest contribution to the uncertainty in the estimate of the total number of ash trees.

Introduction

The emerald ash borer (*Agrilus planipennis* Fairmaire, Coleoptera: Buprestidae) (EAB) is a wood-boring beetle native to Asia that was initially discovered in the United States in June 2002. It most likely entered the country in solid-wood packing material such as crates and pallets and was transported to Detroit, Michigan, at least 10 years before it was discovered there in 2002 (Cappaert *et al.* 2005, Herms *et al.* 2004). Ash trees are the only known host, and damage is the result of larval activity. Once eggs hatch, larvae bore into the cambium and begin to feed on and produce galleries in the phloem and outer sapwood. Larval feeding disrupts the translocation of water and nutrients and eventually girdles the tree. Tree mortality occurs within 1 to 3 years, depending on the severity of the infestation (McCullough and Katovich 2004, Haack *et al.* 2002). All of Michigan's native ash species (*Fraxinus* spp.) and planted cultivars are susceptible (Cappaert *et al.* 2005). Since 2002, southeastern Michigan has lost an estimated 15 million ash trees due to the EAB (Cappaert *et al.* 2005). The natural rate of EAB dispersal is estimated to be less than 1 km per year in low-density sites. Natural dispersal has been enhanced by human transportation of infested firewood, ash logs, and nursery stock. This artificial spread of EAB has initiated the majority of outlier infestations (Cappaert *et al.* 2005). Continued spread outside of the core zone increases the threat to ash trees across the United States.

The objective of the study was twofold: (1) to estimate the uncertainty in forest/nonforest maps, ash tree distribution maps, and intersections of the two maps, and (2) to partition the total uncertainty in areal estimates of the total number of ash trees by source.

Methods

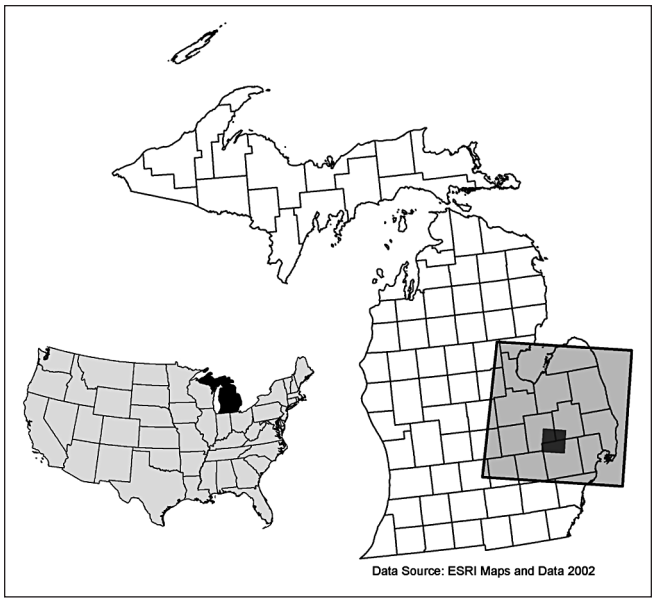
The motivating problem for the study was to calculate an estimate, A_{total}, of the total number of ash trees in a region of southeastern Michigan (fig. 1) that is susceptible to infestation

[1] Mathematical Statistician, U.S. Department of Agriculture (USDA) Forest Service, Northern Research Station, Forest Inventory and Analysis Unit, 1992 Folwell Avenue, St. Paul, MN 55108. E-mail: rmcroberts@fs.fed.us.

[2] Forester, USDA Forest Service, Northern Research Station, Forest Inventory and Analysis Unit, 1992 Folwell Avenue, St. Paul, MN 55108. E-mail: mahatfield@fs.fed.us.

[3] Forester, USDA Forest Service, Northern Research Station, Forest Inventory and Analysis Unit, 1992 Folwell Avenue, St. Paul, MN 55108. E-mail: scrocker@fs.fed.us.

Figure 1.—*State of Michigan, USA with Landsat scene path 20, row 30 (large rectangle) and 30-km x 30-km study area (small rectangle).*



by the EAB. The estimation approach entails intersecting two 30-m x 30-m resolution maps, one depicting the spatial distribution of forest land and the other depicting the spatial distribution of ash trees. The technical objective was to estimate the uncertainty of A_{total} for a selected region. The map depicting the distribution of forest land was derived from a forest probability layer constructed using forest inventory plot observations, Landsat Thematic Mapper (TM) satellite imagery, and a logistic regression model. The ash tree distribution layer was constructed using the same forest inventory plot observations and inverse distance weighted spatial interpolation. Uncertainty in A_{total} was estimated using an analytical technique for estimating the covariances of the logistic regression model parameters, a sample-based technique for estimating the uncertainty in interpolated ash tree counts per hectare, and Monte Carlo techniques for generating forest/nonforest maps and for combining the components of uncertainty.

Data

The study area is wholly contained in Landsat scene path 20, row 30 (fig. 1), for which three dates of Landsat TM/ETM+ imagery were obtained: May 2002, July 2003, and October 2000. Preliminary analyses indicated that Normalized Difference Vegetation Index (NDVI) and the tassel cap (TC) transformations (brightness, greenness, and wetness) (Kauth and Thomas 1976, Crist and Cicone 1984) were superior to both the spectral band data and principal component transformations with respect to predicting forest attributes. Thus, the predictor variables were the 12 satellite image-based variables consisting of NDVI and the three TC transformations for each of the three image dates. Mapping units for all analyses consisted of the 30-m x 30-m TM pixels.

The Forest Inventory and Analysis (FIA) program of the Forest Service, U.S. Department of Agriculture has established field plot centers in permanent locations using a sampling design that is assumed to produce a random, equal probability sample (Bechtold and Patterson 2005, McRoberts *et al.* 2005). The plot array has been divided into five nonoverlapping, interpenetrating panels, and measurement of all plots in one panel is completed before measurement of plots in the next panel is initiated. Panels in the study area are selected on a 5-year rotating basis. Over a complete measurement cycle, the Federal base sampling intensity is approximately one plot per 2,400 ha. The State of Michigan provided additional funding to triple the sampling intensity to approximately one plot per 800 ha. In general, locations of forested or previously forested plots were determined using global positioning system receivers, while locations of nonforested plots were determined using aerial imagery and digitization methods. Each plot consists of four 7.31-m radius circular subplots. The subplots are configured as a central subplot and three peripheral subplots with centers located at a distance of 36.58 m and azimuths of 0, 120, and 240 degrees from the center of the central subplot. Data for 2,995 FIA plots or 11,980 subplots with centers in the selected TM scene that were measured between 2000 and 2004 were available for the study. For each subplot, the proportion of the subplot area that qualified as forest land was determined from field crew observations. The FIA program requirements for forest land are at least 0.4 ha in size, at least 10 percent stocking, at least 36.58 crown-to-crown width, and forest land use. For each subplot, the number of observed ash trees with diameter at breast height of at least 12.5 cm was scaled to a count/hectare basis. The ash tree count/hectare for the i^{th} subplot is denoted A_i^o, where the superscript denotes a subplot observation and

is distinguished from the ash tree count/hectare for the i^{th} pixel which is denoted A_i.

The spatial configuration of the FIA subplots with centers separated by 36.58 m and the 30-m x 30-m spatial resolution of the TM/ETM+ imagery permits individual subplots to be associated with individual image pixels. The subplot area of 167.87 m^2 is an approximately 19-percent sample of the 900 m^2 pixel area, and subplot observations are assumed to adequately characterize entire pixels.

Areal Estimation

The estimate, A_{total}, for a region was calculated in three steps: (1) generate a 30-m x 30-m resolution forest/nonforest map; (2) construct a 30-m x 30-m ash tree count/hectare layer, and for each pixel, estimate the number of ash trees as the product of the ash tree count/hectare and the 0.09 ha pixel area; and (3) estimate A_{total} as the sum over forest pixels from step 1 of pixel-level estimates of the number of ash trees from step 2. Thus, two layers were required: a forest/nonforest layer and an ash tree count/hectare layer. Both layers were constructed specifically for this study to make known their pixel-level uncertainties. Two sets of analyses were conducted. First, forest/nonforest, ash tree count/hectare, and the two associated uncertainty maps were constructed for a 30-km x 30-km study area in the selected TM scene (fig. 1). Second, uncertainty analyses for A_{total} were conducted for a smaller 2-km x 2-km portion of the larger study area (figs. 2 and 3). The restriction of the latter analyses to the smaller area was due to technological constraints as is noted in a later section.

Forest/Nonforest Layer

Because satellite image pixels with different ground covers often have similar spectral signatures, assignment of classes to individual pixels is often probability based. A layer depicting the probability of forest was constructed using a logistic regression model (McRoberts 2006),

$$E(p_i) = f(X_i;\beta) = \frac{\exp\left(\beta_0 + \sum_{j=1}^{12}\beta_j x_{ij}\right)}{1+\exp\left(\beta_0 + \sum_{j=1}^{12}\beta_j x_{ij}\right)}, \quad (1)$$

where p_i is the probability of forest for the i^{th} pixel, X_i is the vector of the 12 spectral transformations for the i^{th} pixel with x_{ij} being the j^{th} element, the βs are parameters to be estimated, $exp(.)$ is the exponential function, and $E(.)$ denotes statistical expectation. When estimating the parameters of (1), only data for the 7,920 completely forested or completely nonforested subplots were used. The covariance matrix for the vector of parameter estimates was estimated analytically as,

$$\hat{Var}\left(\hat{\beta}\right) = \left(Z' V_\varepsilon^{-1} Z\right)^{-1}, \quad (2)$$

where the elements of the matrix Z are defined as,

$$z_{ij} = \frac{\partial f\left(X_i;\hat{\beta}\right)}{\partial \beta_j},$$

the elements of V_e are defined as,

$$v_{ij} = \sqrt{\hat{p}_i\left(1-\hat{p}_i\right)}\sqrt{\hat{p}_j\left(1-\hat{p}_j\right)}\hat{\rho}_{ij},$$

and $\hat{\rho}_{ij}$ is the spatial correlation among the standardized residuals estimated using a variogram (McRoberts 2006).

The most probable forest/nonforest classification of the imagery is constructed by comparing the probability, $\hat{p}$, from (1) for each pixel to 0.5: if $\hat{p} \geq 0.5$, the pixel is classified as forest and assigned a numerical value of 1; if $\hat{p} < 0.5$, the pixel is classified as nonforest and assigned a numerical value of 0; however, because the assignment of forest or nonforest to pixels is based on probabilities, it is uncertain whether this procedure correctly assigns forest or nonforest to individual pixels. Forest/nonforest realizations that reflect the uncertainty in the classification were obtained using a four-step procedure designated Procedure A:

A1. Using the procedure of Kennedy and Gentle (1980: 228-213), generate a vector of random numbers from a multivariate Gaussian distribution with mean 0 and covariance $Var\left(\hat{\beta}_i\right)$ from (2); add these random numbers to the logistic regression model parameter estimates to obtain simulated parameter estimates.

A2. Using the simulated parameter estimates from step 1 with (1), calculate a probability, $\hat{p}$, of forest for each pixel.

A3. For each pixel, generate a random number, r, from a uniform [0, 1] distribution; if $r \leq \hat{p}$, the pixel is designated

forest with a numerical value of 1; if $r > \hat{p}$, the pixel is designated nonforest with a numerical value of 0.

A4. Repeat steps A1 through A3 many times and calculate the mean and variance of the numerical values assigned to each pixel.

The variance of the forest/nonforest classifications for each pixel is a measure of the uncertainty of the pixel's classification.

Ash Tree Distribution Layer

Because the ash tree distribution layer was for forest land, only data for the 1,953 FIA forest subplots were used in its construction. Using these plot data, an empirical variogram was constructed and an exponential variogram model was fit to the data. An interpolated surface was constructed for which the ash tree count/hectare, A_i, for the i^{th} pixel was estimated as:

$$\hat{A}_i = \sum_{j=1}^{1953} \frac{w_j}{W_i} A_j^o , \qquad (3a)$$

where

$$w_{ij} = \begin{cases} \frac{1}{v_{ij}} - \frac{1}{\hat{\lambda}} & \text{if } \frac{v_{ij}}{\hat{\lambda}} \le 0.95 \\ 0 & \text{if } \frac{v_{ij}}{\hat{\lambda}} > 0.95 \end{cases} ; \qquad (3b)$$

v_{ij} is the predicted covariance from the variogram model corresponding to the distance, d_{ij}, between the i^{th} pixel and the j^{th} plot; $\hat{\lambda}$ is the estimate of the variogram sill; and

$$W_i = \sum_{j=1}^{1953} w_j . \qquad (3c)$$

The variance of $\hat{A}_i$ was estimated as:

$$\hat{Var}\left(\hat{A}_i\right) = \sum_{j=1}^{1953} \left(\frac{w_{ij}}{W_i}\right)^2 \left(A_j^o - \hat{A}_i\right)^2, \qquad (4)$$

and is a measure of the uncertainty of the estimate, $\hat{A}_i$. Other equally valid approaches such as kriging could have been used to construct the ash tree county/hectare layer. Realizations of the ash tree count/hectare distribution were obtained by selecting for each pixel a random number from a normal distribution with mean 0 and variance, $Var\left(\hat{A}_i\right)$ from (4), and adding the number to $\hat{A}_i$ from (3a).

Uncertainty Estimation

Uncertainty in the areal estimate, A_{total}, is due to contributions from four sources: (1) uncertainty in the logistic regression model parameter estimates; (2) uncertainty in the pixel-level forest/nonforest classifications, given a set of parameter estimates; (3) uncertainty in the interpolated pixel-level ash tree count/hectare estimates, $\hat{A}$; and (4) spatial correlation in forest/nonforest and ash tree observations not accommodated in the logistic regression model predictions and the ash tree count/hectare interpolated estimates.

The spatial correlation contribution to uncertainty in A_{total} results from two phenomena. First, forest areas tend to be clustered rather than independently and randomly distributed throughout the landscape. Thus, to mimic natural conditions, forest/nonforest realizations generated from the logistic regression model predictions of forest probability should exhibit clustering comparable to that observed among the FIA subplot observations of forest and nonforest. This feature requires that the random numbers used to generate the forest/nonforest realization in step A3 be drawn from a correlated uniform [0, 1] distribution. Second, the errors obtained as the differences between A^o and $\hat{A}$ were expected to be spatially correlated; i.e., if an interpolation, $\hat{A}$, overestimates its true value, other interpolations in close spatial proximity would be expected to overestimate their true values also. For this investigation, however, the range of spatial correlation for the interpolation errors was only slightly more than the 30 m pixel width, regarded as negligible, and ignored.

To generate random numbers from an appropriately correlated uniform [0, 1] distribution as required to accommodate spatial correlation, an eight-step procedure designated Procedure B was used:

B1. Construct an empirical variogram,

$$\hat{\gamma} = \frac{2}{2n} \sum_{n} \left(F_i - F_j\right)^2$$

where F is the numerical designation for forest or nonforest subplot observations, and $n(d)$ denotes a collection of pairs, (F_i, F_j), whose Euclidean distances in geographic space place are within a given neighborhood of d.

B2. Fit an exponential variogram,

$$\hat{\gamma}(d) = \hat{\alpha}_1 \left(1 - \exp\left(-\hat{\alpha}_2 d\right)\right), \quad (5)$$

to the empirical variogram from step B1, where the estimate of the range of spatial correlation is

$$\frac{\ln(0.05)}{\hat{\alpha}_2}.$$

B3. Construct a spatial correlation matrix by assigning to each pixel pair, (i,j), a correlation, r_{ij}, calculated as,

$$\rho_{ij} = \exp\left(\gamma d_{ij}\right),$$

where d_{ij} is the distance between the i^{th} and j^{th} pixel centers and, initially, $\gamma = \hat{\alpha}_2$ from step B2.

B4. Generate a vector of random numbers, one for each pixel, from a multivariate Gaussian distribution with the correlation structure constructed in step B3 using the technique described by Kennedy and Gentle (1980: 228-231).

B5. Convert the Gaussian random numbers from step B4 to Gaussian cumulative frequencies, resulting in a correlated uniform [0, 1] distribution.

B6. Generate a forest/nonforest realization using Procedure A with the correlated uniform [0, 1] distribution from step B5.

B7. Construct an empirical variogram of the forest/nonforest realization from step B5; fit an exponential variogram model; and estimate the range of spatial correlation as in step B2.

B8. Repeat steps B3–B7, adjusting the γ parameter in step B3 each iteration until the range of spatial correlation from step B7 is close to that obtained in step B2.

The exponential variogram model was used in step B2 because of its simplicity and the adequacy of the fit to the data. Construction of the multiviariate Gaussian distribution in step B4 requires the Cholesky decomposition of a covariance matrix corresponding to the correlation matrix constructed in step B3. For a square region, n pixels on a side, the correlation matrix will be n^2 x n^2. Thus, the 30 km x 30 km study area, which has 1,000 TM pixels on a side, would require decomposition of a 10^6 x 10^6 matrix. To accommodate personal computer space and processing limitations, analyses involving spatial correlations were constrained to a 2-km x 2-km region, which is approximately 67 pixels on a side and requires decomposition of a smaller 4,489 x 4,489 matrix (figs. 2 and 3).

Uncertainty in A_{total} for the 2-km x 2-km region was estimated using a six-step Monte Carlo simulation procedure designated Procedure C:

C1. Generate random numbers from a multivariate Gaussian distribution with mean 0 and variance matrix, $Var\left(\hat{\beta}_i\right)$ from (2); add these random numbers to the logistic regression model parameters estimates to obtain simulated parameter estimates; calculate the probability, $\hat{p}$, of forest for each pixel using the simulated parameter estimates with (1).

C2. For each pixel, generate a random number, r, from a correlated uniform [0, 1] distribution using Procedure B; if $r \le \hat{p}$, designate the pixel as forest; if $r > \hat{p}$, designate the pixel as nonforest.

C3. Calculate the total forest area, F_{total}, as the product of the number of forest pixels from step C2 and the 0.09 ha pixel area.

C4. For each pixel, generate a random number from a normal distribution with mean 0 and variance, $Var\left(\hat{A}_i\right)$ from (4); add the random number to the interpolated estimate of ash tree count/hectare, $\hat{A}$, to obtain a simulated ash tree count/hectare; multiply the simulated ash tree count/hectare and the 0.09 hectare pixel area to obtain a simulated ash tree count for the pixel.

C5. Estimate A_{total} as the sum of the simulated ash tree counts from step C4 for forest pixels from step C2;

C6. Repeat steps C1 through C5 many times; calculate the mean and variance of F_{total} and A_{total} over all repetitions; estimate the uncertainties in F_{total} and A_{total} as $\hat{Var}\left(F_{total}\right)$ and $\hat{Var}\left(A_{total}\right)$, respectively.

In Procedure C, the contribution of uncertainty due to the logistic regression model parameter estimates may be excluded by not adding uncertainty in step C1; the contribution of uncertainty in the classification, given the parameter estimates, may be

excluded by skipping step C2 and comparing the probabilities generated in step C1 to 0.5 using Procedure A; the contribution of uncertainty due to spatial correlation may be excluded by generating an uncorrelated uniform [0, 1] distribution in step C2; and the contribution of uncertainty due to the interpolated ash tree counts/hectare may be excluded by not adding uncertainty in step C4. The magnitude of the contributions of individual sources of uncertainty may be estimated by considering $\hat{V}ar(F_{total})$ and $\hat{V}ar(A_{total})$ obtained by including contributions from all sources individually and in combinations.

Results

The analysis of Monte Carlo results indicated that all estimates stabilized to within less than 1 percent by 25,000 simulations (table 1). Therefore, 25,000 simulations were used when applying Procedure C to estimate the contributions of the various sources of uncertainty.

The forest/nonforest maps constructed using logistic regression model predictions produced realistic spatial distributions, although no independent accuracy assessment was conducted (fig. 2a); however, considerable detail was revealed in the uncertainty map; e.g., the field structure and road networks (fig. 2b). Considerably less detail was revealed in the ash tree count/hectare map, but this result was expected because of the fewer FIA plots available and the more continuous nature of the layer (fig. 3a). As expected with biological analyses, the greatest uncertainty in the latter map occurred in the same locations as the greatest estimated values (fig. 3b).

The estimates obtained using Procedure C dramatically revealed that the source of uncertainty making the greatest contribution to uncertainties in the estimates of both F_{total} and A_{total} was spatial correlation in the realizations of the forest/nonforest maps. The magnitude of this effect is highlighted by noting that when uncertainty from this source was included, 95-percent confidence intervals for both F_{total} and A_{total} included, or were close to including, 0. The contribution of the uncertainty in the underlying ash tree count/hectare layer to $\hat{V}ar(\hat{A}_{total})$ was much less than the contribution due to the uncertainty in the forest/nonforest layer.

Conclusions

Three conclusions may be drawn from this study. First, spatial correlation is a crucial contributor to uncertainty in map analyses that aggregate results over multiple mapping units. Ignoring this contribution inevitably leads to underestimates of variances and unwarranted statistical confidence in estimates. Unfortunately, the importance of this source of uncertainty is generally not known to those who conduct map-based analyses, and techniques for estimating its effects are generally unfamiliar. Second, researchers, authors, and university faculty should

Table 1.—*Monte Carlo simulation estimates from Procedure C.*

Source of uncertainty				Estimates			
				F_{tot}		A_{tot}	
Model parameter estimates	Classification	Spatial correlation	Ash tree count/ha interpolation	Mean (ha)	SE* (ha)	Mean (count)	SE* (count)
No uncertainty				113.40		2860.11	
X				110.93	25.37	2801.15	643.23
	X			104.22	1.87	2637.11	47.61
	X	X		104.04	55.32	2541.56	1180.34
X	X	X		105.58	60.89	2633.00	1549.91
			X	113.40		2860.55	47.57
X	X		X	105.53	18.97	2670.79	484.46
X	X	X	X	106.25	61.92	2690.07	1576.47

** SE is standard error and is calculated as the square root of variance.*

Figure 2a.—*Forest/nonforest classification for 30-km x 30-km study area.*

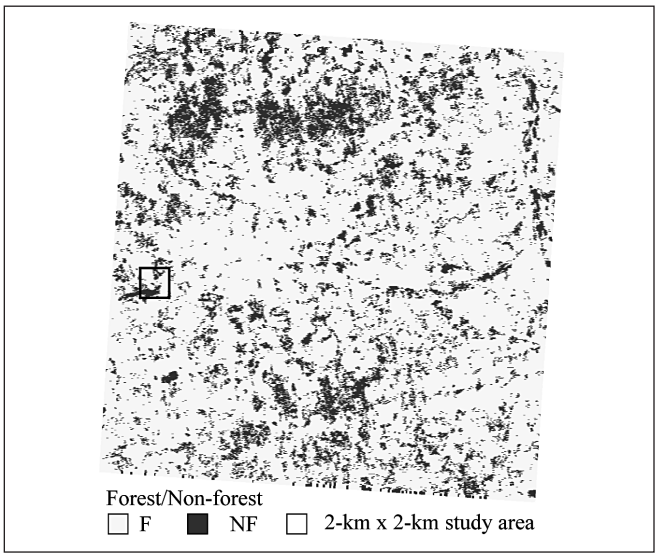


Figure 2b.—*Standard errors of pixel forest/nonforest classifications for 30-km x 30-km study area.*

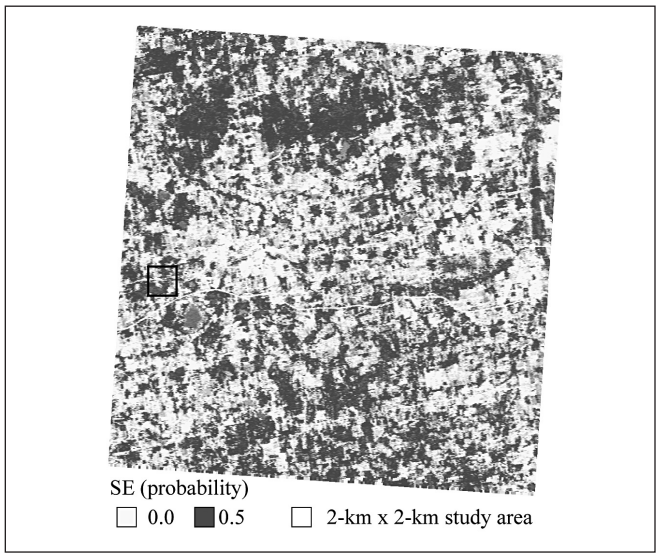


Figure 3a.—*Ash tree count/hectare interpolations for 30-km x 30-km study area.*

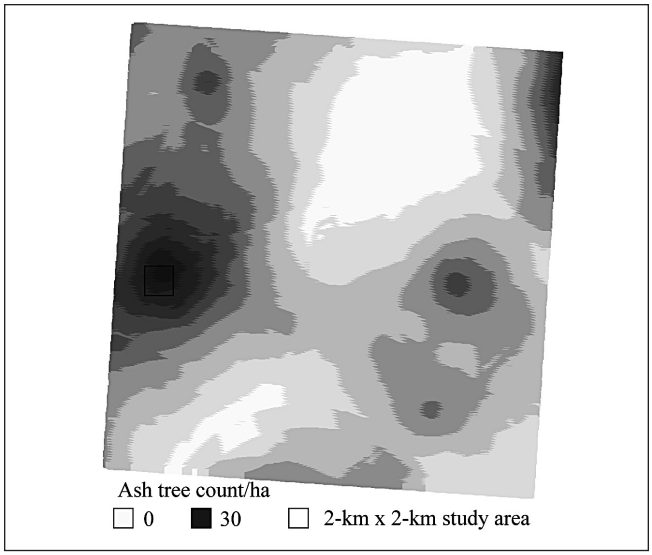
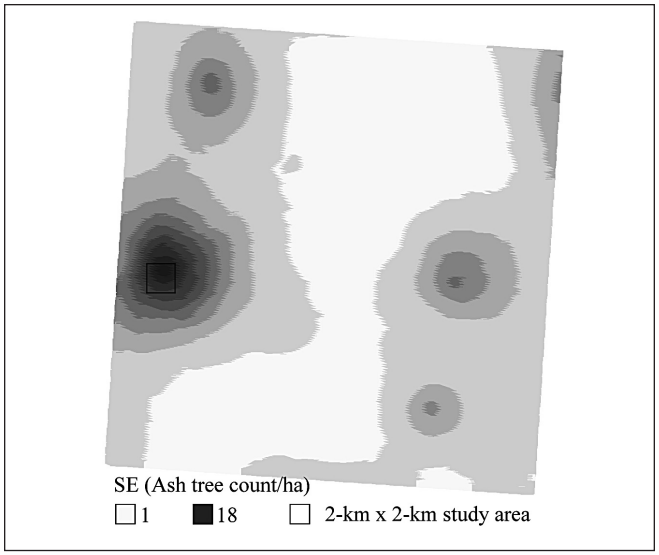


Figure 3b.—*Standard error of pixel ash tree count/hectare interpolations for 30-km x 30-km study area.*



give greater attention to uncertainty estimation. Third, estimation of uncertainty is not trivial, either conceptually or from a technical perspective. The necessity of decomposing very large matrices limits the size of regions that can be analyzed without high-speed computing facilities.

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.

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: 152-165.

Crist, E.P.; Cicone, R.C. 1984. Application of the tasseled cap concept to simulated Thematic Mapper data. Photogrammetric Engineering and Remote Sensing. 50: 343-352.

Haack, R.A.; Jendek, E.; Liu, H.; Marchant, K.R.; Petrice, T.R.; Poland, T.M.; Ye, H. 2002. The emerald ash borer: a new exotic pest in North America. Newsletter of the Michigan Entomological Society. 47: 1-5.

Herms, D.A.; McCullough, D.G.; Smitley, D.R. 2004. Under attack. American Nurseryman. http://www.emeraldashborer.info/educational.cfm.

Kauth, R.J.; Thomas, G.S. 1976. The tasseled cap—a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. Proceedings, symposium on machine processing of remotely sensed data. West Lafayette, IN: Purdue University: 41-51.

Kennedy, W.J., Jr.; Gentle, J.E. 1980. Statistical computing. New York: Marcel Dekker. 591 p.

McCullough, D.G.; Katovich, S.A. 2004. Emerald ash borer. NA-PR-02-04. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, State and Private Forestry, Northeastern Area. 2 p.

McRoberts, R.E. 2006. A model-based approach to estimating forest area. Remote Sensing of Environment. 103: 56-66.

McRoberts, R.E.; Bechtold, W.A.; Patterson, P.L.; Scott, C.T.; Reams, G.A. 2005. The enhanced Forest Inventory and Analysis program of the USDA Forest Service: historical perspective and announcement of statistical documentation. Journal of Forestry. 103(6): 304-308.

Sampling and Mapping Forest Volume and Biomass Using Airborne LIDARs

Erik Næsset[1], Terje Gobakken[2], and Ross Nelson[3]

Abstract.—Since around 1995, extensive research efforts have been made in Scandinavia to develop airborne Light Detection and Ranging (LIDAR) as an operational tool for wall-to-wall mapping of forest stands for planning purposes. Scanning LIDAR has the ability to capture the entire three-dimensional structure of forest canopies and has therefore proved to be a very efficient technique to determine biophysical properties such as stand volume and biomass. In Norway and Sweden, airborne scanning LIDAR is now used operationally to estimate merchantable volume on forest stands remotely across areas from 50 km^2 to 2,000 km^2. Complete scanning LIDAR coverage over larger regions (e.g., counties, States, provinces, and countries) is not economically feasible at this time due to data acquisition and processing costs. Despite these challenges, it has been demonstrated that airborne profiling LIDARs can provide reliable estimates of forest volume and biomass when used in a sampling mode (i.e., a number of flight lines are flown as linear, parallel transects separated by many kilometers). Scanning systems can be similarly employed for regional forest inventory by considering the flight lines as part of a strip sampling design. In this article, we report on a joint research effort by the Norwegian University of Life Sciences, National Aeronautics and Space Administration, Norwegian national forest inventory (NFI), Yale University, Swedish University of Agricultural Sciences, and Swedish NFI to develop and test airborne LIDARs as regional forest sampling tools.

Background

Ever since the first public Norwegian national forest inventory (NFI)—the first NFI in the world—took place 1919 through 1932, the Norwegian NFI has provided vital information of the timber resources. During its 87-year history, the NFI has produced statistics at the national and regional levels and has thus been an important prerequisite for the formation of a national policy for the management of the resources and the control of the policy's implementation in various regions of the country. Even though the main focus of the NFI over all these years has been on the quantification of the timber resources, additional characteristics of the ecosystems have been incorporated as the society has focused on "new" aspects, such as, (e.g.,) the preservation of biodiversity, the effects of air pollution, and the forests' role as sinks and sources for greenhouse gases.

Over the years, the NFI has applied various designs of field-based strip and plot surveys as basic sampling models, and the sampling density has been adjusted from county to county to provide reliable estimates at national and regional (county) levels over time. Currently, the NFI is facing at least three major challenges: (1) to reduce costs by adopting remote-sensing techniques to some of the tasks in which remote sensing can provide reliable and cost-efficient estimates, (2) to provide statistical estimates of the timber resources at local scales (sub-county) to support the local public forest administration, and (3) to provide cost-efficient and reliable estimates of biomass or carbon stocks of forest and nonforest land to meet the Kyoto Protocol requirements. This third challenge includes biomass or carbon estimates of thousands of square kilometers of mountain forest above the official tree line, which is currently not part of the NFI or any other national monitoring systems. This ecotone

[1] Professor, Norwegian University of Life Sciences, Department of Ecology and Natural Resource Management, P.O. Box 5003, NO-1432, Ås, Norway. E-mail: erik.naesset@umb.no.

[2] Associate Professor, Norwegian University of Life Sciences, Department of Ecology and Natural Resource Management, P.O. Box 5003, NO-1432, Ås, Norway. E-mail: terje.gobakken@umb.no.

[3] Physical Scientist, National Aeronautics and Space Administration/Goddard Space Flight Center, Biospheric Sciences Branch, Greenbelt, MD 20771. E-mail: ross.f.nelson@nasa.gov.

is of special interest in the climate change and C sequestration debate because it is likely that a temperature-induced productivity increase could more easily be detected in the mountain forest due to the steep temperature-productivity gradients.

For purposes in which statistical estimates of forest resources of a certain region are sought, airborne Light Detection and Ranging (LIDAR) may be used as a sampling device to collect representative data for a region. For nearly two decades, airborne LIDAR has been used as a tool for research in forest inventory, but a few years ago it was demonstrated for the first time that LIDAR also can be a powerful tool in regional inventory. In 2000, the National Aeronautics and Space Administration (NASA) collected LIDAR data along 56 parallel flight lines across the entire State of Delaware. A so-called profiling LIDAR was used. Unlike a scanning LIDAR that collects data along a corridor of a width of, say, 100 to 1,000 m, a profiling system is only capable of collecting a narrow line of data underneath the aircraft. Along the flight lines, 142 ground samples based on 40-m segments using line intersect sampling were used to provide ground estimates of volume and biomass (Nelson *et al.* 2003; Nelson, Short, and Valenti 2004). These estimates were regressed against canopy structural properties derived from the LIDAR data using parametrically and nonparametrically fit, explicitly linear, and ln-ln models. Stratified and nonstratified versions of these models were considered and used to predict volume and biomass along all 56 flight lines over the State using a stratified sampling scheme accounting for the proportion of different land categories of the State. Year 2000 profiling LIDAR-based results based on 5,159 km of flight data were compared, by county and State, with 1999 U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis (FIA) data based on 225 FIA plots. Results indicate that nonparametrically fit linear models provided results that most closely agreed with FIA timberland estimates. LIDAR-based estimates of merchantable volume and biomass were within 23 percent of FIA estimates at the county level and within 18 percent of FIA estimates at the State level. In all cases, the LIDAR-based estimates were more precise than the comparable FIA estimates were.

Scanning LIDAR is typically used to collect wall-to-wall data of ground topography and forest canopy elevations. Such systems are now being used operationally in forest stand inventory in Scandinavia (Næsset *et al.* 2004), and the procedures used to estimate biophysical stand properties are very similar to those used for data collected by profiling systems (i.e., field training data are related to LIDAR-derived metrics for a number of conventional ground plots distributed throughout the area in question, and estimated regression equations are used to predict biophysical properties of every stand in the target area based on the wall-to-wall coverage of LIDAR data [Næsset 2002, Næsset and Bjerknes 2001]). Extensive testing has indicated that the accuracy of LIDAR-based stand inventory is superior to that of conventional inventories based on fieldwork and/or stereo photogrammetry (Holmgren 2004; Maltamo *et al.* 2006; Næsset 2002, 2004a, 2004b).

Like profilers, however, scanning LIDARs, which capture data along wide corridors, can be used as sampling tools as well. The overall scientific aim of the ongoing study is to develop airborne scanning LIDAR as a strip sampling tool to inventory timber volume and biomass in large areas and compare the accuracy and costs of such an application with what can be obtained by a profiling system. This article presents the project plans and expected output of the work.

Methodology

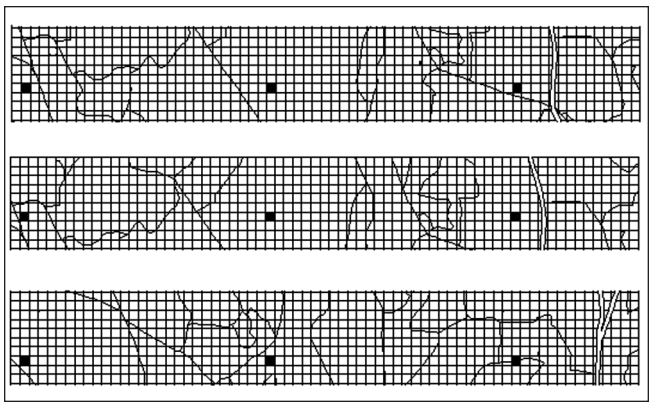
The profiling LIDAR to be tested in this project is the Portable Airborne Laser System (PALS) assembled at NASA/Goddard Space Flight Center (Nelson, Parker, and Hom 2003). PALS has previously been flown over Delaware, where the target variables were merchantable volume and aboveground biomass (Nelson *et al.* 2003). Later, it was flown over parts of U.S. States such as Texas and New Jersey, over Quebec in Canada, and over Japan. In the present project, PALS profiles are flown over NFI permanent forest inventory ground plots to develop parametric (regression) and nonparametric models that will tie the LIDAR measurements to ground-measured volume and biomass. These models will, in a subsequent step, be used to predict timber volume and biomass along all the flight lines. The forest area in question will be stratified according to forest types defined by criteria such as tree species and site productivity. The stratification will be based on existing Geographic

Information System databases of site productivity and land use status while Landsat multispectral image data and nonparametric estimation techniques using the NFI permanent sample plots as training data will guide the classification of the forest into different tree species categories. Thus, the entire inventory area will be assigned to one of several mutually exclusive classes and the classes will be used to correct the stratified estimates of timber volume and biomass derived along the PALS flight lines (ratio estimation).

A number of commercial scanning LIDAR systems are available in the market (Baltsavias 1999), but a disadvantage of these systems, as compared with simple profiling systems, is the high costs for data acquisition. Despite this disadvantage, in regional applications to provide volume and biomass estimates at subcounty, county, or national levels that conform to NFI requirements, even a scanning system can be used as a sampling tool by simply collecting laser scanner data along strips that may be separated by many kilometers.

The concept for a sampling-based method for timber and biomass inventory using airborne scanning LIDAR will, to a large extent, follow the procedure for wall-to-wall mapping of forest stands outlined previously (Næsset 2002, Næsset and Bjerknes 2001). NFI ground plots underneath the scanner flight lines will be used to derive stratified parametric and nonparametric models that tie ground measurements to the LIDAR data. Similar to the application based on profiling, these models will be used to predict volume and biomass along all flight lines. In contrast to profiling LIDAR, where the units used in prediction are short segments of LIDAR data with a length of, say, 40 m, the prediction units for scanner data are grid cells with a size equal to the ground plot size (see fig. 1). As with the profiling-LIDAR–based sampling approach, the scanner-based stratified volume and biomass estimates along the flight-line corridors will be adjusted according to the distribution of the entire area in question on different strata.

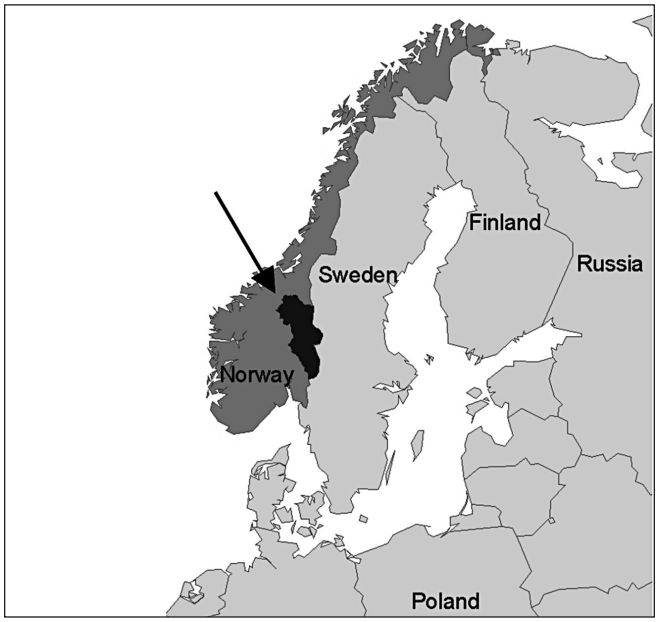
Figure 1.—*An illustration of a scanning Light Detection and Ranging (LIDAR) strip sampling scheme. The three strips of LIDAR data are divided into regular grid cells—the primary unit of the estimations. The black cells indicate systematically distributed field sample plots used as training data. Stratification occurs according to land use classes (irregular solid lines).*



Study Area

The target area of this investigation is Hedmark County, Norway. The total area is approximately 27,000 km^2 (fig. 2). The productive forest area is 13,420 km^2 (Tomter *et al.* 2001). The total standing volume is 134.7 million m^3. The forest is dominated by Scotch pine (48 percent) and Norway spruce (42 percent). Birch is the dominant deciduous tree species (8 percent). No precise estimates of the area covered by mountain forest exist, and, in the ongoing research, we will estimate the area, timber volume, and biomass or carbon stocks of this biome.

Figure 2.—*The study area: Hedmark County, Norway.*



Field Data

NFI circular ground plots will be used to calibrate equations that will relate ground-measured timber volume and biomass to the LIDAR data. Hedmark County is covered by 2,207 permanent NFI sample plots. The 250-m^2 plots, with a plot radius of 8.9 m, are distributed on a 3-x-3-km grid. Twenty percent of the permanent plots are remeasured every year, and the plots are selected within a 45-x-45-km block of plots according to a Latin square design. The NFI ground plots measured in 2005 through 2008 will be used in the calculations.

The NFI comprises all types of land below the coniferous forest limit (i.e., from approximately 120 m above sea level up to 900 m above sea level), but a comprehensive description is made only for forest land. Because it is also an objective to demonstrate that airborne LIDAR can provide data required for reporting on the United Nations Framework Convention on Climate Change and the Kyoto Protocol, approximately 100 additional ground plots will be measured on land other than forest land during the summers of 2006 and 2007.

About 100 additional field plots in the mountain forest above the official tree line, which is currently not a part of the NFI or any other national monitoring systems, will also be measured in 2006 and 2007.

On each plot, all trees with diameter at breast height of more than 4 cm are callipered, and the tree heights of an average of 10 sample trees per plots are measured. The coordinates (x, y) of each plot center are determined with an average accuracy of less than 0.5 m using differential Global Positioning System and Global Navigation Satellite System measurements, according to the procedures suggested by Næsset (2001). Timber volume will be estimated according to standard volume equations for individual trees (see Næsset 2002 for further details) and biomass will be estimated according to allometric equations (Marklund 1988). Estimates of standing carbon will be derived from the dry biomass estimates based on the 0.5 conversion factor from dry biomass to mass of carbon suggested by Gower *et al.* (1997) and Houghton *et al.* (2000).

LIDAR Data

During the summer of 2006, the PALS profiling LIDAR and an Optech ALTM 3100C scanning LIDAR were flown along parallel flight lines over the ground plots. Approximately 9,000 km of flight lines were flown with the PALS LIDAR profiler, whereas scanning LIDAR data were acquired along 4,500 km of flight lines. The spacing between flight lines was 3 and 6 km for the profiler and scanner, respectively, which means that the profiler was flown over all ground plots while the scanner was flown over 50 percent of the plots. The average pulse density for the scanner was 2.5 to 3 m^{-2}. The first and last echoes were recorded by the profiler, while the scanner recorded up to four echoes per laser pulse. The scanner data cover as much as 8 percent of the entire land surface in the county and thus this data set represents a very large sample of the forest in the area.

Discussion

The estimates of timber volume for Hedmark County derived by scanning and profiling LIDAR will be compared with the official statistics based on the NFI for the 1995-to-1999 period. Because the NFI only provides county-level estimates, only totals for the entire county will be compared. The standard error of the official estimate of total standing volume is estimated to 3.0 percent (Tomter *et al.* 2001). Despite these limitations, we hope to be able to compare subcounty estimates derived from the scanning and profiling LIDAR data acquisitions with corresponding estimates from a complete wall-to-wall inventory by scanning LIDAR planned to take place in a 2,000 km^2 area in the county. This inventory is part of an operational project to provide improved digital terrain models and data for forest management planning undertaken by the local mapping authorities and the local forest owners, respectively.

Many sources of errors will be associated with the estimates of volume and biomass at county and subcounty levels derived from the scanning as well as the profiling systems. Some of the major sources of variance are sampling errors, regression errors inherent in the LIDAR data/field data relationships, and errors

in allometric and other equations. These sources probably have different effects on the estimates derived by profiling and scanning LIDAR.

The efforts to improve laser-based inventories will—if the results confirm the hypothesis—reduce the costs and improve the accuracy and applicability of airborne LIDAR in future NFI programs.

Literature Cited

Baltsavias, E.P. 1999. Airborne laser scanning: existing systems and firms and other resources. Journal of Photogrammetry and Remote Sensing. 54: 164-198.

Gower, S.T.; Vogel, J.G.; Norman, J.M.; Kucharik, C.J.; Steele, S.J.; Stow, T.K. 1997. Carbon distribution and aboveground net primary production in aspen, jack pine, and black spruce stands in Saskatchewan and Manitoba, Canada. Journal of Geophysical Research. 102(D24): 29029-29041.

Holmgren, J. 2004. Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. Scandinavian Journal of Forest Research. 19: 543-553.

Houghton, R.A.; Skole, D.L.; Nobre, C.A.; Hackler, J.L.; Lawrence, K.T.; Chomentowski, W.M. 2000. Annual fluxes of carbon from deforestation and regrowth in the Brazilian Amazon. Nature. 403(24): 301-304.

Maltamo, M.; Eerikäinen, K.; Packalen, P.; Hyyppä, J. 2006. Estimation of stem volume using laser scanning-based canopy height metrics. Forestry. 79: 217-229.

Marklund, L.G. 1988. Biomass functions for pine, spruce and birch in Sweden. Report 45. Umeå, Sweden: Swedish University of Agricultural Sciences, Department of Forest Survey. 73 p. In Swedish with English summary.

Næsset, E. 2001. Effects of differential single- and dual-frequency GPS and GLONASS observations on point accuracy under forest canopies. Photogrammetric Engineering & Remote Sensing. 67: 1021-1026.

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

Næsset, E. 2004a. Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. Scandinavian Journal of Forest Research. 19: 164-179.

Næsset, E. 2004b. Accuracy of forest inventory using airborne laser-scanning: evaluating the first Nordic full-scale operational project. Scandinavian Journal of Forest Research. 19: 554-557.

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

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

Nelson, R.; Parker, G.; Hom, M. 2003. A portable airborne laser system for forest inventory. Photogrammetric Engineering & Remote Sensing. 69: 267-273.

Nelson, R.; Short, A.; Valenti, M. 2004. Measuring biomass and carbon in Delaware using an airborne profiling LIDAR. Scandinavian Journal of Forest Research. 19: 500-511. [Erratum. 2005. 3: 283-284].

Nelson, R.; Valenti, M.A.; Short, A.; Keller, C. 2003. A multiple resource inventory of Delaware using airborne laser data. BioScience. 53: 981-992.

Tomter, S.M.; Eriksen, R.; Aalde, H. 2001. Statistikk over skogforholdene og –ressursene i Hedmark. Landsskogtakseringen 1995–99 (Forest statistics of Hedmark, 1995–99). Ås, Norway: Norwegian Institute of Land Inventory. 56 p. In Norwegian with English abstract.

Forest-Cover–Type Separation Using RADARSAT-1 Synthetic Aperture Radar Imagery

Mark D. Nelson[1], Kathleen T. Ward[2], and Marvin E. Bauer[3]

Abstract.—RADARSAT-1 synthetic aperture radar data, speckle reduction, and texture measures provided for separation among forest types within the Twin Cities metropolitan area, MN, USA. The highest transformed divergence values for 16-bit data resulted from speckle filtering while the highest values for 8-bit data resulted from the orthorectified image, before and after performing a histogram stretch. First-order texture derivatives of 8-bit data provided only modest separability, while second-order texture derivatives of 8-bit data provided little, if any, separability. RADARSAT-1 imagery may provide for image separation among forest types provided that preexisting forest/nonforest land cover classifications are incorporated.

Introduction

The classification of forest land cover types from remotely sensed imagery is important for enhancing forest assessments. For example, a forest-cover–type map of the United States derived from Advanced Very High Resolution Radiometer (AVHRR) satellite imagery was used to supplement the Forest and Rangeland Renewable Resources Planning Act of 1974 (which was amended in 1992) Forest Resources Assessment (Powell *et al.* 1993, Zhu and Evans 1994). Päivinen *et al.* (2001) describe a more recent effort to combine AVHRR data with European forest statistics. The utility of these forest-cover–type maps depends on their classification accuracy and their correspondence with inventory-based forest assessments. Thus, developing approaches for improving forest-cover–type classifications provides significant value to forest assessment.

Optical sensors provide the source imagery for most forest-cover–type mapping efforts, but imagery collected by the sensors can be adversely affected by sun angle and atmospheric properties (e.g., clouds, haze, aerosols). Active radar signals are unaffected by these factors, but are sensitive to the moisture content and structural properties of vegetation. Satellite-borne synthetic aperture radar (SAR) data has potential for separating classes of land cover, especially when integrated with optical sensor data (Huang *et al.* 2007).

Imaging radar systems typically are categorized by microwave bands, with the following frequencies and wavelengths representing centers for each of four common bands: P-band (440 MHz, 65 cm λ), L-band (1.25 GHz, 24 cm λ), C-band (5.3 GHz, 5.6 cm λ), and X-band (10 GHz, 3.0 cm λ) (Kasichke *et al.* 1997). Each microwave band is differentially backscattered by objects approximately equal in size to the band's wavelength (Ranson and Williams 1992). Microwave scattering and attenuation in C-band SAR results from interactions with tree canopy leaves, needles, and small secondary branches, but C-band backscatter from tree trunks is small due to minimal canopy penetration (Kasichke *et al.* 1997). C-band SAR data does poorly at discriminating conifer from deciduous forest but shows good capabilities for differentiating some specific northern/temperate forest types (Leckie and Ranson 1998).

SAR data consists of backscatter or intensity values (strength of signal return) and texture (variability of backscatter within adjacent groups of pixels). SAR image texture consists of two components: speckle fluctuations and backscatter fluctuations. Speckle results in a grainy, salt-and-pepper appearance in SAR

[1] Research Forester, U.S. Department of Agriculture (USDA) Forest Service, Northern Research Station, St. Paul, MN.
[2] Forester, USDA Forest Service, Northern Research Station, St. Paul, MN.
[3] Professor, College of Food, Agricultural and Natural Resource Sciences, University of Minnesota, St. Paul, MN.

data. Speckle reduction is recommended prior to performing radar image processing (Collins *et al.* 2000). When analyzing SAR imagery, texture elements can be as important as backscatter data. SAR texture information is derived from first-order texture measures, such as moving window analysis and second-order statistics calculated from the Gray Level Co-occurrence Matrix (GLCM) (Haralick *et al.* 1973).

In summary, C-band SAR data achieves only moderate success at discriminating forest types and is even less useful for discriminating forest from nonforest. C-band SAR, however, may aid in discriminating among tree species when analyses are constrained to previously identified forested pixels and when speckle filtering and texture information are incorporated. In this study we assess the utility of RADARSAT-1 C-band SAR imagery for separating forest-cover types in the Twin Cities metropolitan area (TCMA), MN, USA.

Data and Methods

Study Area

The 770,000-ha study area includes seven counties (Anoka, Carver, Dakota, Hennepin, Ramsey, Scott, and Washington) within the TCMA of Minneapolis-St. Paul, MN, USA. Forest land, most of which is deciduous, constitutes about 90,000 ha (12 percent) of the TCMA (Yuan 2004). We aggregated Yuan's (2004) TCMA map of land cover classes into forest and nonforest classes. The forest/nonforest data set was rescaled to 15-m pixel resolution to accommodate subsequent combinations with 15-m radar data. Subsequent analyses of radar data were constrained to forested pixels of the TCMA.

RADARSAT-1

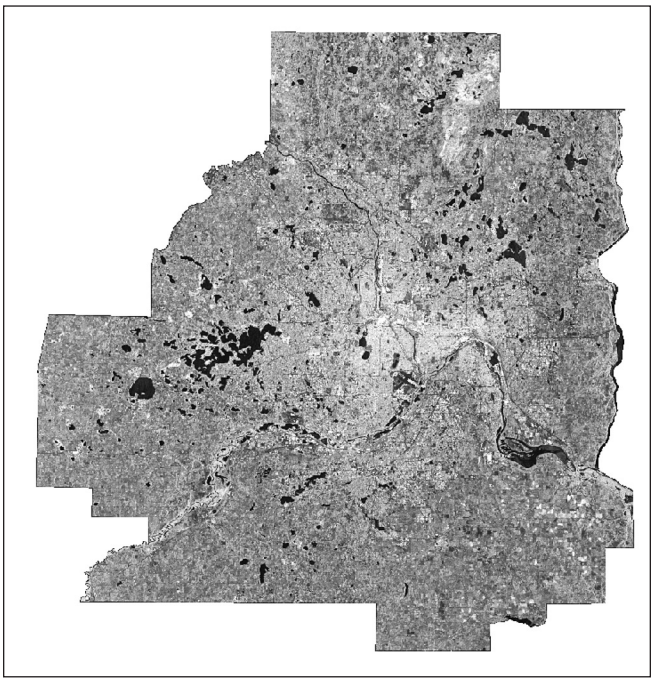
One 16-bit C-band RADARSAT-1 image, which was acquired June 28, 2001, was used in this study. We performed geometric correction using a 10-m digital elevation model, 60 ground control points, a second-order polynomial, and cubic convolution resampling, which resulted in 15-m pixel spatial resolution, orthorectified (ortho) image. Figure 1 displays backscatter values characteristic of the ortho image: low backscatter (black) portrays open water, high backscatter (white) portrays corner reflectors such as vertical buildings adjacent to horizontal impervious surfaces (white), and intermediate backscatter (gray shades) portrays forest land, cropland, and other land covers.

The resampled image was clipped to the geographic extent of the TCMA study area and was masked to exclude pixels identified as nonforest in the land cover classification. Speckle filtering was performed on the 16-bit ortho image, using a 3-×-3 pixel window Frost speckle filter (Frost *et al.* 1982). A variety of first- and second-order texture measures were computed for 3-×-3 and 5-×-5 windows but only the most promising are discussed here: first-order "variance" (actually, standard deviation) and second-order GLCM dissimilarity. Software limitations prevented the calculation of GLCM second-order texture measures from the 16-bit ortho image, so the image was converted to an unsigned 8-bit image and was histogram-stretched prior to computing first- and second-order texture measures.

Reference Data

Field data used for separability analyses are from the Minnesota Land Cover Classification System (MLCCS) (MDNR 2004). The MLCCS reference data include 19 forest types, in-

Figure 1.—*RADARSAT-1 backscatter image of the Twin Cities metropolitan area, MN, USA, June 28, 2001.*



cluding aspen, black ash, box elder, lowland deciduous, maple basswood, red pine, silver maple, tamarack, upland conifer, upland deciduous, and 9 partially overlapping oak subtypes—oak woodland/brushland, oak with impervious cover, white oak, red/white/bur oak, red oak, oak mesic forest, oak forest, oak/mixed deciduous, and bur oak. The oak subtypes were similar to each other so all 9 oak subtypes were aggregated into a single oak class, resulting in 11 MLCCS forest types. Black ash was omitted from first-order texture analyses because covariance matrices for this type were not invertible. Ten forest types were included in first-order analyses and 11 forest types in second-order analyses (table 1).

Forest-Type Class Separability Analyses

We calculated forest-type class separability using average and maximum transformed divergence (TD) values (Jensen 1996). According to Jensen (1996), TD values below 1,700 indicate poor separation, values above 1,900 indicate good separation, and values of 2,000 indicate excellent separation. We assumed that TD values between 1,700 and 1,900 indicate moderate separability.

Results

MLCCS forest types were inseparable when using the 16-bit ortho image, (average TD of 68, maximum TD of 254; table 1). Speckle reduction from the 3×3 Frost filter resulted in average TD of 1,496 and maximum TD of 2,000 (table 1). Average and maximum TD values were 1,716 and 2,000, respectively, before histogram stretching and 1,606 and 2,000, respectively, after histogram stretching of the 8-bit ortho image (table 1). Average and maximum separability of forest types using a first-order texture measure (5-×-5 standard deviation texture after histogram stretching) were 502 and 1,995, respectively (table 1). Forest types were inseparable using a dissimilarity second-order texture measure with a 3-×-3 window and a 3 standard deviation stretch (average TD of 57, maximum TD of 334; table 1).

Discussion

Separability among 1 forest classes increased in TD values after speckle filtering, and was greatest with a 3-×-3 Frost filter. This observation conforms to the results of Kuplich *et al.* (2000), who reported that image filtering produced improved classifications of urban, pasture, and forest classes in Brazil. In addition, rescaling from 16-bit to 8-bit data resulted in markedly improved separability among 11 forest types. First-order texture of 8-bit data appeared to provide separation of forest-type classes, but the separability was no better than using 8-bit data alone. Second-order texture measures appeared to provide little or no separability among pairs of forest types. Although the approach required orthorectification, several image-processing steps, and the availability of a preexisting forest/nonforest land cover classification, RADARSAT-1 imagery can provide for separability among many forest types. This capability may be applied for forest-type mapping, especially where improved discrimination among broad-leaved or needle-leaved types is required or where imagery from optical sensors is degraded by persistent cloud cover.

Table 1.—*Transformed divergence values for RADARSAT-1 imagery and selected derivatives, Twin Cities metropolitan area, MN, USA.*

Type	Band	Texture	Ave.	Max.
16-bit	ortho	none	68	254
16-bit	ortho, 3×3 Frost speckle filter	none	1,496	2,000
8-bit	ortho	none	1,716	2,000
8-bit	ortho, 4 standard deviation stretch	5×5 standard deviation	502	1,995
8-bit	ortho, 3 standard deviation stretch	3×3 dissimilarity	57	334

Ave. = average. Max. = maximum.

Literature Cited

Collins, M.J.; Wiebe, J.; Clausi, D.A. 2000. The effect of speckle filtering on scale-dependent texture estimation of a forested scene. IEEE Transactions on Geoscience and Remote Sensing. 38: 1160-1170.

Frost, V.S.; Stiles, J.A.; Shanmugan, K.S.; Holtzman, J.C. 1982. A model for radar images and its application to adaptive digital filtering of multiplicative noise. IEEE Transactions on Pattern Analysis and Machine Intelligence. 4: 157-166.

Haralick, R.M.; Shanmugam, K.; Dinstein, I. 1973. Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics. SMC-3: 610-621.

Huang, H.; Legarsky, J.; Othman, M. 2007. Land-cover classification using RADARSAT and Landsat imagery for St. Louis, Missouri. Photogrammetric Engineering & Remote Sensing. 73: 37-43.

Jensen, J.R. 1996. Introductory digital image processing: a remote sensing perspective. 2nd ed. Upper Saddle River, NJ: Prentice Hall. 316 p.

Kasichke, E.S.; Melack, J.M.; Dobson, M.C. 1997. The use of imaging radars for ecological applications—a review. Remote Sensing of Environment. 59: 141-156.

Kuplich, T.M.; Freitas, C.C.; Soares, J.V. 2000. The study of ERS-1 SAR and Landsat TM synergism for land use classification. International Journal of Remote Sensing. 21: 2101-2111.

Leckie, D.G.; Ranson, K.J. 1998. Forestry applications using imaging radar. In: Henderson, F.M.; Lewis, A.J., eds. Principles and applications of imaging radar (manual of remote sensing). New York: John Wiley: 436-509.

Minnesota Department of Natural Resources (MDNR). 2004. Minnesota Land Cover Classification System user manual. Version 5.4. Minnesota Department of Natural Resources, Central Region. 273 p. http://files.dnr.state.mn.us/assistance/nrplanning/community/mlccs/mlccs_manual_v5_4.pdf. (9 November 2006).

Päivinen, R.; Lehikoinen, M.; Schuck, A.; Häme, T.; Väätäinen, S.; Kennedy, P.; Folving, S. 2001. Combining Earth observation data and forest statistics. Res. Rep. 14. Joensuu, Finland: European Forest Institute, Joint Research Center—European Commission. 101 p.

Powell, D.S.; Faulkner, J.L.; Darr, D.R.; Zhu, Z.; MacCleery, D.W. 1993. Forest resources of the United States, 1992. Gen. Tech. Rep. RM-234. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station. 132 p.

Ranson, K.J.; Williams, D.L. 1992. Remotely sensed technology for forest ecosystem analysis. In: Shugart, H.H.; Leemans, R.; Bonan, G.B., eds. A systems analysis of the global boreal forest. New York: Cambridge University Press: 267-290.

Yuan, F. 2004. Remote sensing and GIS-based regional land-cover mapping and change analysis in the Twin Cities metropolitan area. St. Paul, MN: University of Minnesota. 131 p. Ph.D. dissertation.

Zhu, Z.; Evans, D.L. 1994. U.S. forest types and predicted percent forest cover from AVHRR data. Photogrammetric Engineering & Remote Sensing. 60: 525-531.

Rapid Mapping of Hurricane Damage to Forests

Eric M. Nielsen[1]

Abstract.—The prospects for producing rapid, accurate delineations of the spatial extent of forest wind damage were evaluated using Hurricane Katrina as a test case. A damage map covering the full spatial extent of Katrina's impact was produced from Moderate Resolution Imaging Spectroradiometer (MODIS) satellite imagery using higher resolution training data. Forest damage detected in MODIS imagery corresponded well with damaged areas mapped in aerial surveys and matched the patterns expected in relation to maximum sustained winds, topography, and forest type. The results were encouraging for the prospects of rapid assessment using MODIS and other environmental data.

Introduction

Hurricanes are a major natural disturbance process impacting forests in the southeastern and northeastern United States (Boose *et al.* 1994, Connor 1998). Although no standardized system for the quantification of wind damage to forests has been developed (Everham and Brokaw 1996), the 2005 hurricane season created awareness of forest managers' and policymakers' needs for extremely rapid assessments of resource damage caused by sudden, large-scale events. The U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis (FIA) program and Remote Sensing Applications Center are therefore preparing methodologies for rapid estimations of land areas and forest resources affected by such events. Data from satellite sensors with high repeat frequency and large spatial footprint (e.g., the Moderate Resolution Imaging Spectroradiometer, or MODIS) appear to hold the most promise for allowing rapid delineation of impacted areas, but the impact of the coarse spatial resolution of these sensors on hurricane damage mapping accuracy has not been evaluated. This work is an initial investigation into the feasibility of producing rapid, accurate maps of hurricane damage using MODIS imagery.

Challenges to any hurricane damage assessment include the continuing accumulation of tree mortality for several years after an event, which has been found both in pines (Armentano *et al.* 1995, Platt *et al.* 2000) and hardwoods (Walker 1995). On the other hand, much immediately apparent damage to hardwoods consists of simple defoliation (Kelly 1993). Recovery of many hardwoods following defoliation or even substantial structural damage can be rapid, with releafing and sprouting often occurring within 3 to 4 weeks (Armentano *et al.* 1995, Loope *et al.* 1994). Understory forest vegetation, particularly vines, can also respond rapidly to increased light levels (Loope *et al.* 1994), confounding attempts to associate forest damage with drops in remotely sensed measures of leaf area. For example, Ramsey *et al.* (1998) observed an abnormal green-up in hardwood-dominated forests in autumn 1992 following the passage of Hurricane Andrew through coastal Louisiana. Defoliated sites experienced a flush of new leaf and sprout growth, while understory plants responded with rapid growth at more severely damaged sites. Walker (1995) stated that accurate mortality estimates require long-term monitoring; estimates made immediately after an event—before recovery from defoliation—may be too high, but subsequent estimates may be lower than the mortality realized after the passage of several years.

Remote Sensing of Hurricane Forest Damage

Hurricane damage to forests has been mapped via satellite-based remote sensing in a number of studies. Landsat Thematic Mapper (TM) data at 30-m resolution were used by Kovacs (2001) to map disturbance in mangrove forests and by Clark *et al.* (2006) to map Hurricane Katrina's impacts on pines and hardwoods in the DeSoto National Forest in southern Mississippi. Ramsey *et al.* (1998) used 1-km Advanced Very High

[1] Remote Sensing Scientist, U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center, 2222 West 2300 South, Salt Lake City, UT 84119. Email: emnielsen@fs.fed.us.

Resolution Radiometer (AVHRR) imagery to map impacts to bottomland hardwood and bald-cypress/tupelo (*Taxodium distichum*/*Nyssa aquatica*) stands in coastal Louisiana. AVHRR imagery was also used by Ayala-Silva and Twumasi (2004) to investigate patterns of damage with relation to distance from the path of Hurricane Georges in Puerto Rico. All of these studies found change in the Normalized Difference Vegetation Index (NDVI) to be a useful indicator of forest damage, but none of the coarser resolution studies had the opportunity to systematically validate their estimates against reference data.

Environmental Patterns of Hurricane Forest Damage

Many studies have assessed hurricane damage in relation to meteorological and environmental variables (referred to henceforth as "auxiliary data"). To the extent that such relationships are consistent across forest types and study areas, they can be helpful in verifying damage estimates derived from remotely sensed data and for refining or applying spatial restrictions to such estimates.

Wind speed is the variable most directly related to forest damage. Doyle *et al.* (1995) found that the severity and spatial extent of hurricane damage declined at greater distances from the storm track, although damaging winds may extend farther from the center on the eastern side of the track (Kelly 1993). Ramsey *et al.* (2001) compared forest damage in Louisiana with modeled wind speeds and found evidence that damage severity depends on the duration for which forests are subjected to wind speeds exceeding a critical threshold. Jacobs (2006) used relationships found between maximum sustained wind speed and forest damage to create rapid assessment maps of forest damage resulting from Hurricanes Katrina and Rita.

The topographic slope and aspect associated with a forest stand determine its degree of exposure to peak winds (e.g., McNab *et al.* 2004). Topography also influences the soil conditions available for tree rooting. Damage has been found to be greater in valleys than on ridges and slopes, probably because of poor soil drainage, shallow rooting depths, and lack of root anchorage (Basnet *et al.* 1992). Boose *et al.* (1984) summed up the influence of both wind and topography by stating that regional scale damage corresponded with the distribution of peak wind speeds, while topographic exposure to predominant wind directions explained the distribution of damage at the landscape scale.

Leininger *et al.* (1997) found that certain highly susceptible tree species were generally associated with severely damaged bottomland hardwood sites in Louisiana. Gresham *et al.* (1991) quantified damage at a species level on the South Carolina coastal plain, finding that susceptible species generally grew in locations with high water tables and had physiognomic characteristics, such as shallow root systems and unbuttressed boles, that predisposed them to damage. Because such characteristics are shared by species among both pines and hardwoods, a map of forest type based on a fine-level of classification would be needed to exploit species-level susceptibility information for predicting storm damage. In general, forest type response patterns are difficult to elucidate in isolation from other auxiliary variables because of co-correlated geographic variation in wind speed (e.g., Kelly 1993) and the association of particular forest types with certain topographic environments that also affect their susceptibility.

The approach here is broken into two parts. First, an assessment is made of whether the MODIS sensor—given its spatial and spectral resolution—is able to accurately detect areas of hurricane damage, given no constraints on how rapidly after an event an assessment is required. Second, a test of the prospects for producing an accurate delineation within days of an event is made using MODIS and auxiliary data such as maximum sustained wind speed, topographic variables, and forest type. This question addresses the issues of image availability and the temporal evolution of forest response after an event.

It is important to note that we used higher resolution reference data here that had been collected specifically in response to Hurricane Katrina. Operational rapid delineation of damage over different regions or forest types in the absence of event-specific training data will depend on the consistency of spectral reflectance responses to forest damage or the repeatability of damage patterns in relation to auxiliary data such as wind speed, topography, and soil types. These are major concerns that require addressing in future work.

Methods

Predictor Data and Study Area Definition

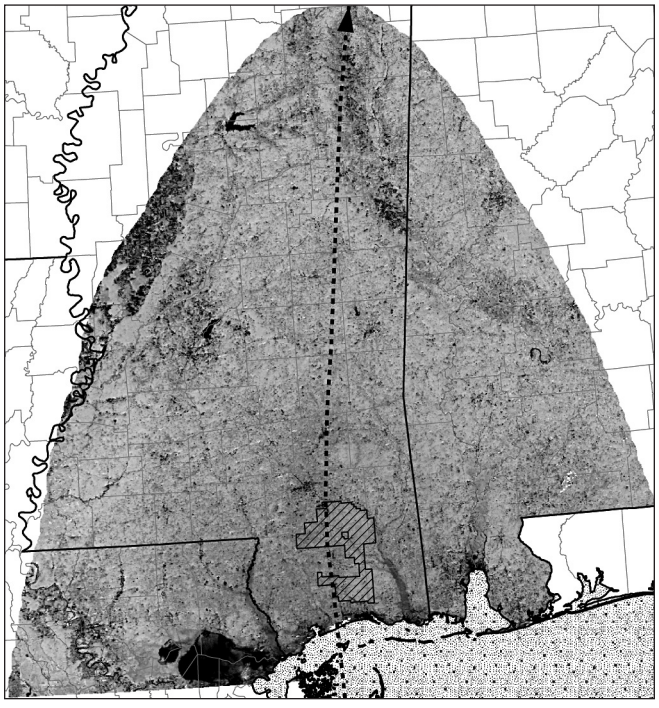
MODIS satellite data are available as tiles covering a fixed spatial area. Data from the tile overlapping the primary affected region in Louisiana, Mississippi, and Alabama were downloaded from the U.S. Geological Survey (USGS) Land Processes Distributed Active Archive Center (http://edcdaac.usgs.gov); the 250-m daily reflectance product (MOD09GQK) was chosen. All images with fairly low cloud proportions from the months of August through early November 2005 were obtained. A prestorm image was formed by mosaicking images collected on August 16 and 21, 2005, just before the hurricane's landfall on August 29, and the final image was collected on November 2, 2005. A forest analysis mask was created using the 30-m National Land Cover Dataset (NLCD) 2001 land cover classification. Forested classes (including wooded wetlands) were extracted from the NLCD and used to calculate forested proportion at the MODIS pixel resolution. Aggregated pixels containing less than 75 percent forest were eliminated from the analysis to diminish the negative impacts of nonforest land cover classes and misregistration error. The proportions of each MODIS pixel represented by each NLCD forest type—deciduous upland, evergreen upland, and wooded wetland—were calculated. A water mask was similarly created to restrict analysis near large bodies of water, where minor spatial registration errors could have major impacts on calculated spectral differences.

Maximum sustained surface wind speed data, available at 3-hour intervals, were downloaded from the Atlantic Oceanographic and Meteorological Laboratory (AOML Hurricane Research Division 2006). For each analysis pixel, the maximum value was extracted from the stack of 3-hour wind speed images. The values were classed into equal wind speed interval categories, and the envelope for each wind speed category was digitized on-screen. The study area (fig. 1) was then defined to incorporate all predominantly forested areas within the region estimated to have experienced tropical storm force winds. Finally, digital elevation model data at 30-m resolution were downloaded from the USGS National Elevation Dataset website (http://ned.usgs.gov/) and slope and aspect were computed using standard algorithms.

MODIS Data Preprocessing

An automated masking procedure was applied to the MODIS images to remove cloud and shadow areas. The procedure used variation from the time-series median reflectance values for each pixel to detect clouds and shadows, so it was necessary to first standardize the data by applying a relative radiometric normalization to correct for view angle variability and residual atmospheric effects. The normalization was performed by applying Reduced Major Axis (RMA) regression (see Cohen *et al.* 2003) to the predominantly forested pixels in each pair of images, using a preliminary threshold-based cloud and shadow mask. Clouds and shadows were then removed from the nor-

Figure 1.—*The study area was defined by the intersection of the region exposed to tropical storm-force winds or greater (the northward-trending oval) with Moderate Resolution Imaging Spectroradiometer (MODIS) tile h10v05 (which sets the southern boundary of the study area). The backdrop is a MODIS Normalized Difference Vegetation Index image acquired on the 10th day after Hurricane Katrina's landfall, with flooding and forest damage evident along the major floodplains. The study area overlaps the states of Louisiana, Mississippi, and Alabama, with Lake Ponchartrain and the flooded New Orleans visible at the southern boundary. The bold dashed line illustrates the path of Katrina's eye. The crosshatched polygon represents the DeSoto District of the DeSoto National Forest, the area from which the damage estimates used as training data were derived.*



malized imagery and replaced with data interpolated from the best previous and subsequent images. The procedure resulted in a complete normalized time-series of MODIS imagery with no missing values.

Reference Data

Clark *et al.* (2006) produced a categorical damage probability map of the DeSoto District in the DeSoto National Forest (see fig. 1), based on change in Landsat TM NDVI between October 15, 2004 and November 3, 2005, and calibrated to damage photointerpreted from aerial photography collected soon after the hurricane's passage by the Forest Service Forest Health Technology Enterprise Team. We determined the mean damage probability represented by each of the mapped damage classes using the statistics provided and then spatially aggregated the probabilities to the MODIS pixel resolution. The aggregated damage probabilities were treated as estimates of the MODIS pixel fraction that had sustained damage and used as reference data here.

Another source of reference data was a Forest Service Forest Health Protection (FHP) aerial sketchmap product that illustrated hurricane damage over several parishes in southeastern Louisiana (Steiner *et al.* 2006). A graphical representation of this map was converted to digital form and georeferenced. This product was the only other spatially explicit reference data discovered that was feasible for use. Although other aerial photography may be available from various sources, its interpretation at the coarse MODIS pixel scale would be extremely time-consuming.

Modeling Approach

Damage Class Discriminant Analysis. Because of the ubiquity of NDVI differencing as a change detection technique, an initial assessment of the capability of the MODIS sensor to detect hurricane damage was made by modeling the TM damage estimates using the MODIS time-series NDVI values. For this purpose, the aggregated forest damage estimates derived from Clark *et al.* (2006) were converted into three discrete classes via an agglomerative clustering technique using Ward's distance criterion. The MODIS NDVI values were converted to a series of difference values by subtracting each of the post-storm time step NDVI values from the initial, prestorm NDVI value for that pixel. Use of difference data only reduced the likelihood of modeling potentially erroneous information in the training data that could have been caused by other land cover change in the year between the prestorm and poststorm TM images. Discriminant analyses were then performed to separate the three TM damage classes on the basis of the MODIS NDVI difference values. The accuracy of the functions was assessed via cross-validation, holding out one sample at a time, and compared with the accuracy expected due to chance alone using a kappa statistic.

Continuous Damage Estimation. A predictor of continuous forest damage based on MODIS data was trained using the aggregated TM damage data. The predictor was constructed using forward stepwise multiple regression on the NDVI difference time series as well as time-series differences based on the raw red and near-infrared MODIS reflectance data. The multiple regression fits were then adjusted to match the mean and variance of the damage training data using RMA regression. Because the application of this predictor to the entire area impacted by the hurricane involved an extrapolation from the small training area, some care was taken to immunize it against potential spatial heterogeneity in the relationship between spectral reflectance and forest damage. In particular, an analysis of the MODIS time-series data indicated that deciduous leaf senescence was well under way in the northern portion of the study area by the date on which the last MODIS image was collected. Therefore, the continuous damage predictor was trained using only the images collected before the date at which senescence was estimated to have begun at any location in the study area. The predictor was then applied to all pixels estimated to have experienced sustained tropical storm force winds, producing a spatially complete estimate of forest damage due to the hurricane.

Damage Estimation Assessment. The resulting damage map was evaluated by comparison with the independent Steiner *et al.* (2006) sketchmap product. Because of spatial uncertainties in the sketchmap, this comparison was done on a simple visual basis. The reasonableness of the regionwide forest damage map was also assessed through statistical assessment of the degree to which expected relationships between the auxiliary data (in

particular, maximum sustained winds, dominant forest type, and slope) and the damage estimate were realized. Because the damage estimate was based solely on MODIS spectral data, assessment of its relationship to the independent auxiliary data constitutes an indirect test of its accuracy. This evaluation was performed by examining the correlation coefficients between the damage estimate and each of the auxiliary variables and by breaking the continuous auxiliary variables into logical discrete classes and determining the significance of differences between class sample means via analysis of variance. Maximum sustained wind speeds experienced by forests across the study area were broken into four classes, corresponding to weak tropical storm force (< 60 mph), strong tropical storm force (60 to 73 mph), Category 1 hurricane force (74 to 95 mph), and Category 2 hurricane force (96 to110 mph), while the aggregated NLCD forest cover percentages were recoded into discrete classes by coding any pixel with greater than 50-percent cover in either deciduous upland, evergreen upland, or wooded wetland forest into the corresponding class.

Rapid MODIS Mapping Feasibility. The prospects for creating rapid, accurate maps of forest damage within days of an event were tested by repeating the previous regression estimation approach using only the spectral differences between the prestorm and immediate poststorm (3 days after landfall) MODIS images rather than the time series used previously. The estimates produced were evaluated using the same approaches described previously.

Rapid Mapping Feasibility Using Auxiliary Data. Finally, the feasibility of producing a rapid forest damage map based solely on auxiliary data (a fairly likely contingency, given the typically cloudy summertime conditions in the southeastern United States) was assessed by using the time-series damage estimate described previously to train a similar multiple regression estimator based on wind speed, NLCD forest type, and slope. The indirect approach of training the auxiliary damage estimation using the MODIS-modeled data was necessary because the small spatial extent of the TM-derived damage map did not incorporate a significant range of variability in the auxiliary data, particularly for wind speed. The auxiliary damage estimate was assessed visually.

Results

Damage Class Discriminant Analysis. The clustering technique produced three forest damage classes, referred to here as little to no damage (0 to 11.35 percent of MODIS pixel damaged), moderate damage (11.35 to 25.8 percent), and severe damage (>= 25.8 percent). The best discrimination between these classes was achieved with a quadratic discriminant function based on both the immediate poststorm and the early November changes in NDVI relative to the prestorm value. The discriminant function produced a cross-validation accuracy of 77.8 percent (Kappa = 0.552) and achieved a fairly even balance between omission and commission errors (table 1). Class discrimination based on immediate poststorm data only was only slightly poorer, with an accuracy of 76.3 percent (Kappa = 0.507). The best linear discriminant function using all time-series data had an accuracy of 73.0 percent (Kappa = 0.488). On the basis of these results, it appears that MODIS data have the inherent spatial and spectral properties needed to produce reasonably accurate maps of forest hurricane damage.

Continuous Damage Estimation. The forward stepwise regression based on all time-series MODIS reflectance and NDVI data indicated a total of 13 predictor variables with significance $p < 0.05$ and a resulting adjusted R^2 of 0.666. No relationships were found between the residuals and the model fits, and residuals were essentially normally distributed. The application of the derived regression equation across the study area yielded

Table 1.—*Cross-validation accuracies of the quadratic discriminant function produced to separate Thematic Mapper-derived forest damage classes on the basis of Moderate Resolution Imaging Spectroradiometer Normalized Difference Vegetation Index change data.*

	True damage class		
Classed as	**Little to none**	**Moderate**	**Severe**
	0–11.35%	11.35%–25.8%	>=25.8%
Little to none	1,205	112	14
Moderate	241	178	44
Severe	11	41	237
Proportion correct	0.827	0.538	0.803

Total proportion correct = 0.778 (95% confidence interval = 0.759 - 0.795); Kappa = 0.552.

the forest damage estimate map shown in figure 2. The damage represented is concentrated in the southern portion of the study area, where wind speeds were highest, and in floodplain areas, in keeping with the general findings of Gresham *et al.* (1991) and Basnet *et al.* (1992).

Damage Estimation Assessment. The visual comparison of the continuous damage map with the Louisiana aerial sketchmap data (Steiner *et al.* 2006) showed a good degree of correspondence (fig. 3). Although the spatial alignment of the data is not perfect, probably due to uncertainties in the sketchmapping process, the general patterns are similar, with extensive damage in the floodplains and scattered damage elsewhere. The expected relationships between the predicted forest damage and the auxiliary variables were found, with a positive correlation (r = 0.545) between maximum sustained wind speed and predicted damage, a negative correlation (r = -0.379) between slope and predicted damage (i.e., greater damage predicted in floodplains), and a positive correlation (r = 0.486) between the percentage of wooded wetland and predicted damage (again, more damage in floodplains). Normal curves fitted to the predicted damage histograms for each of the four wind speed

Figure 2.—*The percentage of forest damage predicted across the southern portion of the study area that was more dramatically impacted. The lightest grays visible indicate an estimated damage fraction of 10 to 15 percent while the darkest colors indicate an estimated damage fraction of 50 percent or more. Damage is concentrated near the coast, where wind speeds were highest, and in floodplains, in agreement with the assessment by Windham (2005).*

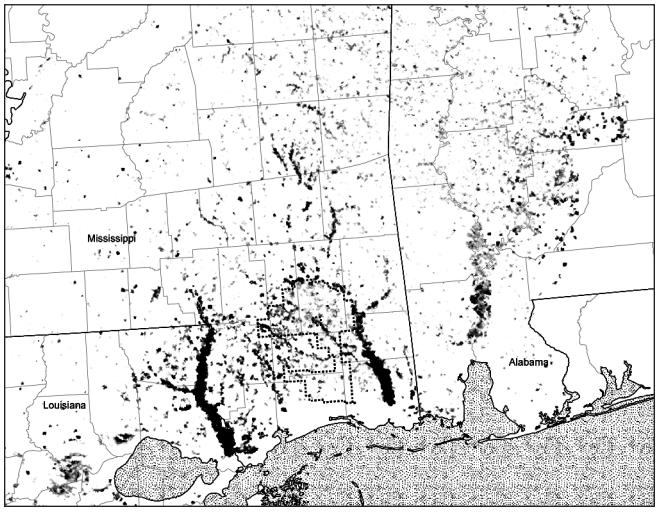
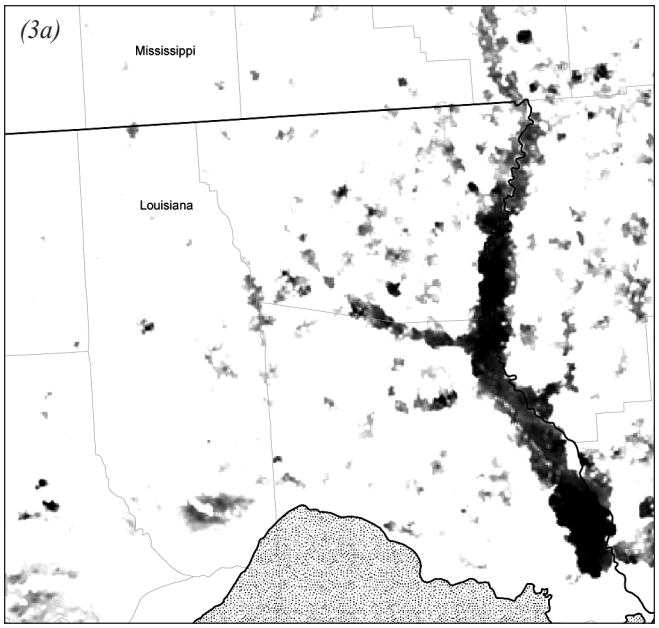
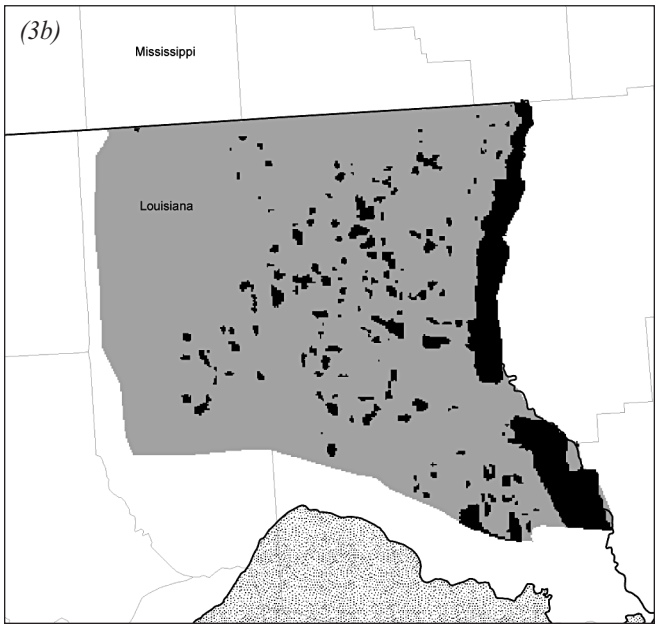


Figure 3.—*The correspondence of forest damage as mapped using Moderate Resolution Imaging Spectroradiometer (MODIS) data (fig. 3a) with Forest Service Forest Health Protection aerial sketchmap data (fig. 3b, Steiner et al. 2006). In the MODIS map, darker grays represent more intense damage. In the sketchmap data, areas in black were mapped as damaged forest, while gray indicates the area flown. Misregistration is present, but the general patterns are similar, with extensive damage in the floodplains of the Pearl River (including a more lightly damaged section in the southern part of the mapped area) and its tributary from the west, the Bogue Chitto, and scattered damage elsewhere.*





classes are shown in figure 4, while similar curves in figure 5 illustrate the damage histograms corresponding to the three forest types. In all cases, the expected associations between the auxiliary data and predicted damage were found, and the 99-percent confidence intervals for all class means were distinct and nonoverlapping.

Rapid Mapping Feasibility. Forward stepwise regression based on only the reflectance and NDVI differences between the first poststorm MODIS image collected—3 days after Katrina's landfall—and the prestorm image indicated that the model was best constructed with only the red and near-infrared reflectance differences and had an adjusted R^2 of 0.522. Application of the equation across the study area yielded a predicted map (not shown) that looked remarkably similar to that based on the complete time-series data. The primary differences between the two maps occurred in less-strongly affected areas in the central portion of the study area. Determination of the relative accuracies of the maps was impossible using the available reference data. The immediate poststorm model, however, matched the Steiner *et al.* (2006) sketchmap data slightly better than the full time-series model. Again, all the expected relationships against auxiliary data were found at statistically significant levels.

Rapid Mapping Feasibility Using Auxiliary Data. Forward stepwise regression found that the best predictive model for forest damage based on auxiliary data was constructed using the maximum sustained wind speed and the fractions of the pixel occupied by deciduous upland forest and wooded wetland forest. The equation generated, which had an adjusted R^2 of 0.434, led to the damage prediction shown in figure 6. Although the overall spatial patterns found in the other predictions are seen here also, the map has a blurriness that could likely be improved only by having more detailed forest type and stand history information. As it is, the auxiliary data prediction provides at least a useful spatial envelope within which most forest damage can be expected to occur.

Figure 4.—*Normal curves fitted to the predicted damage histograms for four maximum sustained wind speed classes, corresponding to weak tropical storm force winds (< 60 mph), strong tropical storm force winds (60 to 74 mph), Category 1 hurricane force winds (75 to 96 mph), and Category 2 hurricane force winds (96 to 111 mph). Higher wind speeds correspond to greater predicted damage, and the 99-percent confidence intervals for the class means are all nonoverlapping.*

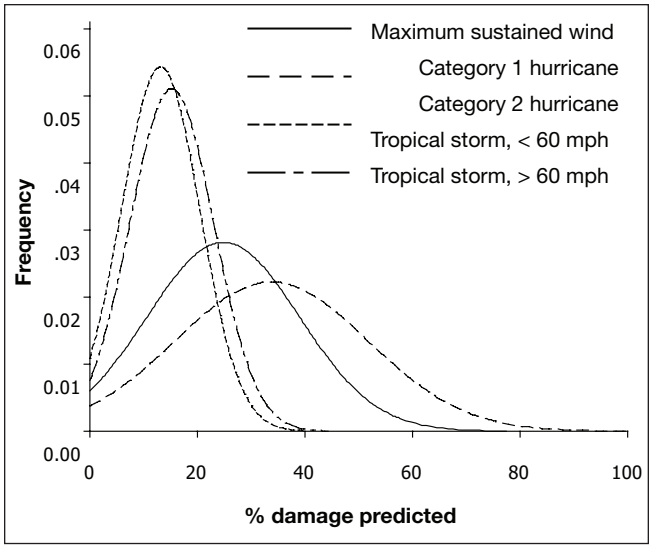


Figure 5.—*Normal curves fitted to the predicted damage histograms for three forest type classes, dominated by deciduous upland forests, evergreen upland forests, and wooded wetlands (predominantly bottomland hardwood forests). Damage is predicted to concentrate in the wooded wetland class, with evergreen upland forests experiencing somewhat more damage than deciduous upland forests. These results are in keeping with the generalizations of Gresham et al. (1991) and Basnet et al. (1992) and with the specific observations of Windham (2005).*

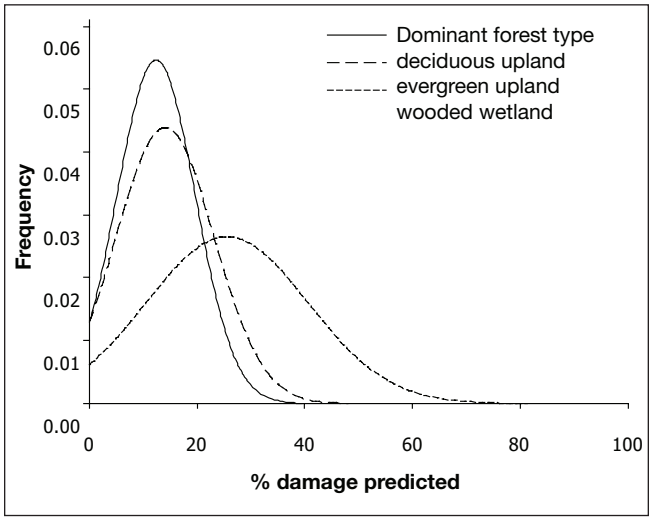
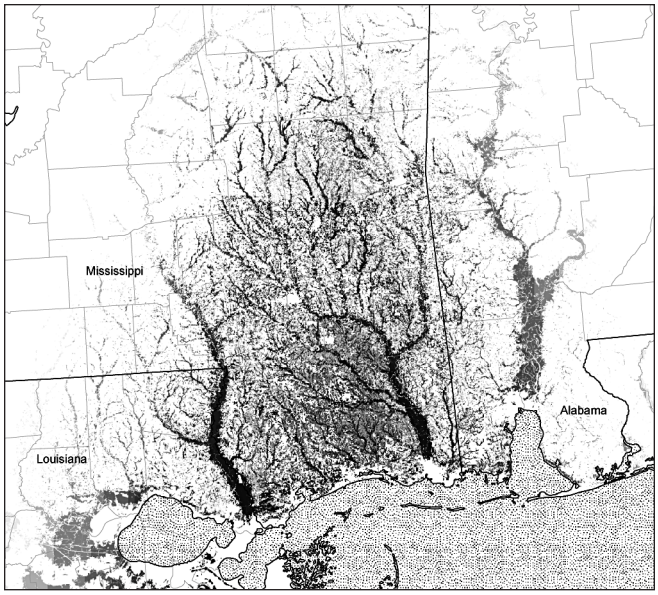


Figure 6.—*Damage estimate predicted on the basis of auxiliary data alone. The model is based on maximum sustained wind speed and pixel percentage occupation by deciduous upland forest and wooded wetland forest. This more general map can serve as a first approximation in the absence of satellite data or as a spatial constraint on predictions based on satellite data.*



Discussion

The results here are encouraging for the prospects of rapid mapping of hurricane damage using MODIS imagery, auxiliary data, or some combination thereof. In fact, immediate postevent mapping has some advantages over use of time-series data: by reducing the analysis window to the shortest possible duration, unrelated land cover change is less likely to become confounded with disturbance due to the event. Considering the variety of possible temporal biological responses to an event, it may be hard to take full advantage of time-series data anyway, and the advantages of rapid mapping may outweigh its disadvantages. In addition, it is possible that the higher R^2 found here for the prediction based on the full time-series data resulted from an overfit to noise or other local phenomena in the training data and did not result in a more accurate map over the full study area.

It is quite likely that rapid postdisturbance mapping will have to occur in the absence of good, clear-sky satellite data. Although the results generated here for predicting damage using auxiliary data provide a starting point, three degrees of separation existed between the primary training data and the auxiliary damage modeling: photo interpretation, derivation of the damage map across the DeSoto District using TM imagery, and extrapolation to the full affected area using MODIS imagery. It would be helpful and probably more reliable to derive relationships based on other results that are more closely tied to field research or detailed air photo work. It is possible that such relationships could be exploited using more detailed forest type maps, which may be forthcoming through the LANDFIRE project or Regional Gap Analysis program.

More accurate rapid mapping could be accomplished by combining satellite imagery with auxiliary data, either through the use of the auxiliary variables as an envelope with which to constrain the spectral estimates or through some more complex multivariate procedure. An approach based on canonical correspondence analysis, in which a spectral damage index is derived through constraint of spectral change scores to linear combinations of auxiliary variables known to influence damage susceptibility, might be helpful.

The fact that floodplain forests appear to be more susceptible to hurricane damage creates one difficulty. These are the same areas that experience heavy flooding in the days and possibly weeks after an event. Change metrics based on simple NDVI differences may be particularly susceptible to the confusion of understory flooding with forest damage, especially in fairly open forests. The strong similarity between the map produced here and the FHP sketchmap is some assurance that the problem is not very severe. Despite this assurance, mid-infrared spectral bands, which are available from the MODIS sensor at 500-m resolution, would likely be helpful in discriminating flooding from forest damage. They could also assist in separating defoliation from more severe structural damage, as indices based on mid-infrared reflectance have been found to have high sensitivity to changes in forest structure (e.g., Jin and Sader 2005). Investigation of the utility of the mid-infrared bands is another matter for further work.

It is important to note that this work makes the assumption that the Landsat TM-based damage maps used as reference data accurately reflect genuine forest damage. However, it is un-

certain whether temporary effects such as defoliation could be fully separated from structural damage based on interpretation of aerial photography collected soon after the storm, and any shortcomings in that process would have propagated through the subsequent TM and MODIS analyses. Conditions during the following year's growing season likely provide a more reliable indication of persistent damage. That is the approach taken by one recently published study of Hurricane Katrina's impacts (Chambers *et al.* 2007). Greater confidence could be placed in these results if comparison with data derived from subsequent years were made.

Conclusions

The results here are encouraging, but only a beginning. To make accurate rapid assessment an operational procedure, it will be necessary to investigate whether such appraisals can be achieved in the absence of event-specific training data. The answer to that question will begin with determining the constancy of forest spectral response to particular damage types and the repeatability of patterns of damage with respect to auxiliary variables. If generalizable relationships cannot be relied upon, the investigation of multivariate techniques by which satellite and auxiliary data can be combined to isolate and reveal unique damage signatures may be profitable.

Acknowledgments

Jess Clark at the Remote Sensing Applications Center (RSAC) produced the damage map that served as the essential training dataset for this work. His work in turn relied on aerial photography collected by the Forest Service Forest Health Technology Enterprise Team. Chris Steiner's work at the Forest Health Protection program enabled the visual validation of the damage map shown here. RSAC's rapid assessment work is funded by the Forest Inventory and Analysis program (FIA), and I thank Dennis Jacobs and other members of the FIA Remote Sensing Band for interesting and amusing discussions.

Literature Cited

Atlantic Oceanographic and Meteorological Laboratory (AOML) Hurricane Research Division. 2006. Katrina 2005 post storm wind analyses. http://www.aoml.noaa.gov/hrd/Storm_pages/katrina2005/wind.html. (October).

Armentano, T.V.; Doran, R.F.; Platt, W.J.; Mullins, T. 1995. Effects of Hurricane Andrew on coastal and interior forests of southern Florida: overview and synthesis. Journal of Coastal Research. 21: 111-144.

Ayala-Silva, T.; Twumasi, Y.A. 2004. Hurricane Georges and vegetation change in Puerto Rico using AVHRR satellite data. International Journal of Remote Sensing. 25: 1629-1640.

Basnet, K.; Likens, G.E.; Scatena, F.N.; Lugo, A.E. 1992. Hurricane Hugo: damage to a tropical rainforest in Puerto Rico. Journal of Tropical Ecology. 8: 47-55.

Boose, E.R.; Foster, D.R.; Fluet, M. 1994. Hurricane impacts to tropical and temperate forest landscapes. Ecological Monographs. 64: 369-400.

Chambers, J.Q.; Fisher, J.I.; Zeng, H.; Chapman, E.L.; Baker, D.B.; Hurtt, G.C. 2007. Hurricane Katrina's carbon footprint on U.S. Gulf Coast forests. Science. 318: 1107.

Clark, J.; Megown, K.; Jacokes-Mancini, R. 2006. Hurricane Katrina damage assessment on lands managed by the DeSoto National Forest using multi-temporal Landsat TM imagery and high resolution aerial photography. In: Greer, J., ed. Proceedings, 11th biennial remote sensing conference, RS-2006. Salt Lake City, UT: U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center.

Cohen, W.B.; Maiersperger, T.K.; Gower, S.T.; Turner, D.P. 2003. An improved strategy for regression of biophysical variables and Landsat ETM+ data. Remote Sensing of Environment. 84: 561-571.

Doyle, T.W.; Keeland, B.D.; Gorham, L.E.; Johnson, D.J. 1995. Structural impact of Hurricane Andrew on the forested wetlands of the Atchafalaya Basin in South Louisiana. Journal of Coastal Research. 21: 354-364.

Everham, E.M.; Brokaw, N.V.L. 1996. Forest damage and recovery from catastrophic wind. Botanical Review. 62: 113-185.

Gresham, C.A.; Williams, T.M.; Lipscomb, D.J. 1991. Hurricane Hugo wind damage to southeastern U.S. coastal forest tree species. Biotropica. 23: 420-426.

Jacobs, D.M. 2006. Rapid assessment model for determining extent of hurricane damage. Poster presentation at the Forest Health Monitoring Program work group workshop; 2006 January-February; Charleston, SC.

Jin, S.; Sader, S.A. 2005. Comparison of time series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances. Remote Sensing of Environment. 94: 364-372.

Kelly, J.F. 1993. Hurricane Andrew forest damage assessment. World Resource Review. 5: 401-408.

Kovacs, J.M. 2001. Mapping disturbances in a mangrove forest using multi-date Landsat TM imagery. Environmental Management. 27: 763-776.

Leininger, T.D.; Wilson, A.D.; Lester, D.G. 1997. Hurricane Andrew damage in relation to wood decay fungi and insects in bottomland hardwoods of the Atchafalaya Basin, Louisiana. Journal of Coastal Research. 13: 1290-1293.

Loope, L.; Duever, M.; Herndon, A.; Snyder, J.; Jansen, D. 1994. Hurricane impact on uplands and freshwater swamp forest. Bioscience. 44: 238-246.

McNab, W.H.; Greenberg, C.H.; Berg, E.C. 2004. Landscape distribution and characteristics of large hurricane-related canopy gaps in a southern Appalachian watershed. Forest Ecology and Management. 196: 435-447.

Platt, W.J.; Doren, R.F.; Armentano, T.V. 2000. Effects of Hurricane Andrew on stands of slash pine (*Pinus elliottii* var. densa) in the Everglades region of south Florida (USA). Plant Ecology. 146: 43-60.

Ramsey, E.W.; Chappell, D.K.; Jacobs, D.M.; Sapkota, S.K.; Baldwin, D.G. 1998. Resource management of forested wetlands: hurricane impact and recovery mapped by combining Landsat TM and NOAA AVHRR data. Photogrammetric Engineering and Remote Sensing. 64: 733-738.

Ramsey, E.W.; Hodgson, M.E.; Sapkota, S.K.; Nelson, G.A. 2001. Forest impact estimated with NOAA AVHRR and Landsat TM data related to an empirical hurricane wind-field distribution. Remote Sensing of Environment. 77: 279-292.

Steiner, C.A.; Bruce, W.W.; Johnson, C.W.; Petty, S.D. 2006. Results of aerial surveys of forested lands damaged by Hurricanes Katrina and Rita in southeast and southwest parishes, Louisiana. Poster presentation at the Forest Health Monitoring Program work group workshop; 2006 January-February; Charleston, SC.

Walker, L.R. 1995. Timing of post-hurricane tree mortality in Puerto Rico. Journal of Tropical Ecology. 11: 315-320.

Windham, J.W. 2005. Forest vegetation analysis, Hurricane Katrina damaged tree removal and hazardous fuels treatment, DeSoto National Forest. Jackson, MS: U.S. Department of Agriculture, Forest Service, National Forests in Mississippi. 6 p.

LIDAR Forest Inventory With Single-Tree, Double- and Single-Phase Procedures

Robert C. Parker[1] and David L. Evans[2]

Abstract.—Light Detection and Ranging (LIDAR) data at 0.5- to 2-m postings were used with double-sample, stratified inventory procedures involving single-tree attribute relationships in mixed, natural, and planted species stands to yield sampling errors (one-half the confidence interval expressed as a percentage of the mean) ranging from ±2.1 percent to ±11.5 percent at α=0.05. LIDAR sample trees were selected with focal filter procedures and heights were computed as the difference between interpolated canopy and digital elevation model surfaces. Tree diameter at breast height (d.b.h.) and height data were obtained on LIDAR ground samples ranging from a 5:1 ratio on 0.08-ha rectangular strips to a 10:1 ratio on 0.02-ha circular plots established with a real-time Differential Global Positioning System. D.b.h.-height and ground-LIDAR height models were used to predict d.b.h. from adjusted LIDAR height and compute phase 2 ground and LIDAR estimates of basal area and volume. Phase 1 LIDAR estimates were computed by randomly assigning heights to species classes using the probability distribution from ground plots in each inventory strata. Phase 2 LIDAR estimates were computed by randomly assigning heights to species-product groups using a Monte Carlo simulation for each ground plot. No statistical difference was present between double-sample, mean volume estimates from 0.5-m and 1-m LIDAR posting densities with and without height bias adjustment or on smoothed and unsmoothed LIDAR canopy surfaces. Volume estimates from single-phase LIDAR inventory procedures using existing tree attribute and LIDAR-ground height bias relationships were obtained with sampling errors of 1.8 percent to 5.5 percent for full and minimized data sets to test minimum LIDAR inventory requirements.

Introduction

Light Detection and Ranging (LIDAR) is a relatively new remote-sensing tool that has the potential for use in the acquisition of measurement data for inventories of standing timber. LIDAR systems have been used in a variety of forestry applications (Magnussen and Boudewyn 1998, Lefsky *et al.* 1999, Means *et al.* 2000) for the quantification of biomass (Nelson *et al.* 2003), basal area, and tree and stand height estimates. Stand-level LIDAR inventory procedures involving average values of tree attributes such as dominant height, mean diameter, basal area, and volume have been applied to obtain unbiased stand-level predictions (Naesset 2002, 2004; Popescu *et al.* 2002). Because LIDAR has the capability to detect individual trees and measure tree height with predictable bias when correlated with ground measurements (Holmgren 2004, Persson *et al.* 2002), strata-level inventory estimates involving individual tree, double-sample inventory procedures have been used by researchers from Mississippi State University in conifer and mixed hardwood stands in the Northwest and Southeast (Collins 2003; Parker and Evans 2004, 2006; Parker and Glass 2004; Parker and Mitchel 2005; Williams 2006). The individual tree approach to stand inventory, when combined with double-sample, ground procedures, permits relatively precise estimates of volume with a simple prediction function for ground-LIDAR height bias and ground-based attribute relationship functions for tree diameter and total height that can be used with any standard, standing tree volume function. Stand-level approaches

[1] Professor, Forest Biometrics, Department of Forestry, Forest and Wildlife Research Center, Mississippi State University, Mississippi State, MS 39762. Phone: 662–325–2775. E-mail: rparker@cfr.msstate.edu.

[2] Professor, Spatial Technologies, Department of Forestry, Forest and Wildlife Research Center, Mississippi State University, Mississippi State, MS 39762. Phone: 662–325–2796. E-mail: devans@sitl.cfr.msstate.edu.

involving average tree attribute values for sampling units require more sophisticated prediction models than an individual tree approach and procedures that differ radically from traditional ground-based inventory methods. The objective of this article is to summarize and discuss the procedures, models, advantages, and disadvantages of the single-tree approach to using LIDAR data in double- and single-phase forest inventory methods.

Flight Planning for a LIDAR Inventory

Small-footprint, multireturn LIDAR data have been acquired with various sensors to attain nominal posting spacings of 0.5 to 2.0 m, 0.25 to 4 point/m^2, and footprint sizes of 0.122 to 0.330 m for two returns per pulse (table 1). Aircraft altitudes of 600 to 1,000 m and swath widths of 189 to 609 m were used. The minimum required density of LIDAR hits is a function of the crown size, average height, and spatial density of the sample trees in the primary canopy. Acceptable sampling statistics were attained for sparse densities of large-crown conifers in Idaho (± 11.5 percent sampling error at the 95-percent confidence level with a standard error of ± 5 m^3, Parker and Evans [2004]) with 0.25 points/m^2; however, 1 point/m^2 was required to achieve acceptable inventory results in natural pine and mixed pine-hardwood stands (± 7.6-percent sampling error, Parker and Glass [2004]) and 2 points/m^2 in young (6+ years) pine plantations (± 2.2-percent sampling error, Parker and Evans [2006]) in the Southeast. Increasing LIDAR density from 2 to 4 points/m^2 did not statistically improve the volume estimation precision and the increased "noise" in the high-density LIDAR data translated into additional sampling error about the volume estimate.

Target aircraft altitude is a function of the desired swath width and scan angle for the LIDAR pulse generator and sensor and the technical ability of the sensor to achieve the desired posting density. The swath width diminishes as the desired posting density increases, but reasonable swath widths can be achieved with 2 to 4 points/m^2. An important factor influencing desired swath width was the size of the ground-based sample plots used in the inventory procedure. The swath should be sufficiently wide to encompass the sample ground and LIDAR plots within the center one-third of the swath so as to minimize the "edge effects" of the LIDAR data. Tree attribute measurements are severely compromised at the extremes edges of the swath and scan angle.

Percentage of LIDAR coverage is a function of economics and inventory design. LIDAR data are relatively expensive to obtain and complete area coverage is normally not required for most timber inventory designs. In some instances, the cost of complete LIDAR coverage to produce an accurate, up-to-data digital elevation model may be more justifiable than the expense for a timber inventory. The use of a current Geographic Information System (GIS) to locate flight lines that cross the desired sampling strata can minimize the percentage of coverage of the LIDAR area. Most forested areas to be inventoried can be flown with 10 percent or less LIDAR area coverage by orienting flight lines so as to cross the target inventory strata at desired flight-line intervals.

Field and LIDAR Plot Design and Procedures

Inventory design for the single-tree LIDAR applications involved the use of circular (Parker and Evans 2006, Parker and

Table 1.—*LIDAR specifications for sample single-tree inventory projects in conifer and mixed.*

LIDAR specification	Example 1 (Idaho)	Example 2 (CP LA)	Example 3 (FW LA)	Example 4 (CP LA)
Points per square meter	0.25	1	1.9	4
Nominal spacing (m)	2.0	1.0	0.7	0.5
Footprint size (m)	0.330	0.213	0.250	0.122
Aircraft altitude (m)	1,000	1,067	1,000	610
Swath width (m)	600+	609	243	189
Tract size (ha)	2,023	485	18,000	485
Percentage of LIDAR coverage (%)	1.43	100	10	100

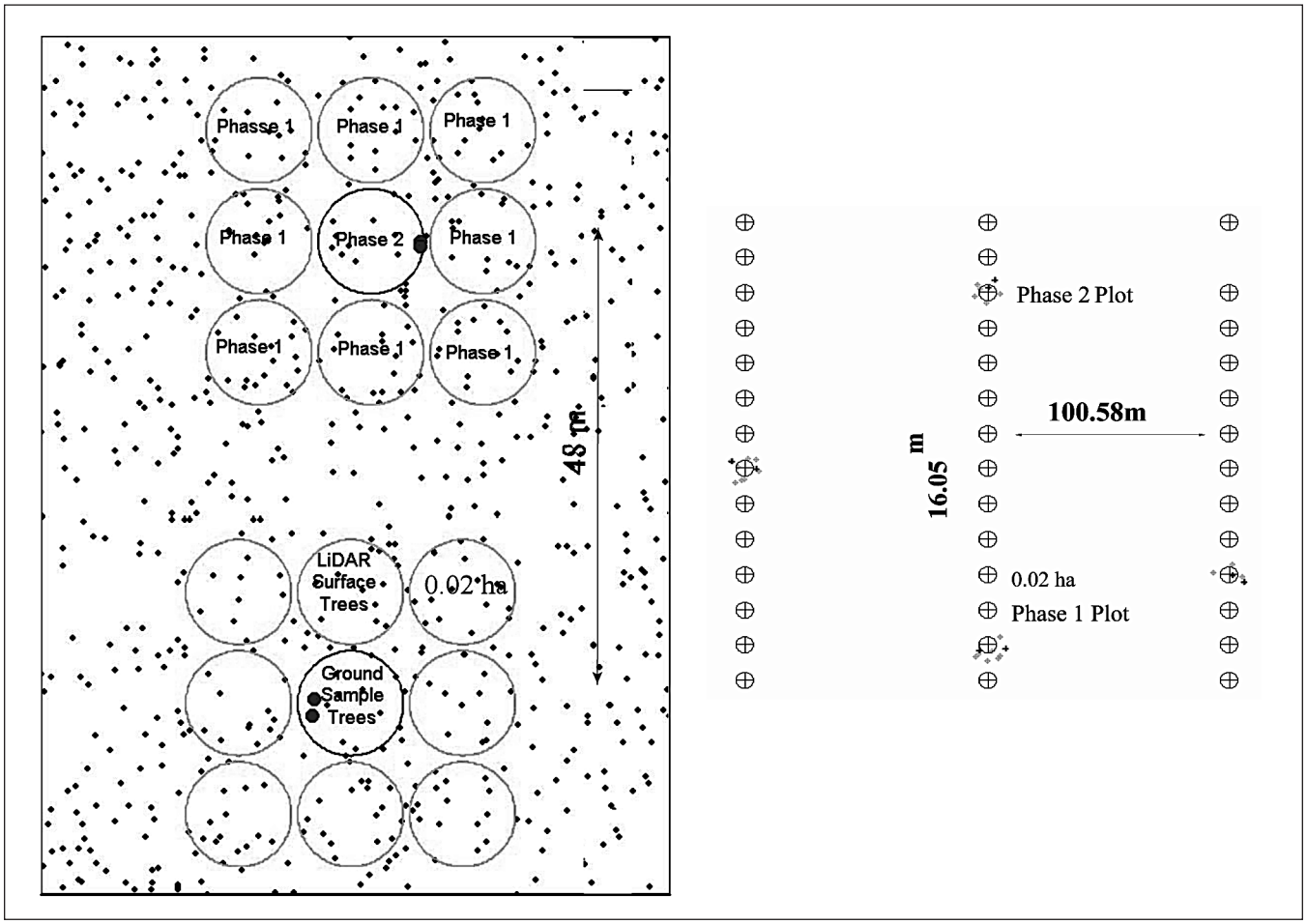
CP = coastal plain. FW = flatwood. LA = Louisiana. LIDAR = Light Detection and Ranging.

Glass 2004) or rectangular plots (Parker and Evans 2004) with all plots being phase 1 LIDAR plots and every rth plot being a phase 2 ground plot. Field designs varied from a 9:1 ratio of LIDAR to ground circular plots in a nested arrangement to a 10:1 ratio with rectangular or circular plots along a flight line (fig. 1). Universal Transverse Mercator (UTM) coordinates were established at the center of each circular phase 2 plot or at the endpoints of rectangular plots for navigation with a real-time Differential Global Positioning System (DGPS). Differential corrections from either the U.S. Government Wide Area Augmentation System or private-enterprise OmniSTAR geostationery satellite were obtained satisfactorily under tree canopies by using a large dome antenna. Based on informal field tests on surveyed bench marks, field locations were obtained with approximately 1-m accuracies with both systems.

LIDAR Surfacing for Tree Location and Height Determination

The LIDAR data were processed to produce a ground surface or digital terrain model (DTM) and a tree surface for the determination of sample tree locations and tree heights within the sample field and LIDAR plot areas. LIDAR data sets were surfaced to produce first-return canopy and last-return DTM with 0.2-m cell sizes using a linear interpolation technique. Tree locations and heights were determined with algorithms and focal filter procedures developed by McCombs *et al.* (2003) that used a variable search window radius based on relative density. These procedures used moving 2.5-, 4.0-, or 5.5-ft radius search windows to identify each tree peak as the point that is higher than 85 percent of the surrounding maxima from one of the three search window radius files. Tree height was interpreted as

Figure 1.—*Sample designs for 0.02-ha plots with 9:1 and 10:1 ratios of Light Detection and Ranging plots (phase 1) to ground (phase 2) circular plots in a nested arrangement and plots along a flight line.*



the difference between canopy and DTM z-values at each tree peak location. Tree heights were converted to point coverages and clipped to sample area boundaries using UTM coordinates to describe sample plot locations and sizes.

A spatial filtering technique derived from image analysis called smoothing was used to reduce commission errors by minimizing the abrupt elevation changes in the initial canopy surface. The Focal Analysis option in ERDAS' IMAGINE software performed smoothing based on user-defined inputs for window size and preferred statistical procedure. A 5-by-5 pixel window was used to create a 1-m^2 filter that would avoid removal of small peaks in the canopy surface (small trees) while maximizing the smoothing function. The filter moved across the LIDAR canopy surface, pixel by pixel, averaged the values within the window, and placed the result in the center pixel.

Smoothing heights on LIDAR surfaces improved the relationship between LIDAR and ground tree heights in terms of R^2, reduced height biases for hardwoods, increased height biases for pines, and improved target recognition in terms of trees-per-acre estimates; however, no statistical differences (α =0.05) occurred between double-sample regression volume estimates with smoothed versus unsmoothed LIDAR surfaces from low- or high- density LIDAR. Standard errors and sampling errors of the regression estimates were lower for all unsmoothed LIDAR data models than with smoothed data models. Thus, smoothing heights on LIDAR surfaces did not produce a statistical gain for volume estimation using double-sample procedures.

Double-Sample, Regression Estimator Procedures

The double-sample model widely used with ground-based point sampling (Avery and Burkhart 2002) and aerial photogrammetric inventories and adapted for these studies was the following:

$$\overline{Y}_{lr} = \overline{y}_2 + \beta(\overline{X}_1 - \overline{x}_2) \quad (1)$$

With traditional aerial photogrammetric inventories, the X_{1i} and x_{2i} variables are photographic volume/unit area and ground volume/unit area from phase 1 and phase 2 plots, respectively, and β is the regression slope coefficient for y_i (ground volume) over x_{2i} (photo volume) on ground plots. Thus, a phase 1 (large sample) variable such as remotely sensed (i.e., photographic or LIDAR-derived) volume has a strong, identifiable relationship with a phase 2 (small sample) variable such as ground volume.

In applications of the double-sample model with single-tree LIDAR data, phase 2 sample tree measures of diameter at breast height (d.b.h.) and height were used to derive height-d.b.h. and d.b.h.-height equations of the model type:

$$H_{gr} = b_0 + b_1[\ln(\text{d.b.h.})]^{b2}[\text{age}]^{b3} + \varepsilon \quad (2)$$

$$\text{d.b.h.} = b_0 + b_1[\ln(H_{gr})]^{b2}[\text{age}]^{b3} + \varepsilon \quad (3)$$

where H_{gr} was ground measured height, d.b.h. was ground measured d.b.h., and age was average stand age (years) from GIS data. The age variable in models (2) and (3) was removed when age did not contribute significantly to the relationship. Models (2) and (3) were derived from the ground-measured sample trees, but one is not a back transformation of the other. Model (2) was applied to ground-plot trees where d.b.h. was measured on all trees and heights on a subsample. Model (3) was applied to LIDAR-derived tree heights to obtain a d.b.h. for single-tree volume computation.

Generally, only two trees per ground plot were measured for height; d.b.h. was measured on all trees. The height-d.b.h. model (2) was applied to trees on the ground plots for which height was not measured to obtain a height for single-tree volume computation. Sample tree heights from the phase 2 ground plots were used to predict the ground height of target trees identified on LIDAR surfaces. The d.b.h.-height model (3) was applied to the bias-adjusted, single-tree LIDAR height from the ground-LIDAR height bias model:

$$H_{gr} = b_0 + b_1(H_{Li}) + \varepsilon \quad (4)$$

where H_{gr} was measured ground height of trees on phase 2 plots and H_{Li} was interpolated height of the same trees from the LIDAR surface.

Derived d.b.h. on LIDAR plots and derived height on ground plots permitted the use of a standard, standing-tree volume equation with d.b.h. and height as variables to predict volume. Thus, the double-sample models used in this study involved

LIDAR mean estimates of basal area (LiBA from phase 1and liba from phase 2 with matching ground plot) and volume (LiVOL from phase 1 and livol from phase 2 with matching ground plot) for the x-variables as

$$\overline{Y}_{lr} = \overline{y} + \beta(LiBA - liba) + \varepsilon \quad (5)$$

$$\overline{Y}_{lr} = \overline{y} + \beta(LiVOL - livol) + \varepsilon \quad (6)$$

with variance

$$s^2_{\overline{y}_{lr}} = \frac{s^2_{yx}}{n_2} + \frac{s^2_y - s^2_{yx}}{n_1} \quad (7)$$

where $\overline{y}$ was phase 2 mean ground volume, β was the regression slope coefficient for y_i (ground volume/unit area) over x_{2i} (LIDAR volume/unit area or basal area/unit area on ground plot), and x_{1i} was volume or basal area on the LIDAR plot. Data were fitted to models (5) and (6) for all data combined (i.e., nonstratified), each age-class strata, and combined strata.

Combined strata, linear regression estimates of volume and associated standard error of each double-sample model were obtained by using the following equation:

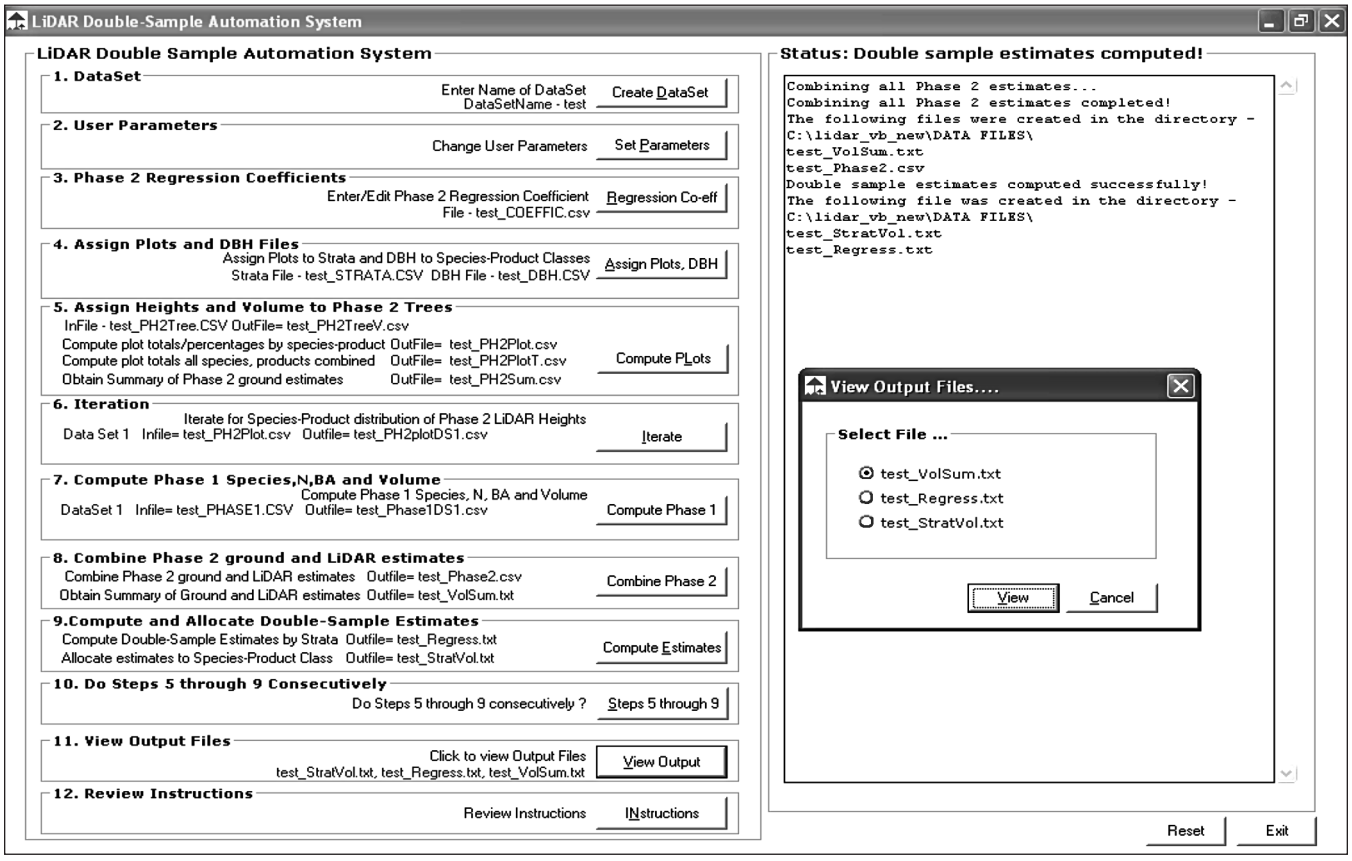
$$\overline{Y}_{lr,c} = \sum \frac{(n_{1i} + n_{2i})}{N} (\overline{Y}_{lr,i}) \quad (8)$$

$$S_{\overline{Y}_{lr,c}} = \left[\sum \left(\frac{(n_{1i} + n_{2i})}{N} \right)^2 (S^2_{\overline{y}_{lr,i}}) \right]^{0.5} \quad (9)$$

where n_{1i} and n_{2i} were phase 1 and 2 sample sizes, respectively, for stratum i, i=1 to X strata.

All double-sample volume computations were performed with the Windows-based software program LiDAR Double Sample Automation System (fig. 2) developed by Parker (2005). The software allowed the user to specify d.b.h. limits for species-product classes, regression coefficients for the d.b.h.-height and ground height-LIDAR height models, and stratum definitions of beginning and ending plot numbers and average age and enter

Figure 2.—*Sample computer screen for the Light Detection and Ranging (LIDAR) Double Sample Automation System used for LIDAR double-sample inventory procedures.*



comma-delimited data files of phase 1 LIDAR heights and phase 2 ground-plot trees (species, product, d.b.h., and height of sample trees). LIDAR heights in the phase 1 data were allocated in a Monte Carlo simulation to species-product classes on each matching phase 2 ground plot on the basis of percentage of distribution by numbers on the ground plot. Because species and d.b.h. of the LIDAR-derived trees are unknown, the Monte Carlo simulation (50 iterations) would randomly allocate the LIDAR-derived trees (d.b.h. predicted from adjusted LIDAR-to-ground height) to species-product classes and obtain a mean basal-area and volume estimate for the species-product class. Thus, basal-area and volume estimates from phase 1 LIDAR plots that had a matching phase 2 ground plot became phase 2 LIDAR plots. Phase 1 LIDAR heights that did not have a matching phase 2 ground plot were randomly allocated to encountered species classes in each stratum in a single iteration and used to compute mean estimates of numbers of trees, basal area, and volume. Phase 2 tree measures of d.b.h. and height were used to compute LIDAR estimates of mean basal area and volume by using field-derived d.b.h.-height equations to predict d.b.h. from LIDAR height and volume. Predicted d.b.h. and height were used in a single-tree volume function to predict individual/single tree volume. Double-sample volume estimates and associated precision statistics were computed with models (5) and (6) for each stratum and with models (8) and (9) for combined strata.

Single-Phase Inventory Procedures

A recent study by Williams (2006) investigated the minimal data inputs for a LIDAR based timber inventory with single-phase procedures. Because the ground phase of a double-sample field procedure is both expensive and time consuming, Williams investigated the feasibility of using LIDAR data in single-tree approach to obtaining volume estimates by stratum with a single-phase inventory procedure. Previous studies have shown that LIDAR can provide precise but biased estimates of tree numbers and heights. If the assumptions are made that the LIDAR height bias is known and relatively constant for a given species-origin class (i.e., pine plantations) and previously established tree attribute relationships are also known, inventory estimates of volume can be obtained with a single-tree approach and single-phase procedures from LIDAR data only.

The tree attribute relationship from model (3) was developed from ground measurements of d.b.h. and height within continuous forest inventory (CFI) plots and from the phase 2 ground plots in the double-sample approach by Parker and Evans (2006). Sample trees were randomly selected from the original data sets in groups of 75 and fitted to tree attribute model (3) under the assumption that ground data were available from previous studies. The ground-LIDAR height bias equation obtained from model (4) (Parker and Evans 2006) was assumed to be constant and known. LIDAR-derived heights from phase 1 plots were adjusted for bias with model (4) then used with model (3) to obtain single-tree d.b.h. estimates for use with a single-tree volume function in a conventional, single-phase inventory processor. The stratum volume estimates and precision statistics were compared to estimates obtained from the phase 2 ground plots by Parker and Evans (2006).

The single-tree, single-phase volume estimates from LIDAR data compared favorably with ground-plot estimates (table 2). At the tract level for 20 age-class strata on 10,443 ha, no statistical difference between the single-phase LIDAR estimates and the ground-plot estimate was present. Single-phase volume estimates were obtained for the full data set of phase 1 LIDAR

Table 2.—*Comparison of single-tree, single-phase, volume predictions from LIDAR data using 0.02-ha circular plots on 10,443 ha of pine plantations with ground-plot volume estimates where sampling error was half the (1-α) confidence interval expressed as a percentage of the mean.*

Data set description	Number of plots	Sampling error (%)	Tract volume compared with control
Ground plots—phase 2	842	2.8	Control
LIDAR plots—phase 1	7,562	1.8	Not statistically different
LIDAR plots—phase 2	842	5.2	Not statistically different
LIDAR plots—phase 1	5:1 ratio of Phase 1:2	5.5	Not statistically different in three of five iterations

LIDAR = Light Detection and Ranging.

plots, the phase 1 LIDAR plots that had a matching phase 2 ground plot, and five iterations of reducing the LIDAR data set to a 5:1 ratio with ground plots within each stratum. The Williams (2006) study found no statistical difference (α=0.05) between tract-level, single-phase volume estimates where the single-tree relationship model was developed with 1,539 sample trees from the phase 2 ground plots by Parker and Evans (2006), 1,509 trees from the regional CFI plots, or 5 iterations of 75 randomly selected trees from the phase 2 data set. The study concluded that precise, single-phase LIDAR inventory estimates are feasible with minimal inputs of ground data for establishing tree attribute relationships. A potential application of the single-tree, single-phase inventory procedure would be the rapid post-thinning inventory of pine plantations and periodic inventories of forested holdings.

Conclusions about Single-Tree LIDAR Inventory Procedures

LIDAR provides precise x, y, and z coordinate data that can be used to extract tree heights and locations; however, several sources of bias exist that can impact the accuracy of a per-unit area volume estimate. Height bias is primarily caused by the failure of the laser pulse to hit the terminal leader, but this bias can be predicted with acceptable success in confers but not in hardwoods with broad, rounded crowns. Height bias can also be introduced by the interpolation of tree heights with mixed linear and nonlinear procedures within the same data set. Tree count bias has at least two sources of origin: (1) trees in the mid and lower canopy layers are hidden from the laser pulse by a dominant canopy, and (2) tree maxima locations may not be interpreted correctly during the LIDAR surfacing and height extraction process.

The precision of volume estimates with single-tree LIDAR procedures in a double-sample process is not affected by the height or tree count bias inherent in LIDAR data. These biases are effectively adjusted during the double-sample inventory procedures. The height bias can be adjusted before single-tree volume computations with LIDAR-derived heights or afterwards during the double-sample volume adjustment process. Height bias correction before volume computation improves the accuracy of the resulting per-unit area volume estimate. The tree count bias caused by canopy coverage or LIDAR surfacing or processing is also effectively adjusted through the double-sample volume computations. Tree count bias, however, has a major impact on the accuracy of volume estimates in a single-tree procedure computed with single-phase inventory methods. Thus, if any doubt exists about the validity of the tree counts or height bias during LIDAR processing, a double-sample volume computation process should be used.

Ground sample tree measurement is needed to establish the ground-LIDAR height bias and the relationship between standing tree height and d.b.h. The number of sample trees is dependent upon the number of parameters in the regression models and variation in the data. A reasonable rule of thumb is 25 samples per parameter estimated in a regression model. Because the LIDAR height bias is relatively constant for a species in a local area and the d.b.h.-height relationship for a give set of species-site conditions is also relatively stable, the sample trees can be obtained from either the LIDAR inventory or surrounding forested areas.

Establishing the LIDAR to ground tree height bias requires the matching of trees on the ground and on the LIDAR surface. Past experience has shown that the location of a plot center must be done with a real-time DGPS and the distance and direction to the sample trees from the plot center should be obtained with a laser so that the x, y coordinates of the sample trees can be located on the LIDAR surface.

Sample plot size and shape on the ground and on the LIDAR surface should be a function of tree density on the ground and the LIDAR processing procedures employed. Experience has shown that the ground-plot size should be adjusted such that a minimum of 6 and a maximum of approximately 15 trees should be selected. The minimum number is associated with the with- and between-plot variation and the maximum is a logistical consideration for minimizing omission or commission errors in tallying trees. Rectangular plots are easier to handle during the LIDAR processing but are more difficult to establish in the field and to use in establishing distance and direction to sample trees from a DGPS location. The cost of a LIDAR inventory can be minimized by flying only a portion of the

desired inventory area. As long as the LIDAR swaths cover the desired strata, ground plots could be located randomly or systematically within strata within swaths. DGPS permits the location of sample plots and trees with relative ease and precision.

Literature Cited

Avery, T.E.; Burkhart, H.E. 2002. Forest measurements. 5th ed. New York: McGraw-Hill. 456 p.

Collins, C.A. 2003. Comparing integrated LiDAR and multi-spectral data with field measurements in hardwood stands. Starkville, MS: Mississippi State University. 158 p. M.S. thesis.

Holmgren, J. 2004. Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. Scandinavian Journal of Forest Research. 19(6): 543-553.

Lefsky, M.A.; Cohen, W.B.; Acker, S.A.; Parker, G.T.; Spies, A.; Harding, D. 1999. LIDAR remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. Remote Sensing of Environment. 70(3): 339-361.

Magnussen, S.; Boudewyn, P. 1998. Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. Canadian Journal of Forest Research. 8: 1016-1031.

McCombs, J.W.; Roberts, S.D.; Evans, D.L. 2003. Influence of fusing LIDAR and multispectral imagery on remotely sensed estimates of stand density and mean tree height in a managed loblolly pine plantation. Forest Science. 49(3): 457-466.

Means, J.E.; Acker, S.A.; Fitt, J.B.; Renslow, M.; Emerson, L.; Hendrix, C.J. 2000. Predicting forest stand characteristics with airborne scanning lidar. Photogrammetric Engineering & Remote Sensing. 66(11): 1367-1371.

Naesset, E. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sensing of Environment. 80(1): 88-99.

Naesset, E. 2004. Accuracy of forest inventory using airborne laser scanning: evaluating the first Nordic full-scale operational project. Scandinavian Journal of Forest Research. 19(6): 554-557.

Nelson, R.; Valenti, M.A.; Short, A.; Keller, C. 2003. A multiple resource inventory of Delaware using airborne laser data. BioScience. 53(10): 981-992.

Parker, R.C. 2005. Computer automation of a LIDAR double-sample forest inventory. Bull. FO275. Starkville, MS: Mississippi State University, Forest and Wildlife Research Center. 22 p.

Parker, R.C.; Evans, D.L. 2004. An application of LIDAR in a double-sample forest inventory. Western Journal of Applied Forestry. 19(2): 95-101.

Parker, R.C.; Evans, D.L. 2006. Industrial application of a LIDAR double-sample forest inventory. Bull. FO317, Starkville, MS: Mississippi State University, Forest and Wildlife Research Center. 17 p.

Parker, R.C.; Glass, P.A. 2004. High versus low-density LIDAR in a double-sample forest inventory. Southern Journal of Applied Forestry. 28(4): 205-210.

Parker, R.C.; Mitchel, A.L. 2005. Smoothed versus unsmoothed LIDAR in a double-sample forest inventory. Southern Journal of Applied Forestry. 29(1): 40-47.

Persson, A.; Holmgren, J.; Soderman, U. 2002. Detecting and measuring individual trees using an airborne laser scanner. Photogrammetric Engineering & Remote Sensing. 68(9): 925-932.

Popescu, S.; Wynne, W.H.; Nelson, R.F. 2002. Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window size. Computers and Electronics in Agriculture. 37: 71-95.

Williams, E.C. 2006. Minimum data input for a LiDAR based timber inventory. Starkville, MS: Mississippi State University. 38 p. M.S. thesis.

Cartographic Standards To Improve Maps Produced by the Forest Inventory and Analysis Program

Charles H. (Hobie) Perry[1] and Mark D. Nelson[2]

Abstract.—The Forest Service, U.S. Department of Agriculture's Forest Inventory and Analysis (FIA) program is incorporating an increasing number of cartographic products in reports, publications, and presentations. To create greater quality and consistency within the national FIA program, a Geospatial Standards team developed cartographic design standards for FIA map products. We present an overview of FIA's proposed cartographic requirements and guidelines, descriptions of specific map elements, and examples of map templates for State reports. The full set of FIA cartographic design standards are expected to be published in a Research Station publication and on the Internet.

Introduction

The Forest Service, U.S. Department of Agriculture's Forest Inventory and Analysis (FIA) program produces data, information, and knowledge about the extent, condition, status, and trends of the Nation's forest resources across all land ownership categories (Smith 2002). Traditionally, FIA data and information have been provided as tabular summaries for estimation units such as States and counties (e.g., Schmidt 1997). Possessing greater understanding of Geographic Information Systems (GISs), current FIA customers require enhanced geospatial analysis of forest resources. FIA scientists and analysts are meeting this need by incorporating an increasing number of maps into their reports, publications, and presentations (e.g., McWilliams *et al.* 2005; Woodall *et al.* 2005).

The interagency Federal Geographic Data Committee (FGDC) was established by the Office of Management and Budget in 1990 for the express purpose of "promot[ing] the coordinated development, use, sharing, and dissemination of geospatial data on a national basis" (FGDC 2006a). With the perspective that consistent standards make it easier to develop and use spatial data, FGDC actively solicits participation from governmental and private entities in standards development (FGDC 2006b). In addition to developing standards for maintaining geospatial data, many diverse Federal agencies have developed standards to meet their cartographic needs (i.e., the visual display of geospatial data within map products). For example, the U.S. Geological Survey maintains National Mapping Program Standards for printed maps as well as their underlying digital data (USGS 2006). Likewise, the Oregon office of the U.S. Bureau of Land Management and Natural Resources Canada also provide standards for cartographic products (Natural Resources Canada 2006a, 2006b; BLM 2006). Most of the standards provided by these organizations are very detailed; some go so far as to specify the exact symbology (e.g., shape, color, line width) required to represent every feature on a map.

FIA recognizes a need to ensure quality and consistency in their cartographic products (e.g., map standards for State reports). Unfortunately, none of the traditional Federal agencies mentioned above have created standards that meet these needs. In response, FIA's Remote Sensing Band established a cross-band task team known as the Geospatial Standards Team (GeoTeam) to recommend (1) cartographic standards associated with FIA national and regional map products and (2) a list of relevant GIS-base layers for the FIA national program. Our objective in this article is to introduce a working draft of the GeoTeam's cartographic design standards and encourage their use by FIA analysts and partners.

[1] Research Soil Scientist, U.S. Department of Agriculture (USDA) Forest Service, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108. E-mail: charleshperry@fs.fed.us.

[2] Research Forester, USDA Forest Service, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108.

Methods

Our process for standards development began with a casual review of existing cartographic products released by a range of FIA analysts. Some of these products were included in State reports while others were created for specific national tasks. It was clear from this review that many analysts were capable users of GIS, but they lacked training in traditional cartography and principles of design. The agencywide availability of ArcGIS software made it easier to create maps, but no standards existed to facilitate the creation of high-quality maps.

A number of excellent cartographic resources are available as texts and Web sites (Brewer 2005, Geographer's Craft Project 2006, Krygier 2005, Slocum 2005). A review of these resources reinforces the notion that effective maps have several elements in common (e.g., a depiction of scale, a legend, a list of data sources, and a title). These elements should be addressed in any cartographic standards document.

A comparison of extant cartographic products with traditional cartographic theory suggests the rough structure of the future standards. The cartographic standards for the FIA program would need to be different than those currently promulgated by other Federal agencies. The national FIA program is implemented as a confederation of regional programs, and analysts want to express creativity in their cartographic products. Furthermore, most FIA maps are designed to augment State reports and other publications, not to be mass-produced as quadrangle sheets or visitor maps. A set of highly prescriptive standards like those released by the U.S. Geological Survey (2006) would lead to consistency and quality, but they would not be practical or well received. Fortunately, we discovered that King County, WA, had encountered a similar dilemma, and its local task team developed a set of cartographic standards (Cartographic Standards Workgroup 2006) that would serve as a model for our efforts.

Standards

The standards we developed consist of both requirements and guidelines. Together, these requirements and guidelines will facilitate high-quality maps that "brand" FIA cartographic products in much the same way that everyone recognizes a U.S. Geological Survey quadrangle map or a national forest map.

The standards pertain to State maps in FIA reports, national maps, and other FIA map products. Requirements are prescribed for all maps produced for public distribution. Generally, these requirements are written for those elements essential to high-quality cartography or are required by the national FIA program (e.g., plot security). Guidelines are provided for some map elements to assist analysts with map making; guidelines also address design features where creativity could be used to create unique maps that still fall within our hopes for an FIA "brand."

Figures 1 through 4 present State map templates (using Wisconsin as an example) with text boxes that describe the current proposed requirements (10 in all) and a sample of the guidelines (12 of 24), linked to specific elements on the example maps. As can be seen in the figures, FIA's cartographic design standards pertain not only to traditional county-based choropleth maps, but also to dot maps and pixel-based land cover maps. Note that the standards do not address every detail of every map element. Rather, the standards provide for more consistency, efficiency, and quality while allowing the map producer some flexibility and creativity.

Summary

Effective cartographic communication depends upon following some basic principles. By considering the balance, proportion, and emphasis of different map elements, we developed cartographic standards for the FIA program; our intent is to facilitate the creation of good maps. The implementation of consistent cartographic standards provides a common 'look-and-feel' for map products and thus creates a visual brand for the FIA program. A more complete description of each FIA cartographic requirement and guideline will be documented in a forthcoming publication.

Figure 1.—*A subset of the proposed cartographic standards (two requirements and three guidelines) using a reference map for illustration.*

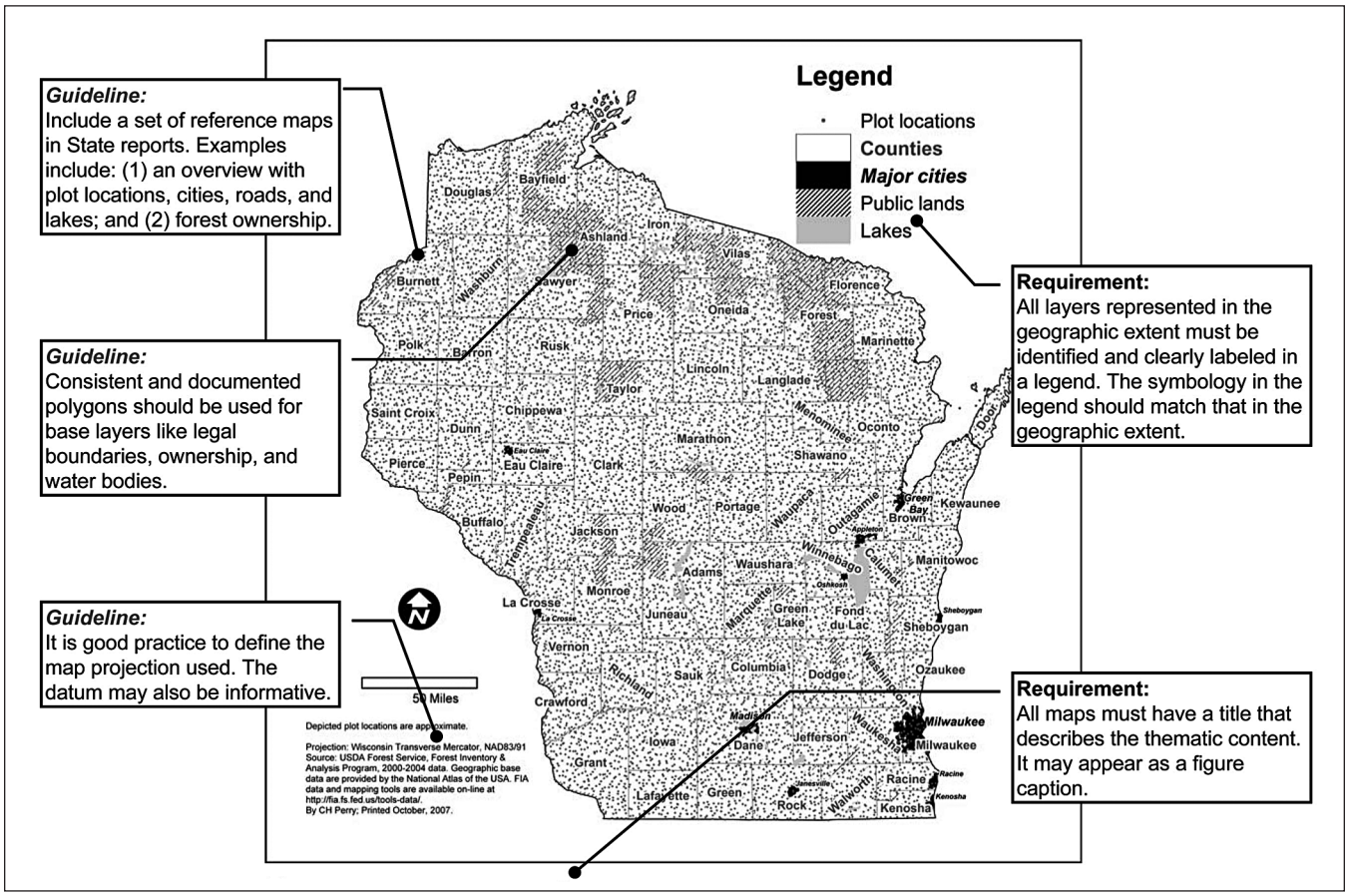


Figure 2.—*A subset of the proposed cartographic standards (three requirements and three guidelines) using a county-based map for illustration.*

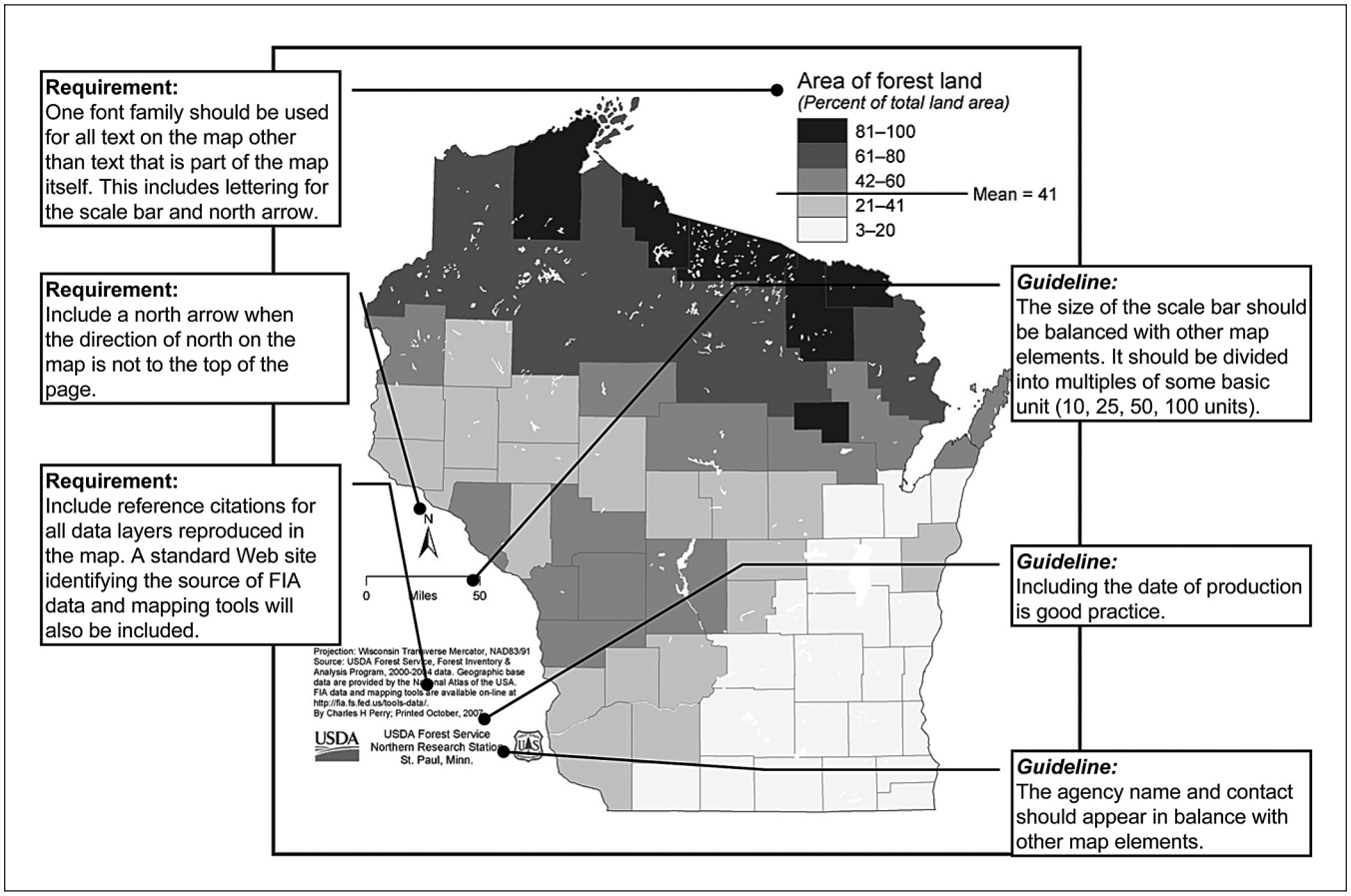


Figure 3.—*A subset of the proposed cartographic standards (two requirements and three guidelines) using a pixel-based map for illustration.*

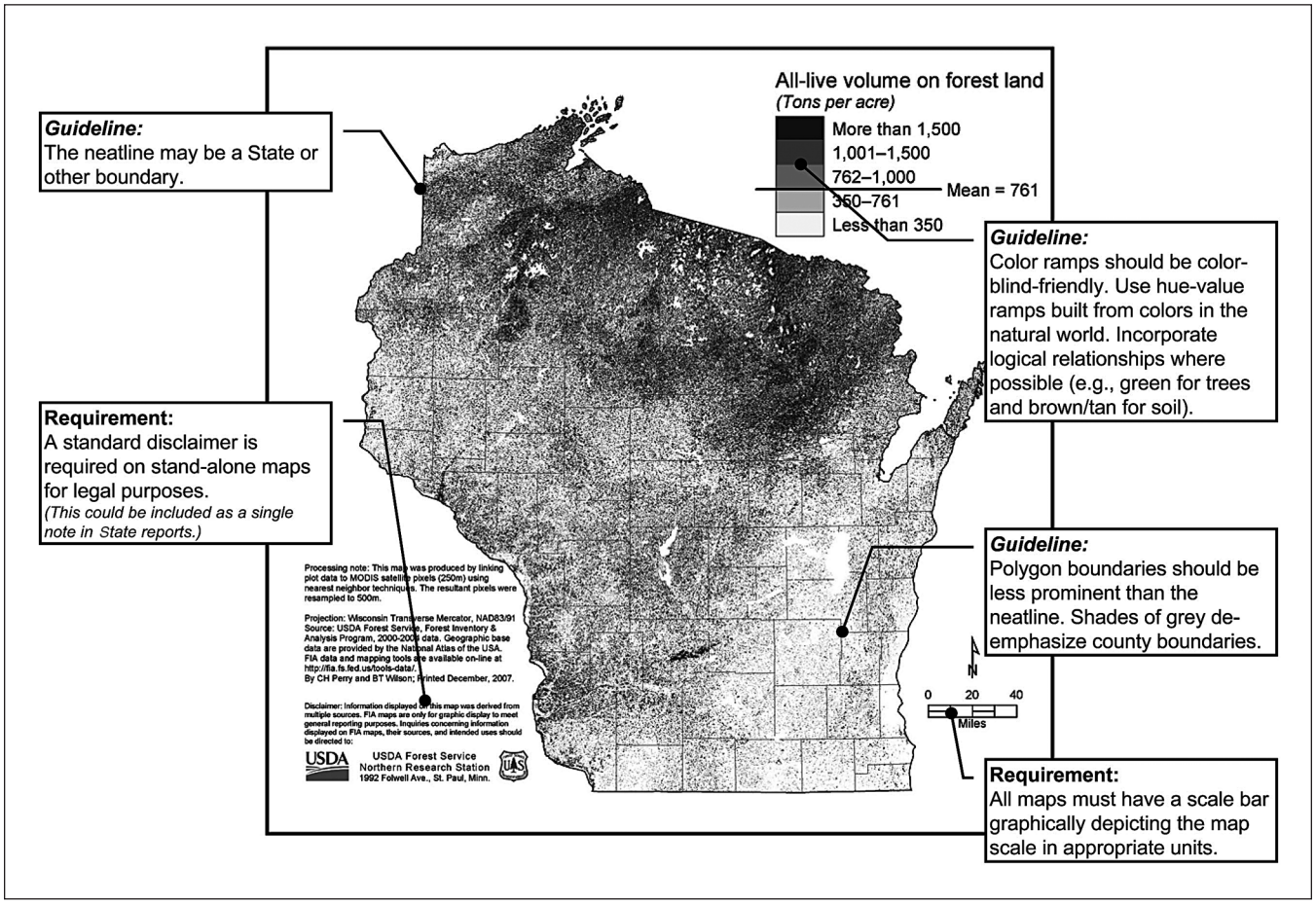
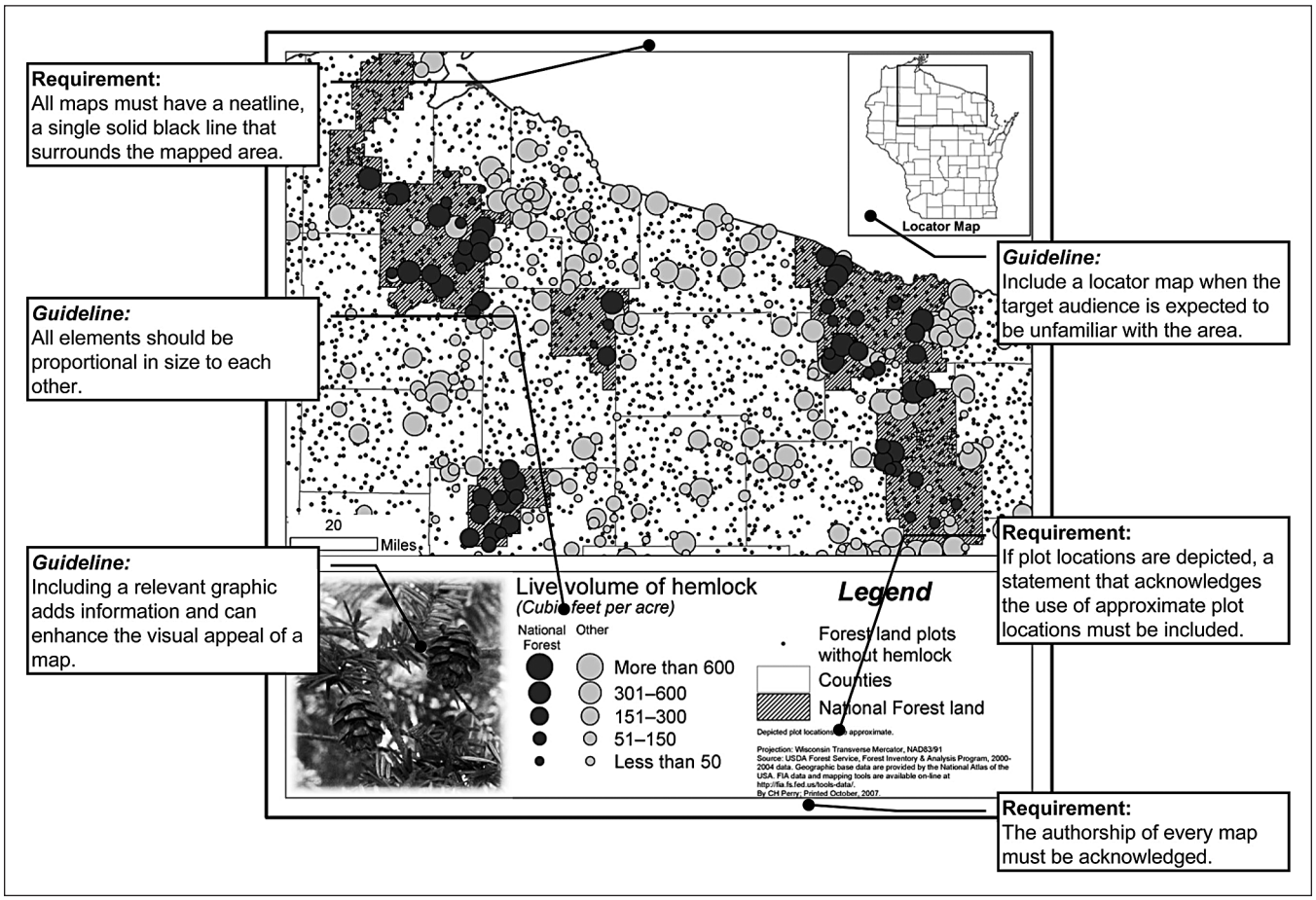


Figure 4.—*A subset of the proposed cartographic standards (three requirements and three guidelines) using a dot map for illustration.*



Acknowledgments

We thank other members of the Geospatial Standards Team for their contributions to the content of the cartographic standards: John Chase, Jamie Cochran, Dennis Collins, Dale Gormanson, Liz LaPoint, Andy Lister, Sam Lambert, Rachel Riemann, and Rich Warnick.

Literature Cited

Brewer, C.A. 2005. Designing better maps: a guide for GIS users. Redlands, CA: ESRI Press. 203 p.

Cartographic Standards Workgroup. 2006. King County GIS cartographic standards. http://www.metrokc.gov/gis/kb/Content/CartoStandards.htm. (1 December).

Federal Geographic Data Committee (FGDC). 2006a. Federal Geographic Data Committee. http://www.fgdc.gov/. (1 December).

FGDC. 2006b. Standards, Federal Geographic Data Committee. http://www.fgdc.gov/standards/. (1 December).

Geographer's Craft Project. 2006. Geographer's craft. http://www.colorado.edu/geography/gcraft/contents.html. (1 December).

Krygier, J. 2005. Making maps: a visual guide to map design for GIS. New York: Guilford Press. 303 p.

McWilliams, W.H.; Butler, B.J.; Caldwell, L.E. *et al*. 2005. The forests of Maine, 2003. Resour. Bull. NE-164. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station. 188 p.

Natural Resources Canada. 2006a. Cartographic standards and specifications: 1/50 000. http://maps.nrcan.gc.ca/cartospecs/MainIndexE50.htm. (1 December).

Natural Resources Canada. 2006b. Cartographic standards and specifications: 1/250 000. http://maps.nrcan.gc.ca/cartospecs/MainIndexE250.htm. (1 December).

Schmidt, T.L. 1997. Wisconsin forest statistics, 1996. Resour. Bull. NC-183. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station. 150 p.

Slocum, T.A. 2005. Thematic cartography and geographic visualization. 2nd ed. Upper Saddle River, NJ: Pearson/Prentice Hall. 518 p.

Smith, W.B. 2002. Forest inventory and analysis: a national inventory and monitoring program. Environmental Pollution. 116: S233-S242.

U.S. Bureau of Land Management (BLM). 2006. Geographic sciences: mapping standards. http://web.or.blm.gov/or957/mapping/cartography/mapstandards.asp. (1 December).

U.S. Geological Survey (USGS). 2006. National Mapping Program standards. http://nationalmap.gov/gio/standards/index.html. (1 December).

Woodall, C.W.; Johnson, D.; Gallion, J. *et al.* 2005. Indiana's forests 1999-2003 (Part A). Resour. Bull. NC-253A. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station. 94 p.

Automatic Crown Cover Mapping To Improve Forest Inventory

Claude Vidal[1], Jean-Guy Boureau[2], Nicolas Robert[2], Nicolas Py[2], Josiane Zerubia[3], Xavier Descombes[3], Guillaume Perrin[3, 4]

Abstract.—To automatically analyze near infrared aerial photographs, the French National Institute for Research in Computer Science and Control developed together with the French National Forest Inventory (NFI) a method for automatic crown cover mapping. This method uses a Reverse Jump Monte Carlo Markov Chain algorithm to locate the crowns and describe those using ellipses or ellipsoids. The system works well when the crowns can easily be distinguished (e.g., in old, high forests or in poplar plantations). The outputs of such a system can be used as post-stratification criteria to enhance the precision of NFI results. It will also allow computing results at larger scales and provide new data for a better understanding of forest ecosystems.

Introduction

Since its creation in 1958, the French National Forest Inventory (NFI) has used aerial photographs to collect information during the first phase of an inventory. Using color infrared (CIR) pictures since the 1980s, photo interpreters determine land cover and land use as well as forest structure and major species on sample plots. The information is used for the statistical evaluation of the forest and other wooded land areas and to decide whether the plot should be assessed in the field. Until now, since it was a manmade operation, only descriptive information was entered. Because it was too time-consuming, the count and cartography of trees was out of the scope.

In 2003, the French National Institute for Research in Computer Science and Control (INRIA) started a research program to build an automatic system for tree crown extraction using aerial photographs in close cooperation with the French NFI. Such an automatic analysis of images is interesting to enhance forest inventory capacities. The French NFI conducted the first steps to validate the method and give an assessment of the result's accuracy. According to these preliminary results, some uses and enhancements are proposed.

Development of a Method to Map Tree Crowns

New Needs for Data on Forest Ecosystems

The French NFI changed its forest assessment method in 2004. Observations are now made annually of a systematic sample covering the whole country (Vidal *et al.* 2005). This new sampling design is adapted to produce data concerning a larger area and can include the entire country. The former sampling design was applied at the scale of the *department* (1/90th of the whole country), so the new design is no longer able to provide data at the local level without specific measurement. Since it was unable to measure systematically more plots without considerable expense, the French NFI investigated alternative solutions. Automatic analysis of documents required for usual work, such as aerial photographs, is one way to provide the required additional information for wood volume estimation at a marginal cost.

The French NFI was asked for further information to gain a better understanding of the forest ecosystems and their dynamic. The evaluation of the number of crowns, their size, and the height of the trees on large areas are interesting material for such an analysis.

[1] Inventaire Forestier National, Château des Barres, 45290 Nogent-sur-Vernisson, France. E-mail: claude.vidal@ifn.fr.

[2] Inventaire Forestier National, Château des Barres, 45290 Nogent-sur-Vernisson, France.

[3] Institut National de Recherche en Informatique et en Automatique. 2004 route des Lucioles BP93 - 06902 Sophia Antipolis Cedex, France.

[4] During his Ph.D. thesis, the work of Guillaume Perrin was supported partly by the Mathematics Applied to Systems Laboratory from École Centrale Paris and partly by the French National Institute for Research in Computer Science and Control. Thanks to the Action de Recherche Concertée, "Mode de Vie" tests were conducted by the French National Forest Inventory to validate the method and give an assessment of the result's accuracy. According to the preliminary results, some uses and enhancements are proposed.

Result of Close Cooperation Between Foresters and Computer Science Specialists

In 1998, INRIA, National Center for Scientific Research (CNRS), and the Nice-Sophia Antipolis University launched the Ariana project to provide image processing tools for Earth observation and cartography using remotely sensed data. Under this project, several probabilistic models combined with stochastic algorithms were developed for an automated analysis of airborne photographs or satellite images. The major objective was to analyze the structure of the pictures and to extract objects. Specific work has been done in cooperation with the French NFI to analyze the crown cover. The French NFI provided pictures and knowledge about the forest photo interpretation. The Ariana project team proposed the method to create an algorithm. A first program was set up to test the quality of the method.

Overview of the Method

The principle to map tree crowns using aerial infrared photographs is to fit a set of objects (circles, ellipses, etc.) on the original picture. The proposal is created using a random process (van Lieshout 2000). The fitting is obtained using a reversible iterative algorithm: Reverse Jump Monte Carlo Markov Chain (RJMCMC) (Green 1995).

To evaluate the quality of the proposal, two different criteria based on photo interpretation are measured (Perrin *et al*. 2005):

1. An external criterion resulting from the comparison between the original picture and the obtained set of objects after rasterization;
2. An internal criterion corresponding to prior information on the size, shape, and spatial organization of the crowns modeled by local interactions between objects.

To create the predicted proposal and evaluate both aspects, a description of crowns (color, shape, etc.) is required. To describe the crowns in a simple mathematical representation, shapes are used.

Model Varieties

The analysis of the photo-interpretation process led to the use of different shapes, depending on the general characteristics of the crown in the analyzed image. Two- and three-dimensional (3D) shapes are used. In the two-dimensional (2D) case, ellipses are used (Perrin *et al*. 2005, Perrin *et al.* 2006a). Each shape is described by the following random variables named marks (van Lieshout 2000), defining the geometry of the object (fig. 1):

1. Position of its center (*u, v*).
2. Length of the two orthogonal axes (*a, b*).
3. Orientation (*a*).

The 2D models are efficient for dense cover, when the shadows of the trees are not entirely visible and crowns are very close. These models roughly describe the size and shape of the crown's visible part.

Four subfamilies of 2D models are used (fig. 2), depending on the quality of the picture (saturated colors or visible differences

Figure 1.—*Parameters of the ellipses used to model crowns.*

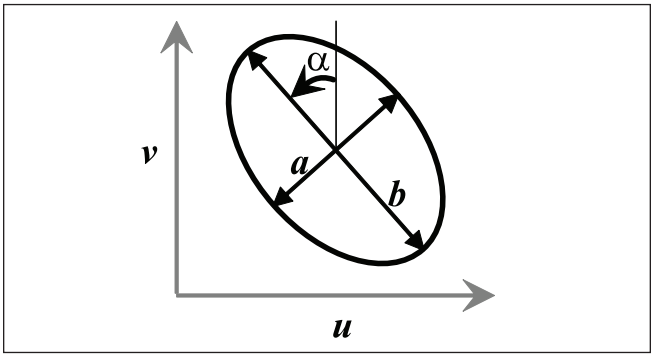
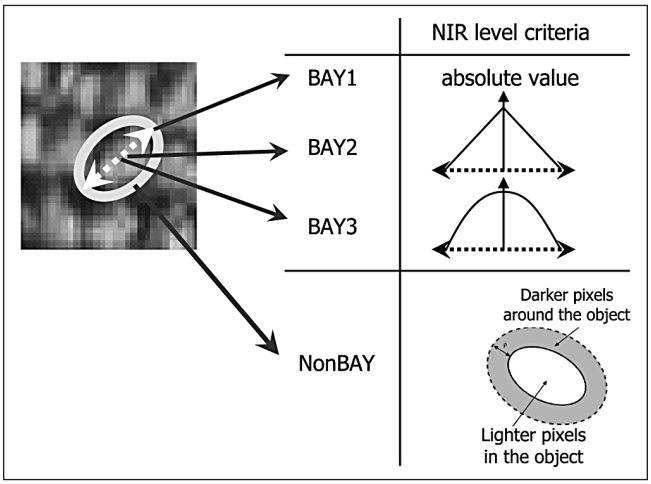


Figure 2.—*Four two-dimensional models to describe the crown appearance on aerial photographs.*



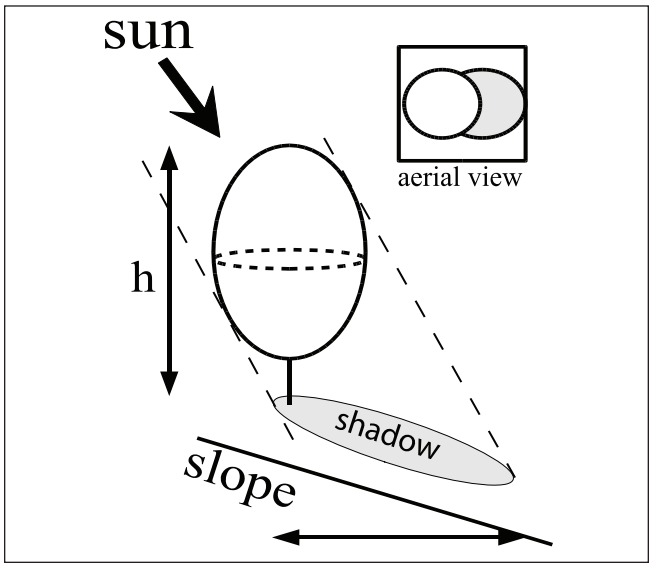
between the border and the center of the crown), the usual shape of the trees (conic top such as sessile oak, or triangle top such as spruce), and the surrounding shadow.

The first model, called BAY1 (for Bayesian model 1), takes into account the absolute reflectance value in the near infrared (NIR). It is assumed that crown's shapes are close to ellipses and are strongly contrasted in the NIR. The second and third models consider a variation of the radiometry inside the ellipse, with higher reflectance in the center of the crown. This variation can be either linear (BAY2) or conic (BAY3). The last shape is combined with a non-Bayesian model (NonBAY), and analyzes the contrast between the outer and the inner parts of the predicted crown and tries to identify a shadow area around this predicted crown.

The 3D shape is an ellipsoid (fig. 3) described by the same parameters as the 2D model, plus height (*h*). This shape is efficient for sparse trees, when the shadows are fully visible, or to determine the height of stands in the edge, when sunshine conditions are sufficient. Such a model considers the high reflectance of the vegetation in NIR, the orientation of the sun, and the relief and shadows around trees.

The type of objects is manually selected, depending on the differentiation level between crowns and shadows. This choice has a strong influence on the results.

Figure 3.—*Ellipsoid shape used in 3D models.*



RJMCMC Algorithm

To fit a random proposal of tree location and characteristics to the original image, an RJMCMC algorithm is used (Green 1995). An image corresponding to the random proposal is predicted. An energy is attached to the system "predicted picture—original picture" to measure the quality of the prediction. The better the proposal, the lower the system's energy. The goal of the iterative process is to find the most accurate prediction (equivalent to the lowest energy).

A random initialization using a rough number of trees creates the initial proposal. The energy of this system is calculated. At each step of the interactive process, a new proposal is built, starting from the previous one, and changes are proposed. These changes consist of the creation or the suppression of the tree crowns, moves of the crown centers, or modification in the size of crowns. Then, the energy of this new system is evaluated and compared to the previous one. If the energy is lower, then the new proposal is taken to the next step. If the energy increases, the new proposal is either rejected, and the next step starts with the former proposal, or accepted with a given probability depending on the model and a temperature parameter. A proper acceptation ratio (Green 1995) ensures the convergence of the process to a configuration that minimizes the energy.

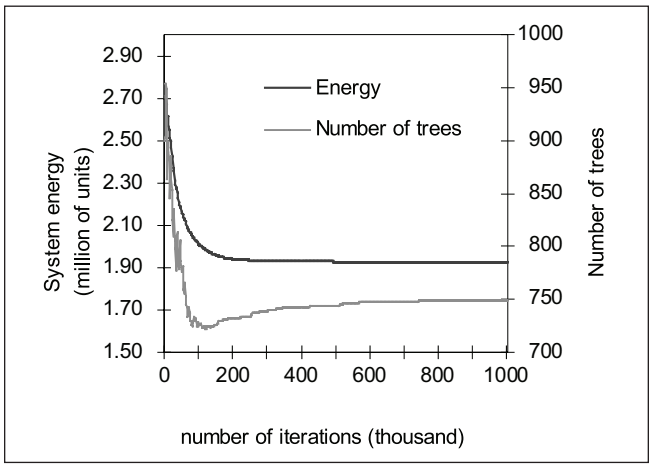
Changes between iterations are randomly proposed, and both the number and the extent of changes are regulated through a term called temperature: the lower the temperature, the fewer the amount of changes. The algorithm starts with high temperatures so as to consider the higher number of possibilities. The temperature is progressively turned low so that few changes are proposed, converging to an optimal solution. The algorithm ensures the convergence of the process to a configuration that minimizes the energy (see example in figure 4).

Assessment of the Quality of the Prediction

Test Sites

Three sites, representing an important part of French forest, have been selected. These sites are examples of hardwood or coniferous mature high forest.

Figure 4.—*Variation of the energy of the system and the number of detected trees.*



1. Site 1: "Indre-et-Loire." This oak, regular high forest in the Loches State Forest is located in a flat zone in the center of France. According to the zones, this managed forest presents various intensities of thinning, from full cover to very clear stands. Corresponding CIR photography was recorded in June 1995 (scale of 1/20000, f = 152 mm).
2. Site 2: "Moselle." Located in the Vosges massif in the northeast of France, this mature, regular high forest of fir is also a managed forest with variable densities of cover from one stand to another and the presence of some clear cuts. As in the test site above, it is easily possible to separate tree crowns. CIR photography was recorded in August 1989 (scale of 1/17 000, f = 213 mm).
3. Site 3: "Loir-et-Cher." Composed of maritime and Scots pines, this forest is located in the "Grande Sologne" in the center of France. This picture is composed of two parts. In the first part, the structure of afforestation is still quite visible. The shape and space distribution of trees are regular, dense, and homogeneous. In the second part, which is less dense, the tree crowns have more heterogeneous shapes. CIR photography was recorded in August 1993 (scale 1/20 000, f = 214 mm).

Collection of Reference Data

The evaluation of the quality of the model requires precise identification of the trees on the image. These data were obtained by simple photo interpretation. The identification of trees was realized with high density (800 dpi) but not an orthorectified digital image. A stereoscopic analysis was carried out in parallel. Each tree crown was identified by a point placed as near as possible to its top (fig. 5).

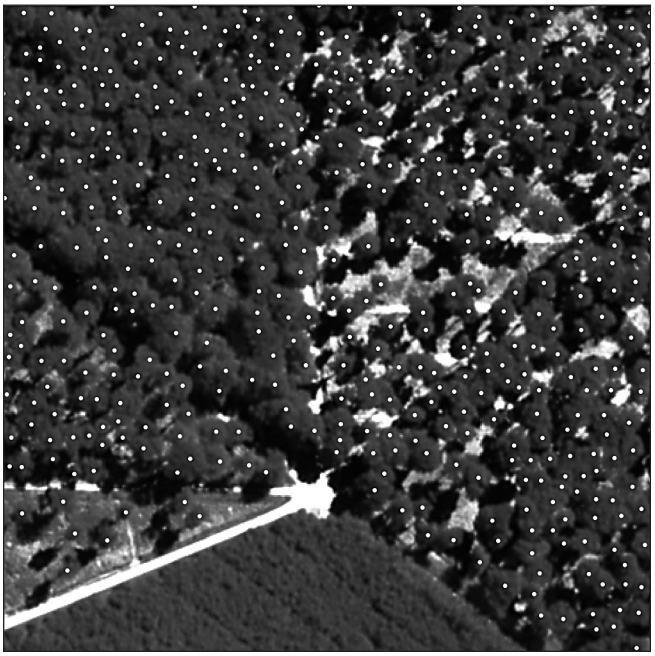
This method offers numerous advantages: orthorectification is not necessary, the information is easily and quickly obtained, and its cost is very low; however, only the mature trees are easily identifiable. When the canopy cover is too dense, group of trees tend to form only one summit but with a specific form and dimension different from insulated trees. Lastly, the partly dominated trees are hardly identifiable.

These limitations, compared to those in field observation, do not constitute a problem insofar as the processing is carried out on the same image. The aim of the model is to identify trees on an image and not to represent the real situation.

Test Protocol

Model Selection. The model selected is the classic BAY 1 (cf. *Model Varieties*), because of the contrast and the image structure. Pixels on the image are split: vegetation or no vegeta-

Figure 5.—*Example of photo-interpretation results on an image presenting a high diversity of densities (in Indre-et-Loire, France).*



Note: A total of 515 crowns were recognized on this part of the image. The lower part was not photo-interpreted because of the too high density of the trees.

tion. The vegetation class corresponds to highest reflectance in NIR. Only the absolute value of this reflectance is considered. These classes are described by their average and standard deviation, which are used by the model to distinguish crowns from shadows and nonstocked areas.

Parameter Definition. The approximate minimum and maximum sizes of the tree crowns must be specified. These values (number of pixels) have been determined by measuring some trees on the digital images. The values obtained for each site are as follows:

1. Indre-et-Loire, oak forest: from 6 to 16 pixels = 3.6 to 9.6 m.
2. Moselle, fir forest: from 6 to 20 pixels = 3 to 10 m.
3. Loir-et-Cher, pine forest: from 4 to 10 pixels = 2.4 to 6 m.

In the realized test series, the number of iterations was set between 1 and 30 million so as to evaluate the number of required iterations. After 1 million of iterations, the system is approximately stabilized. A very good stabilization is reached between 3 and 5 million. Finally, around 10 million iterations, the temperature reached by the system is very low and no change is noticeable. On the site of Indre-et-Loire, for example (fig. 6), in an oak high forest, the number of trees detected after 10 million or after 30 million iterations is practically identical (749 and 753, respectively).

Validation of the Results

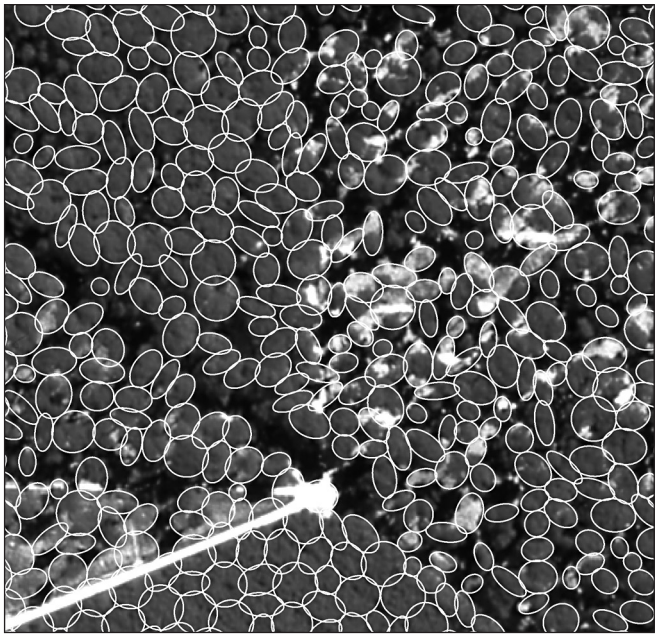
Processing Time. The noted processing times vary with the number of iterations, the image size, and the number of objects. It takes generally 1 to 4 hours to process 5 to 10 ha (12.4 to 24.7 acres) and with an iteration number varying from 10 to 30 million.

Model Stability. To test the model stability, the application was launched five times with the same parameters on the whole fir plantation in the Moselle site. Data have been recorded and, therefore, it is possible to compare the results in terms of number of trees, ellipse coordinates (*x, y*), parameters and areas of ellipses (axis lengths and orientation angle), and computing time.

To compare the results, five stability indicators have been calculated:

1. Indicator 1: Average and standard deviation of the number of detected trees.
2. Indicator 2: Detected crown cover rate. Stability of the image percentage detected as tree crown.

Figure 6.—*Example of results obtained using the BAY 1 model (piece of the Indre-et-Loire image).Note: In the lower part of the image, the prediction of many large crowns covering the whole sector is due to the BAY 1 model that does not use shadows and cannot work in the very high-density regular cover of the young regular high forest. In this part of the image, it is also almost impossible to manually map the crowns.*



3. Indicator 3: Stability of the ellipse parameter distribution. Comparison of averages, variances, and standard deviation for the small axis, the large axis, the orientation, and the area.
4. Indicator 4: Stability of the barycenter for ellipse centers and Euclidean distance of the points to the ellipse centers.
5. Indicator 5: Visualization of images and predicted ellipses for the five simulations. Observation of consistencies and inconsistencies.

The results provided by the model for indicator 1 show a rather good stability: average of 822 trees, standard deviation of 9. For indicator 2, stability is obvious (61% ± 0.1%). The results obtained for indicators 3, 4, and 5 show the same overall stability with variation coefficients generally quite lower than 2 percent.

Comparison Between Photo Interpretation and Automatically Obtained Results. The results obtained are gathered in table 1. If the number of detected trees on the Moselle site is very close to the one seen by photo interpretation, the results will seem

Table 1.—*Comparison between the number of detected trees and the number of trees counted by photo interpretation.*

Site	Area (ha)	Number of detected trees	Number of photo-interpreted trees
Indre-et-Loire	9.0	672	940
Moselle	4.3	667	698
Loir-et-Cher	6.5	1,264	1,962

worse on the two other sites. An unsuited parameter setting, in particular maximal and minimal sizes of the tree crowns, could at least partly explain this wrong detection.

These general results mask rather significant disparities on each site according to the local context. We note either an over- or underdetection according to the species (hardwood or coniferous), the density of the trees, and the nature—chlorophyllian or not—of the ground.

1. Underdetection is often observed in continuous cover but without sufficient shadow to separate the crowns even if each tree could be distinguished in visual analysis. Underdetection is frequent also with small crowns (see fig. 7).
2. On the other hand, an overdetection can occur, for example, with NIR reflective ground spots surrounded by shadow. The model considers them as trees.

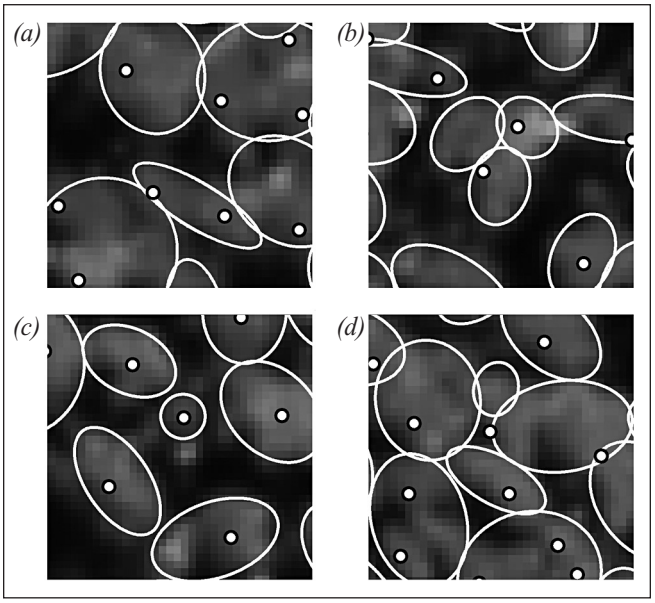
The use of Normalized Difference Vegetation Index (NDVI) instead of the infrared (IR) channel only and the improvement of the thresholding of images would improve these first results.

Possible Use in NFI

Context

Today, the results of very high resolution (decimeter) digital image processing provide information not only at the stand scale but also at the tree scale. Information includes number of stems, local density, dimension and form of the crowns, type of space distribution, and georeferencing of the trees if the image was first orthorectified. Beyond this elementary information, data such as the origin of the forest, its age class, and the present group of species can thus be estimated. This information is impossible to obtain by traditional photo interpretation because

Figure 7.—*Example of results and comparison between photo-interpreted (PI) center of the crowns (points) and predicted ellipses (lines).*
(a) several PI centers in ellipses.
(b) no PI center in ellipses.
(c) good results—one PI center in each ellipse and no empty ellipse nor PI center outside ellipses.
(d) PI centers outside ellipses resulting from problems with size of detected objects.



of costs or the characteristics of visual interpretation. This method is able to synthesize much information but remains not very powerful for a quantitative analysis of the image, especially at the tree scale. A real complementarity now exists between the results of the visual analysis and those of the automatic processing, at least for mature forests.

The requirements in spatialized information, at the tree scale, are increasing for (1) a better knowledge of the ecosystem functioning, (2) the survey of landscape evolution at a very large scale, and (3) a low-cost improvement of the precision of the inventory results. The French NFI is interested in such an automated system, because it allows a more indepth analysis of existing CIR photographs to enhance the quality of the results and the diversity of collected data at a marginal cost.

Improvement of the French NFI's Main Missions

Forest resource assessment is one of the main objectives of the French NFI. The automatic crown mapping program is

expected to improve resource assessment accuracy through the automated analysis of numerous aerial photographs taken all over France. The application could be used to analyze the forestry cover around the photo-interpreted inventory plots (55,000 plots a year, including more than 15,000 forestry plots). It will provide additional information concerning the plots such as number of crowns, density, and repartition. These data can be used as criteria in a post-stratified computation and especially when computing number of stems or tree volume. This will enhance the estimators, even if the confidence level of the application results is low. Using this tool should increase the precision of the results at the national and regional levels. In addition, the increase in precision could possibly provide data at lower scales (e.g., at the *department* level).

Such a tool can also be used for small study areas to produce rough estimators of the crown cover, number of stems, and other variables. Combined with field observations, it might be possible to produce information at the scale of the forest stand.

Within the forests, poplar stands are very specific. Because these stands are usually regularly planted with well-separated crowns, they can be analyzed easily using the automated system (Perrin *et al*. 2005). In these cases, the system will provide good results for the cover rates, the number of trees, and the tree location. It will give a better knowledge of (1) the space heterogeneity of these plantations, (2) their distribution in age groups according to the dimension of the crowns, and (3) their spacing.

Finally the French NFI also carries out an inventory of (1) hedges, (2) tree rows and other trees, and (3) sparse trees in the agricultural landscape. The number of trees and the length of linear elements are statistically estimated. Currently, no systematic map of these trees out of forest has been established. The automatic detection of the crowns enables the analysis at the tree scale within these formations to improve the estimate of tree volumes.

Other Missions

Automatically obtained tree scale information can enrich traditional approaches (visual image analysis and ground measurements) for a better knowledge of the diversity of wooded ecosystems and their evolution. The definition and calculation of many biodiversity indicators, for example, could benefit from these results in digital processing of very high resolution images. Currently, the stand structure, a very important component of forest biodiversity, is rarely available in cartographic form. The detection of individual trees within a stand would enable the characterization of the composition, the spatial organization of mixed forest, and the density of the stems, and other heterogeneities of the forest cover at different scales. We will thus be able to establish main characteristics of vertical and horizontal structures of forest stands and be able to obtain a spatial representation of the forest mosaic. For example, by having access to tree-level information and stand-spatial structure, these first results would allow a qualitative and quantitative step toward understanding links between stand characteristics and plant richness.

The analysis of the evolution of the tree presence in a mosaic landscape answers a similar need. The stands' dynamic (i.e., their natural or anthropic evolution) can thus be approached and mapped by applying the same processing to images registered at different dates. New descriptors calculated on a great number of elementary sampling surfaces would allow the followup of these evolutions at the tree scale.

Future Improvements

This method was developed using a selection of high-quality pictures. Now it must be tested in real conditions to determine the limits of the model. Future developments such as a mapping of the crown shape and a combination with radiometry to recognize main tree species categories are also expected.

Reference data can be obtained, as presented above, by simple photo interpretation. The system can also use orthorectified images so that results are comparable to the Global Positioning System measurement of the tree positions. This method requires a perfect adequacy between the accurate correction of the image and the georeferencing of the trees on the ground. Tests are conducted to specify the interest of this approach depending on the forest types, both for a validation of the model and for its later use.

The first validation tests presented here do not permit an analysis of the model behavior under varied conditions of use. The effects of topography, angle of solar incidence, distance to nadir, spatial resolution, and local contrast of image, in addition to those of spectral bands or type of sensor, also have to be analyzed. It seems, however, that the slope is a secondary problem if the angles of illumination and sighting allow a good observation of the trees.

Similarly, the model has been validated in mature and monospecific forest stands managed in regular high forest, but it has only been tested in mixed forest with a strong vertical heterogeneity (Perrin *et al*. 2006a).

The species identification is based simultaneously on radiometric and crown shape criteria. For the model presented here, post-processing is possible after obtaining the ellipses. This would permit getting information on the real shape (or shape class) of the crowns. This result combined with reflectance values would lead to a better identification of the species (or groups of species).

Conclusions

The major advantages of this method are (1) a limited cost of the inputs (photographs are already available), (2) a short computing time for limited size images, and (3) reproducible results. The results can be used to improve the statistical quality of the French NFI's results, especially number of stems, growing stock in high forests (representing 45 percent of the forest area in France), or standards (80 percent of the growing stock). Better knowledge of the heterogeneity of the top of the crown cover and the height differences can be interesting for landscape and biodiversity purposes.

The French NFI uses segmentation tools development by the French National Geographic Institute to map forests and tree rows. Thanks to a combination of those tools and the crown mapping application, it will be possible to better estimate the standing volume in hedges, using the predicted number of trees, tree height, and crown size. This would help (e.g., in enhancing the wood fuel resource evaluation).

Literature Cited

Green, P. 1995. Reversible Jump Markov Chain Monte Carlo computation and Bayesian model determination. Biometrika. 57: 97-109.

Perrin, G.; Descombes, X.; Zerubia, J. 2005. Point processes in forestry: an application to tree crown detection. Res. Rep. 5544. Sophia Antipolis, France: French National Institute for Research in Computer Science and Control. http://hal.inria.fr/inria-00070463.

Perrin, G.; Descombes, X.; Zerubia, J. 2006a. A non-Bayesian model for tree crown extraction using marked point processes. Res. Rep. 5846. Sophia Antipolis, France: French National Institute for Research in Computer Science and Control. http://hal.inria.fr/inria-00070180.

Perrin, G.; Descombes, X.; Zerubia, J. 2006b. Forest resources assessment using stochastic geometry. Res. Rep. Sophia Antipolis, France: French National Institute for Research in Computer Science and Control. ftp://ftp-sop.inria.fr/ariana/Articles/perrin_emmcvpr.pdf. (11 January 2009).

van Lieshout, M.N.M. 2000. Markov Point Processes and their applications. London, United Kingdom: Imperial College Press.

Vidal, C.; Bélouard, T.; Hervé, J.C.; Robert, N.; Wolsack, J. 2005. A new flexible forest inventory in France. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C.; McWilliams, W.H., eds. Proceedings, seventh annual forest inventory and analysis symposium. Gen. Tech. Rep. WO-77. Washington, DC: U.S. Department of Agriculture, Forest Service: 67-74.

Additional Reading

Project Ariana: http://www.inria.fr/ariana/.

French NFI: http://www.ifn.fr/.

The Finnish Multisource National Forest Inventory: Small-Area Estimation and Map Production

Erkki Tomppo[1]

Abstract.—A driving force motivating development of the multisource national forest inventory (MS-NFI) in connection with the Finnish national forest inventory (NFI) was the desire to obtain forest resource information for smaller areas than is possible using field data only without significantly increasing the cost of the inventory. A basic requirement for the method was that it provide applicable information for forestry decision making; e.g., volume estimates—possibly by subclasses such as tree species—timber assortments, and stand-age classes. In an optimal case, the method had to provide all the same estimates for small areas as the field data-based method provides at the national and subnational level. A nonparametric k-NN estimation method fulfills at least part of these requirements. It is simple to apply in its basic form, and the final estimation method is similar to the method that uses field data only. The input data for the Finnish MS-NFI are NFI field data, satellite images, and digital map data of different types, e.g., basic map data and soil data, as well as a digital elevation model. The first operational results were computed in 1990. The method has been modified continuously and new features added. The k-NN method has several advantages but also limitations; e.g., the field plot data should cover the variation of the field variables in the target area. Application of the k-NN estimation method also presumes the selection of estimation parameters. Particularly, predictions of volumes by tree species may be biased if the area of interest is large and covers several different vegetation zones with different tree species compositions. The biases can be reduced if the set of potential nearest neighbors are restricted or the selection of the neighbors directed. A challenging task related to k-NN method is development of an analytic error estimation method for a target area of an arbitrary size. Recent work with the Forest Inventory and Analysis Program of the Forest Service has shown promise in this task. Progress and problems related to the development and application of the k-NN method are discussed, as well as demonstration of the results.

Introduction

The development of the Finnish multisource national forest inventory (MS-NFI) began in 1989; the first operational results were obtained in 1990 (Tomppo 1990, 1991, 1996, 2006). The driving force behind the development of a multisource method was the need to obtain forest resource information for smaller areas than would be possible with field data in an inexpensive manner. Furthermore, new natural resource satellite images provided new possibilities for increasing the efficiency of the inventories at relatively small additional costs.

A basic requirement of the method was that it should be able to provide information applicable to forestry decisionmaking. Thus, methods that are often used in satellite-image-aided approaches but that produce only maps of forest types or land use classes were not considered satisfactory. Methods were sought that would be able to provide area and volume estimates, possibly by subclasses, such as tree species, timber assortments, and stand-age classes. In the optimal case, the method had to be able to provide all the same estimates for small areas as the field-data-based method provides at national and regional levels. The number of variables measured in the field is usually high, typically ranging from 100 to 400. Estimates for additional variables are calculated from these measured ones.

[1] Professor of Forest Inventory, Finnish Forest Research Institute, Vantaa Research Centre, Unioninkatu 40, FIN-00170 Helsinki Finland. E-mail: erkki.tomppo@metla.fi.

Since the first implementation of the method, it has been modified continuously and new features added (Katila *et al.* 2000; Katila and Tomppo 2001, 2002). The core of the current method is presented in Tomppo and Halme (2004). Any digital land use map or land cover data can be used to improve the accuracy of the predictions (Tomppo 1991, 1996). The list of references of the *k*-Nearest Neighbor (*k*-NN) applications and tests in forest inventories is given in Tomppo (2006). Another nationwide application for Sweden is given in Reese *et al.* (2003).

Application of the *k*-NN estimation method presumes the selection of estimation parameters for each satellite image and for the other data used with the image (Katila and Tomppo 2001). Operational application of the method has also shown that the predictions may be biased if the area of interest is large and covers several vegetation zones with different tree species compositions. Varying imaging conditions within the area of a satellite image can also alter the covariance structure between the field data and the image data. The biases will be reduced if the set of potential nearest neighbors can be restricted to the areas corresponding to the vegetation structure and imaging conditions of the pixel in question. In the first operational applications of the MS-FNFI, a subset of field plots was selected for potential nearest neighbors in the image space for each pixel, usually field plots within a certain geographical horizontal and vertical distance from the pixel in question (Katila and Tomppo 2001).

Tomppo and Halme (2004) presented another method for guiding the selection of field plots that has been in operational use since early 2000. This method uses additional elements in the distance metric vector to guide the selection of nearest neighbors. The elements are variables describing large-area variations in forest characteristics, e.g., mean volumes by tree species, and are map-form predictions of those variables. A relevant, practical variation scale for these variables and their predictions would range between 40 and 60 km. Variation on this scale can be computed from field data only, e.g., from field data acquired in the current or preceding inventory of the area.

The method also uses band transformations in addition to the original image bands because band ratios are assumed to improve the identification of tree species. An optimization method based on a genetic algorithm was developed to find the weight vector, a method that considerably reduces the errors both at the pixel level and in areas of different sizes. The method is called the i*k*-NN method (improved *k*-NN method) in Tomppo and Halme (2004).

One of the open problems related to the *k*-NN method is that until recently, analytical methods for estimating the standard error of any estimate for an area of an arbitrary size have been lacking. However, current progress with a model-based approach has shown promise for error estimation for the *k*-NN method (Kim and Tomppo 2006, McRoberts *et al*. 2007).

Input Data Sets

Field Data

The basic computation unit in image processing is a picture element called a pixel. The pixel size used with Landsat Thematic Mapper (TM) images, for example, is 25 by 25 m. Therefore, it is more convenient to work with volumes per unit area than with volumes of tallied trees. Volumes per ha are estimated for each sample plot by tree species and by timber assortment classes based on the tally tree volumes.

Satellite Images

Images from the Landsat 5 TM or Landsat 7 Enhanced Thematic Mapper Plus (ETM+) sensors have been the most suitable for operational application by virtue of the fairly large coverage area of each image combined with moderate spatial and spectral resolution. These images have been given priority when choosing satellite images to cover an area. If these images are not available, e.g., due to clouds, either Spot 2-4 XS HRV images or IRS-1 C LISS images have been used.

The land area of Finland is 30.4473 million ha, and the total area together with lakes and rivers is 33.8145 million ha. This area was covered by 36 Landsat 5 TM images and 2 Spot 2 XS HRV images in the eighth NFI (NFI8) and its updating in of southern Finland (1990–94), and by 40 Landsat 5 TM or Landsat 7 ETM+ images and 4 IRS-1 C LISS images in the ninth NFI (NFI9) (1996–03).

Areas corresponding to the cloud-free parts of satellite images are used in operational applications. Forests under clouds and in cloud shadows are assumed to be similar on the average to those on the cloud-free part of the same area unit (e.g., municipality). All images are rectified to the national coordinate system using the nearest neighbor resampling method with a pixel size of 25 by 25 m.

Digital Map Data

Digital map data are used to reduce the errors in the estimates. The errors in both the area and total volume estimates can be reduced significantly by the multisource method if the distinguishing forest land from nonforest land can be supported by digital map information in addition to satellite images. The effect of possible map errors on the estimates can be reduced by two alternative statistical methods (Katila *et al.* 2000, Katila and Tomppo 2002). The first is a calibration method using a confusion matrix derived from the land use class distributions on the basis of field plot data and map data, and the second uses stratification of the field plots on the basis of map data. The map information is used to separate forestry land from other land use classes, such as arable land, built-up areas, roads, urban areas, and single houses. In addition, a map is used to stratify the forest land area and corresponding field plots into a mineral soil stratum and a peatland soil stratum (spruce mires, pine mires, open bogs, and fens).

A digital elevation model is used in two ways: for stratification on the basis of elevation data and for correcting the spectral values by reference to the angle between solar illumination and the terrain normal. The latter method is described in detail by Tomppo (1992).

The basic computation unit in the MS-NFI is the municipality. The number of municipalities in the entire country is approximately 500 and their land areas range from around 1,000 ha to some hundreds of thousands of hectares. Digital municipality boundaries are used to delineate the units (Tomppo 1996).

Large-Area Forest Resource Data

The large scale variation of forest variables used in the i*k*-NN method is presented in the form of maps derived from field data using spatial interpolation (Tomppo and Halme 2004). The number of field plots on land in the entire country in NFI9 was 81,249, including 67,264 on forestry land, 62,266 on combined forest land and poorly productive forest land, and 57,457 on forest land alone. All the plots on forest land and poorly productive forest land were used for the large-area maps. The variables were selected in such a way that their values indicate the areas in which the covariance structure between field variables and image variables would be approximately constant. Volumes by tree species (m^3/ha) on forest land and poorly productive forest land were therefore selected as variables.

i*k*-NN Estimation

The Principles of the Method

A nonparametric *k*-NN estimation was used for the MS-NFI calculations during NFI8 and at the beginning of NFI9. The improved method, i*k*-NN, has been applied since 2000 and is described in Tomppo and Halme (2004) and Tomppo (2006). We recall that each field plot has a certain area representativeness, a plot weight, sometimes called a plot expansion factor when forest inventory estimates are calculated from pure field data. This plot weight can be the total land area divided by the number of field plots on land if either systematic or systematic cluster sampling is used (Tomppo 2006). In the MS-NFI, new plot weights (not equal for each plot) are calculated for each plot on an areal unit, e.g., municipality (Tomppo 1996). The weights are calculated for each field plot $i \in F$, where F is the set of field plots on forest land. These plot weights are sums of satellite image pixel weights over the forest land mask pixels.

The pixel weights are in turn calculated by a nonparametric *k*-NN estimation method that uses the distance metric *d*, defined in the feature space of the satellite image data (Tomppo 1991, 1996, 2006). The *k* nearest field plot pixels (in terms of *d*)—i.e., pixels that cover the center of a field plot $i \in F$ —are sought for each pixel *p* under the forest land mask of the cloud-free satellite image area. A maximum distance was usually set in the basic *k*-NN method in both a horizontal and a vertical direction to avoid selecting the nearest plots (spectrally similar plots) from a region in which the response of the image variables to the field variables is not similar to that of the pixel under

consideration. In the improved i*k*-NN method, only vertical maximum distance is applied. The use of horizontal distance is replaced by the use of large-scale variation of forest variables in equation (2) (Tomppo and Halme 2004, Tomppo 2006). All elements of the distance metric are finally weighted by means of optimization based on a genetic algorithm.

Denote the nearest feasible field plots by $i_1(p),...,i_k(p)$. The weight $w_{i,p}$ of field plot *i* on pixel *p* is defined as

$$w_{i,p} = \begin{cases} \dfrac{1}{d^t_{p_i,p}} \Big/ \displaystyle\sum_{j\varepsilon(i_1(p),...,i_k(p))} \dfrac{1}{d^t_{p_j p}} \\ 0 \end{cases} \quad (1)$$

if and only if $i\varepsilon\{i_1(p),...,i_k(p)\}$ *otherwise*

The power *t* is a real number, usually $t \in (0,2]$. The distance metric *d* in i*k*-NN method is

$$d^2_{p_j,p} = \sum_{l=1}^{n_f} \omega_{l,f}{}^2 (f_{l,p_j} - f_{l,p})^2 + \sum_{l=1}^{n_g} \omega_{l,g}{}^2 (g_{l,p_j} - g_{l,p})^2 \quad (2)$$

where:

f_{l,p_j} = the l^{th} normalised image variable $f_{l,p_j} = f^0_{l,p_j} / \cos^r(\alpha)$.

f_{l,p_j} = the original intensity of the spectral band l.

α = the angle between the terrain normal and the solar illumination.

r = the applied power due to non-Lambertian surface.

n_f = the number of spectral features.

$g_{l,p}$ = the large-area prediction for the l^{th} forest variable.

n_g = the number of coarse-scale forest variables.

ω_f and ω_g = the weight vectors for the image features and coarse-scale forest variables, respectively. A pixel size of 1 by 1 km is used in the variables $g_{l,p}$ (Tomppo and Halme 2004, Tomppo 2006).

The values for the elements of the weight vector to be estimated are derived from an optimization using a genetic algorithm, as given below. The first phase of i*k*-NN is to run the optimization algorithm, possibly by strata in the applications, e.g., for the mineral soil stratum and mire and bog stratum separately. The procedure then returns to the basic *k*-NN estimation.

To estimate forest parameters for areal units, the field plot weights for the pixels, $w_{i,p}$, are added for the areal units (e.g., municipalities) in an image analysis process extending over the pixels belonging to each unit. The weight of plot *i* in areal unit *u* is denoted by

$$c_{i,u} = \sum_{p \in u} w_{i,p}. \quad (3)$$

Reduced weight sums $c^r_{i,u}$ are obtained from equation (3) if clouds or their shadows cover part of the areal unit *u*. The real weight sum for plot *i* is estimated by means of the formula

$$c_{i,u} = c^r_{i,u} \frac{\hat{A}_{s,u}}{\hat{A}^r_{s,u}}, \quad (4)$$

where $\hat{A}_{s,u}$ is the estimated area of forest land in unit *u,* and $\hat{A}^r_{s,u}$ the estimated area of forest land in unit *u* not covered by the cloud mask. The areas can be taken from digital maps or estimated by means of field plots. It is thus assumed that the forestry land covered by clouds in areal unit *u* is on average similar to the rest of the forest land in that unit with respect to the forest variables (Tomppo and Halme 2004).

Equations (3) and (4) are calculated separately for the mineral soil stratum and peatland stratum within the forest land, and also for other land use classes such as arable land, built-up land, roads and water bodies if a stratification-based map correction method is used (Katila and Tomppo 2002). Alternatively, a statistical calibration and confusion matrix can be used to reduce the effect of map errors on the estimates (Katila *et al.* 2000).

After the final field plot weights on the areal units have been calculated, ratio estimation is used to obtain the estimates (Cochran 1977). In this sense, the estimation procedure is similar to that using field plot data only. Volume estimates, for example, for areal unit *u* and reference unit *s* are calculated in the following way. Mean volumes are estimated by the formula

$$v = \frac{\sum_{i \in I_s} c_{i,u} v_{i,t}}{\sum_{i \in I_s} c_{i,u}}, \quad (5)$$

where $v_{i,t}$ is the estimated volume per hectare of timber assortment (log product) *t* on plot *i* and I_s the set of field plots belonging to stratum *s*. The corresponding total volumes are obtained by replacing the denominator in equation (5) with 1. The forest

variable estimators for areal unit *u* thus utilise information from outside unit *u*.

Examples of estimates obtained with MS-NFI are given in table 1. These area and volume estimates are based on the NFI8 inventory field data and satellite images from 1992 for the Kainuu Forestry Centre District (Tomppo *et al.* 1998). Totals for the entire forestry center district are given in two ways, one based on the MS-NFI and the other based on the NFI only. The standard errors for the forestry center totals are based on NFI (e.g., Heikkinen 2006). In addition to table 1, the following other tables were given in MS-NFI8 for all the municipalities in Finland: areas of mineral soil and peatland soils on forest land; poorly productive forest land and unproductive land separately; tree species dominance on forest land and poorly productive forest land, separately; areas of age classes on forest land; mean volumes (m^3/ha) by age classes on forest land; areas of development classes on forest land; mean volumes (m^3/ha) by development classes on forest land; mean and total volumes by tree species; timber assortment classes on forest land and on forest land and poorly productive forest land combined; and some relative distributions for the area and volume estimates.

Predictions of certain (optional) forest variables are distributed in the form of a digital map during the procedure; e.g., the land use classes outside forestry land are transferred to mapform predictions directly from the digital map file. Within forest land, the variables are predicted from the weighted averages of the *k* nearest neighbors (see Tomppo 1991, 1996).

A pixel-level prediction $\hat{m}_p$ of variable *M* for pixel *p* is defined as

$$\hat{m}_p = \sum_{i \in F} w_{i,p} m_i, \quad (6)$$

where m_i is the value of the variable *M* on plot *i*.

The mode or median value is used instead of the weighted average for categorial variables, i.e., land use class, site fertility class, stand age, mean diameter of stand, mean height of stand, volumes by tree species (e.g., pine, spruce, birch, other broadleaved trees), and by timber assortment class. The total number of maps is thus more than 20.

Optimizing the Variable Weights

The overall aim of the i*k*-NN method is to minimize the errors attached to predictions based on the MS-NFI, both at the pixel level and particularly at higher areal levels (from several tens of thousands of hectares up to several millions of hectares).

Two modifications of the *k*-NN estimation method were introduced:

1. The use of supplementary ancillary variables in addition to spectral data for selecting neighbors.
2. The use of optimal weights for both the image features and the ancillary information.

A vector consisting of these elements is called a vector of explanatory variable weights and denoted by ω. A optimization method based on a genetic algorithm was developed for using ancillary data and finding the optimal explanatory variable weights.

Table 1.—*Mean and total volume of growing stock on forest land and on poorly productive forest land.*

	Forest land			Poorly productive forest land		
	ha	m^3/ha	1000 m^3	ha	m^3/ha	1000 m^3
Hyrynsalmi	113,370	67.4	7,638	14,435	12.0	173
Kajaani	92,850	68.8	6,386	7,798	11.1	87
Kuhmo	388,155	74.9	29,083	49,706	11.4	567
Paltamo	73,713	83.2	6,133	6,429	12.8	82
Puolanko	191,331	69.7	13,341	30,259	13.2	399
Ristijärvi	69,148	68.7	4,748	6613	11.1	73
Sotkamo	222,928	75.7	16,870	16,846	11.5	194
Suomussalmi	389,616	66.1	25,747	65,185	10.4	678
Vaala	88,493	64.0	5,661	18,862	13.3	251
Vuolijoki	52,272	74.6	3,900	8,344	10.2	85
Total, MS-NFI	1,681,876	71.1	119,507	224,477	11.5	2,589
Total, NFI	1,659,701	70.8	117,000	222,675	12.6	2,800
Standard error of NFI	13,895	1.4	2,494	8,969	0.7	198

Volumes by tree species were selected as the variables, because tests with age class distributions, for example, and earlier experiments had shown that the errors in the predictions for other variables are reduced when those attached to volumes by tree species are minimized (Tomppo *et al.* 1998).

Optimization was conducted solely at the pixel level with the expectation and later checking that larger area errors would decrease once the weights were optimized.

A weighted sum of pixel-level biases and root mean square errors (RMSEs) of the predictions was selected as the objective function. The weights are called fitness function weights and denoted by γ (7). The variables used were (1) total volume, (2) volume of pine, (3) volume of spruce, (4) volume of birch, and (5) volume of other broadleaved tree species. These 10 variables have also been used in operational applications of the method. The fitness (objective) function to be minimized with respect to ω is:

$$f(\omega,\gamma,\hat{\delta},\hat{\bar{e}}) = \sum_{j=1}^{n_e} \gamma_j \hat{\sigma}_j(\omega) + \sum_{j=1}^{n_e} \gamma_{j+n_e} \hat{\bar{e}}_j(\omega), \qquad (7)$$

where:

$\gamma > 0$ = user-defined coefficients for the pixel level standard errors $\hat{\sigma}_j$ and biases $\hat{\bar{e}}_j$ in forest variable j (applied in a genetic algorithm).

ω is the weight vector to be estimated (eq. 2).

W = the feasible set of weight vectors.

The pixel-level biases and errors based on cross-validation,

$$\hat{\sigma}_m = \sqrt{\frac{\sum_{i\in F}(\hat{m}_i - m_i)^2}{n_F}} \quad \text{and the bias} \quad \hat{\bar{e}} = \frac{\sum_{i\in F}(\hat{m}_i - m_i)}{n_F},$$

have been applied. Here m_i is the observed value of the variable to be estimated (e.g., total volume), $\hat{m}_i$ its estimate on plot i, and n_F the number of field plots.

The fitness function weights, bias weights, and RMSE weights were experimentally given values and then fixed. This weighted sum was the criterion in the search for good weight vectors for image features and ancillary information.

The main goal in introducing the i*k*-NN method was to improve the accuracy of the predictions of the volumes by tree species, as well as the estimates in different area units. In practical runs of the genetic algorithm (in estimating the weight vector ω), the pixel-level biases are almost totally removed. An example is taken from Tomppo and Halme (2004) for an area in east Finland with the details provided in the original paper. Examples of pixel-level (field-plot-level) biases in predicting volumes by tree species are given for different methods in table 2: the basic *k*-NN prediction, *k*-NN predictions using large-area variables, and i*k*-NN prediction. Leave-one-out cross-validation and field-data-based volume predictions $\hat{V}_F$ as a reference are employed.

Spruce volume was significantly underestimated with the *k*-NN method. The addition of large-area variables to *k*-NN did not alone reduce the biases, but a reduction was noticeable with i*k*-NN, although all predictions were somewhat lower for birch and other broadleaved tree species than for pine and spruce (columns a/b in table 2).

The predictions are validated at the level of groups of municipalities as follows. The area in question is divided into subareas with forest and other wooded land areas, ranging typically between 150,000 ha and 300,000 ha. An example of the predictions for

Table 2.—*Examples of biases attached to k-NN predictions, k-NN predictions with large-area variables (k-NN, la), and ik-NN predictions at the pixel level (field-plot level) on the mineral soil stratum, leave-one-out cross-validation, using field-data-based volume predictions $\hat{V}_F$ as a reference. 1,953 field plots, k = 5, upper bounds used.*

		Bias	Standard error of bias	Bias (a)	Standard error of bias	Bias (b)	Standard error of bias	Reduction
		k-NN	k-NN	k-NN, la	k-NN, la	ik-NN	ik-NN	(a) / (b)
Volume	*m³/ha*	*m³/ha*	*m³/ha*	*m³/ha*	*m³/ha*	*m³/ha*	*m³/ha*	*%*
Pine	63.750	2.430	1.648	2.230	1.570	– 0.002	1.539	99.925
Spruce	38.883	– 3.167	1.304	– 4.725	1.293	– 0.005	1.260	99.891
Birch	15.903	– 0.961	0.684	– 1.571	0.696	– 0.199	0.701	87.346
O. br. l.	3.874	– 0.382	0.376	– 0.430	0.383	– 0.133	0.389	69.057
Total	122.303	– 2.021	1.827	– 4.432	1.764	– 0.259	1.800	94.152

mean volumes by tree species (m^3/ha) is given for a municipality group with an area of forest and poorly productive forest land of 241,200 hectares in table 3. The table also gives standard errors for the field-data-based predictions. The table enables multisource predictions to be compared with the field data estimates and assessed in terms of the field-data-based standard errors. The i*k*-NN method gave lower deviations from the field-data-based predictions, and thus more accurate predictions. Use of information on large-area variations in forest variables in conjunction with the i*k*-NN method noticeably reduces the problem of distinguishing pine dominant stands from spruce dominant ones, for instance, or of estimating the volumes by tree species.

Conclusions

Since 1990, the Finnish NFI has been using a satellite-image-aided multisource method to obtain results for smaller areas than is possible using field data only. The entire country has been covered more than twice by this method. The method is under continuous refinement. During NFI9, the method was enhanced by using new features: (1) large-area forest variables for directing the selection of nearest neighbors, (2) an optimization method based on a genetic algorithm to weight both large-area forest variables and satellite image variables, and (3) two optional methods to remove the effect of map errors on the estimates. The new i*k*-NN method performs noticeably better than the original *k*-NN method. The use of information on large-area forest variables considerably reduces the problem of distinguishing stands with different tree species, or tree species composition, and reduces the errors entailed in the estimates of volumes by tree species. Any relevant data, such as soil data or vegetation zone data, can be used as ancillary data.

For several reasons, the pixel-level and stand-level errors of the estimates are rather high with current satellite images. The error sources in pixel-level predictions of forest variables have been listed in many papers (e.g., Katila 2004, Tomppo 2006).

Several methods for assessing pixel-level errors exist. Leave-one-out cross-validation has been used in many cases; Kim and Tomppo (2006) applied variogram modelling to the spectral space. The finding of a generally applicable error estimation method for areas larger than a pixel is a challenging task. Because the error in the predictor of a variable depends on the true value of the variable, errors are spatially correlated, and spatial dependences in the image itself make the error structure even more complex. Lappi (2001) presented a different, calibration-type approach to multisource estimation, together with a variogram-based variance estimator, and some other interesting variogram approaches are currently under development (McRoberts *et al.* 2007).

Practical applications of the MS-NFI technique are also currently facing other problems. A serious problem relates to the optical-area images, in particular, the availability of images obtained under cloud-free conditions. The most applicable satellite sensor, Landsat 7 ETM+, has suffered from a scan line corrector failure since 2003. Several correction methods have been introduced, but the quality of the product is not the same as before (USGS 2005). One advantage of the *k*-NN method is that it is applicable to all image material. The precision of the estimates depends on the spectral, spatial, and radiometric resolution of the sensor, however, and some image material may presume the use of other image material as an intermediate step between the field data and the final image data (Tomppo *et al.* 2002). Furthermore, the precision of the estimate will depend

Table 3.—*Estimates of the volume of growing stock (m3/ha) on forest and other wooded land (a) and its standard error (aer) by tree species based on field data and on the k-NN method (b), ik-NN method, (c) and ik-NN method when the resulting large-area weights have been multiplied by 10 (d) for a municipality group with an area of forest land and other wooded land of 241,200 ha. The multisource estimates are compared with the field-data-based estimates.*

Tree species	a	a_{er}	b	b-a	c	c-a	d	d-a
Pine	48.6	2.9	53.8	5.2	47.5	– 1.1	49.7	1.1
Spruce	38.5	2.5	35.6	– 2.9	41.9	3.4	40.3	1.8
Birch	15.7	1.2	15.7	– 0.0	15.8	0.1	15.9	0.2
Other br. 1.	4.3	0.6	3.7	– 0.6	3.6	– 0.7	3.1	– 1.2
Total	107.2	3.3	108.8	1.6	109.0	1.8	108.9	1.7

on how the *k*-NN method is applied, as seen above. Numerous research work has been conducted to analyze the errors and improve the precision of the estimates; the process is ongoing.

Literature Cited

Cochran, W.G. 1977. Sampling techniques. 3rd ed. New York: John Wiley. 428 p.

Heikkinen, J. 2006. Assessment of uncertainty in spatially systematic sampling. In: Kangas, A.; Maltamo, M., eds. Forest inventory, methods and applications. Managing forest ecosystems. Vol. 10. London: Springer Dordrecht: 155-176.

Katila, M. 2004. Controlling the estimation errors in the Finnish multisource National Forest Inventory. The Finnish Forest Research Institute, Research Papers 910. 36 p. Ph.D. dissertation.

Katila, M.; Heikkinen, J.; Tomppo, E. 2000. Calibration of small-area estimates for map errors in multisource forest inventory. Canadian Journal of Forest Research. 30: 1329-1339.

Katila, M.; Tomppo, E. 2001. Selecting estimation parameters for the Finnish multisource National Forest Inventory. Remote Sensing of Environment. 76: 16-32.

Katila, M.; Tomppo, E. 2002. Stratification by ancillary data in multisource forest inventories using k-nearest-neighbor estimation. Canadian Journal of Forest Research. 32(9): 1548-1561.

Kim, H.J.; Tomppo, E. 2006. Model-based prediction error uncertainty estimation for k-nn method. Remote Sensing of Environment. 104: 257-263.

Lappi, J. 2001. Forest inventory of small areas combining the calibration estimator and a spatial model. Canadian Journal of Forest Research. 31: 1551-1560.

McRoberts, R.E.; Tomppo, E.O.; Finley, A.O.; Heikkinen, J. 2007. Estimating areal means and variances of forest attributes using the k-nearest neighbors technique and satellite imagery. Remote Sensing of Environment. 111: 466-480.

Reese, H.; Nilsson, M.; Granqvist Pahlén, T.; Hagner, O.; Joyce, S.; Tingelöf, U.; Egberth, M.; Olsson, H. 2003. Country-wide estimates of forest variables using satellite data and field data from the National Forest Inventory. Ambio. 32: 542-548.

Tomppo, E. 1990. Designing a satellite image-aided national forest survey in Finland. In: The usability of remote sensing for forest inventory and planning. Proceedings, SNS/IUFRO workshop. Umea, Sweden: Swedish University of Agricultural Sciences, Remote Sensing Laboratory, Report 4: 43-47.

Tomppo, E. 1991. Satellite image based National Forest Inventory of Finland. International Archives of Photogrammetry and Remote Sensing. 28(7-1): 419-424.

Tomppo, E. 1992. Satellite image aided forest site fertility estimation for forest income taxation. Acta Forestalia Fennica. 229.

Tomppo, E. 1996. Multi-source National Forest Inventory of Finland. In: Vanclay, R.; Vanclay, J.; Miina, S., eds. New thrusts in forest inventory. Proceedings, IUFRO XX World Congress. Joensuu, Finland: European Forest Institute: 27-41.

Tomppo, E. 2006. The Finnish Multi-Source National Forest Inventory—small area estimation and map production. In: Kangas, A.; Maltamo, M., eds. Forest inventory, methods and applications. Managing forest ecosystems. Vol. 10. London: Springer Dordrecht: 195-224.

Tomppo, E.; Halme, M. 2004. Using coarse scale forest variables as ancillary information and weighting of variables in k-nn estimation: a genetic algorithm approach. Remote Sensing of Environment. 92: 1-20.

Tomppo, E.; Katila, M.; Moilanen, J.; Mäkelä, H.; Peräsaari, J. 1998. Kunnittaiset metsävaratiedot 1990-94. Metsätieteen aikakauskirja. Folia Forestalia. 4B/1998: 619-839. In Finnish.

Tomppo, E.; Korhonen, K.T.; Heikkinen, J.; Yli-Kojola, H. 2001. Multisource inventory of the forests of the Hebei Forestry Bureau, Heilongjiang, China. Silva Fennica. 35: 309-328.

Tomppo, E.; Nilsson, M.; Rosengren, M.; Aalto, P.; Kennedy, P. 2002. Simultaneous use of Landsat-TM and IRS-1C WiFS data in estimating large area tree stem volume and aboveground biomass. Remote Sensing of Environment. 82: 156-171.

Topographic Database of Finland. 1998. The National Land Survey of Finland. http://www.nls.fi/kartta/maps/topodb.html. (1 January 2003).

U.S. Geological Survey (USGS). 2005. Landsat Project. http://landsat7.usgs.gov/slc_enhancements/.

Prefield Methods: Streamlining Forest or Nonforest Determinations To Increase Inventory Efficiency

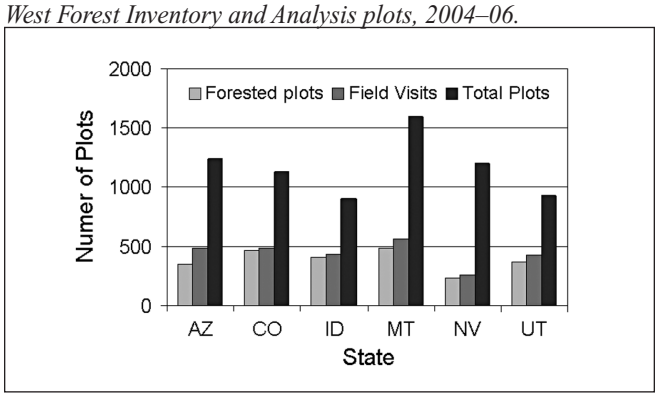
Sara Goeking[1], Gretchen Moisen[2], Kevin Megown[3], and Jason Toombs[4]

Abstract.—Interior West Forest Inventory and Analysis has developed prefield protocols to distinguish forested plots that require field visits from nonforested plots that do not require field visits. Recent innovations have increased the efficiency of the prefield process. First, the incorporation of periodic inventory data into a prefield database increased the amount of information available for making accurate forest or nonforest determinations. Second, acquisition of low-altitude aerial photography proved to be a cost-effective method of verifying forest or nonforest status. Third, tools based on geographic information system technology have decreased the time required to complete the prefield process. Comparisons of field data with prefield determinations verify the prefield process as a cost-effective method of focusing data collection efforts on forested plots.

Introduction

The Forest Service, U.S. Department of Agriculture's (USDA's) Forest Inventory and Analysis (FIA) program, Interior West (IW-FIA) is currently implementing the annual measurement protocol following the national grid design described by Reams *et al.* (2005). FIA's primary mandate is to inventory only those lands that meet its definition of forest. Previous inventories show that less than 50 percent of all IW-FIA plots were actually forested (fig. 1); in 2005, only 34 percent of all plots were found to be forested. Visiting every plot in the inventory grid, including nonforest plots, as part of a forest inventory is logistically and economically infeasible because of the rugged terrain, the patchy access networks, and the sheer amount of nonforest area in the IW-FIA. The cost of data acquisition could be reduced, however, by focusing field efforts on plots that are known, or likely, to be forested. IW-FIA has developed processes, referred to as prefield protocols, to distinguish forested plots that require field visits from nonforested plots that do not require visits.

Figure 1.—*Number of forested, field-visited, and total Interior West Forest Inventory and Analysis plots, 2004–06.*



Prefield Protocols: An Overview

The primary purpose of prefield operations is to reduce unnecessary visits to plots that are known to lie within developed land, rangeland, sparse woodland, agricultural land, remote areas above treeline, or other nonforest lands. A second purpose is the preparation of field materials to aid in field crews' navigation and data collection efforts. Addressing both purposes require the compilation of aerial imagery; topographic maps; previous periodic inventory data, if available; and any other relevant ancillary data, such as species distribution maps. Once these data have been compiled, prefield observers determine whether each plot might meet the definition of forested land. Forest or nonforest determinations are based largely on photo-

[1] Ecologist, U.S. Department of Agriculture (USDA) Forest Service, Rocky Mountain Research Station, Ogden, UT 84401. E-mail: sgoeking@fs.fed.us.
[2] Research Forester, USDA Forest Service, Rocky Mountain Research Station, Ogden, UT 84401.
[3] Forest Biometrician, USDA Forest Service, Remote Sensing Applications Center, 2222 W. 2300 South, Salt Lake City, UT 84119.
[4] Forestry Technician, USDA Forest Service, Rocky Mountain Research Station, Ogden, UT 84401.

interpretation with some consideration of aspect, elevation, and available ancillary information. In addition, prefield observers have experience in collecting FIA field data and thus possess a great deal of local knowledge of forest types and tree species' distributions. Within a prefield database, each plot is classified into one of three categories: (1) forest plots, (2) nonforest plots, and (3) checker plots. IW-FIA's definition of forest land includes areas that meet minimum width and area requirements and that (1) are 10 percent or more stocked by trees of any size now or in the past or (2) have at least 5 percent canopy cover of tree species now or in the past. A checker plot is defined as a location that does not appear to be forest, but the observer cannot tell with certainty that the plot is nonforest. All plots classified as either forest or checker plots will be visited by field crews, while plots that are unquestionably nonforest will not be visited.

Although prefield procedures traditionally involved interpretation of hard-copy aerial photographs and topographic maps, recent innovations have greatly increased the efficiency of the prefield process. Technological tools based on relational databases, low-altitude photography of marginal forest plots, and geographic information system (GIS) tools have all increased the amount of information available for making forest or nonforest determinations while increasing the efficiency of the prefield process, thus decreasing inventory costs.

Database Tools

IW-FIA has developed two database tools that increase the amount of available information for making accurate forest or nonforest determinations and decrease the time required to make such determinations. First, the prefield database includes active links to aerial imagery databases. Many plots in the annual inventory design were never visited during previous periodic inventories, and these locations will be established during the first annual inventory cycle. The quality and accessibility of imagery influences the ability of prefield observers to minimize unnecessary field visits to nonforest plots. The prefield database currently includes links to digital orthoquads (DOQs) with 1-m resolution. An extension of this innovation is the use of the ArcGIS-based Nationwide Select Image Server (ESRI 2006a), which provides access to the most current high quality imagery available for the contiguous 48 States, coupled with advances in ArcGIS software that allow rapid display of precise coordinates at a specified scale (ESRI 2006b).

Second, IW-FIA recently incorporated periodic inventory data into the prefield database. For plots that were visited during periodic inventories in the past, these legacy data include information about tree species' presence, understory vegetation, disturbance, and land use. Both of these database tools enable prefield observers to quickly and easily access useful information for making forest or nonforest determinations.

Low-Altitude Photography

The use of low-altitude aerial photography has refined observers' ability to remotely differentiate forest from nonforest plots. While DOQs and the BOTUS Image Server provide adequate resolution for making initial forest or nonforest determinations, their typical 1-m resolution is insufficient for conclusively separating questionable checker plots into forest or nonforest categories. The acquisition of low-altitude aerial photography for specific plots in the Nevada Photo-Based Inventory Pilot (Moisen 2006) has provided a cost-effective source of high-resolution imagery for verifying forest or nonforest status, thereby reducing the number of field visits to questionable plots.

IW-FIA has also used aerial photography to further refine forest or nonforest classifications. In 2004 and 2005, low-altitude photography with a resolution of about 2 in (15 cm) was acquired and digitally georeferenced for 77 checker plots in Nevada at a cost of $200 per plot for a total cost of $15,400. These 77 plots were categorized as checker plots because they could not conclusively be categorized as forest or nonforest using available DOQs. Based on photointerpretation, 26 plots were identified as nonforest and therefore did not require field visits; 51 plots were verified as forest and were visited at an estimated average cost of $2,000 per plot for a total of $102,000. Thus, the total cost of the imagery acquisition and processing, plus the field costs of visiting verified forest plots, was $117,400. Visiting all 77 plots, at an average cost of $2,000 per plot, would have been $154,000. Therefore, the net savings of acquiring low-altitude photography to reduce unnecessary visits to checker plots was estimated at $36,600. This preliminary cost-benefit analysis shows that low-altitude photography may be a cost-effective tool for refining forest or

nonforest determinations. Further study is needed to perform quality control on these plots and to refine our cost estimates for visits to checker plots since data collection at low-tally plots and nonforest plots typically requires less time than at high-tally plots.

Geographic Information System Tools

One of the most time-consuming parts of the prefield process has been the compilation of aerial photograph stereopairs that field crews use for navigation and plot documentation. Within IW-FIA, most plot locations require aerial photographs from the Aerial Photograph Field Office, which is a division of the Farm Service Agency within USDA, typically consisting of either Forest Service Resource Photography or the photos from the U.S. Geological Survey National Aerial Photography Program (USGS 2007). Until recently, customers of these programs received hard-copy flightline indices or internet-based tools that required the user to individually identify the appropriate flightlines and frames for each plot location. Each year, this process took, on average, 80 person-hours per State, or more than 500 person-hours for all of IW-FIA.

IW-FIA has automated this process by using GIS tools to compile lists of relevant aerial photographs by flightline and station number. Based on proximity analysis in ArcGIS, the flightline and adjacent stations nearest to each plot are identified. This output comprises a list of stereopairs for each plot. The entire process consumes roughly 10 person-hours per State and thus has greatly decreased the time required to prepare field materials.

Quality Assurance Program

The integrity of forest inventory data depends on the ability to accurately distinguish forested from nonforested areas in the context of a specific definition of forested lands. Errors in the classification of plots into forest and nonforest categories can translate into large errors in estimates of total forest area. Errors of commission in the prefield process should be rectified during field data collection because plots that are classified as forest in the office may be found to actually be nonforest; the data are then correctly recorded as nonforest. On the other hand, if a plot is classified as nonforest but is actually forested, it will not be visited during field data collection. This type of error would go unnoticed, and the effect is that total forest area will be underestimated. For this reason, IW-FIA has proposed a quality assurance program for prefield protocols. The design would consist of a random sample, stratified by ecoregion, of a certain percentage of plots that were categorized as nonforest by prefield observers. Ecoregions are meaningful strata for sampling because the sources of error in pinyon-juniper woodland in the Great Basin Desert, for example, are different than those in heavily timbered and logged areas in the northern Rocky Mountains.

A recent nonforest pilot inventory, the Forest Service Region 1 All Condition Pilot (ACP) (O'Brien 2006), provided preliminary data concerning the accuracy of prefield assessments that categorize plots as nonforest. The purpose of the ACP project was to visit all nonforest plots on National Forest System (NFS) lands and collect vegetation data from these plots; the ACP used prefield observations from IW-FIA to identify the sample. Within NFS lands in Montana in Region 1, 126 plots from the existing annual inventories (2003–2006) were classified as nonforest by prefield observers. Field crews found three of these plots to be forested in 2006; these three plots represent errors in the prefield process. During the same inventory period (2003–2006), field crews visited 1,067 forest plots on NFS lands in Montana. Since the prefield assessment for the ACP would have omitted three forest plots from this inventory of 1,067 plots, the error rate was 0.28 percent. This information provides analysts with an estimate of the error associated with forest estimates based on IW-FIA data on NFS lands in Montana.

One goal of the quality assurance program is to calculate error rates for each ecoregion within IW-FIA; another is to use the results to further refine prefield methods. Detection of errors affords the opportunity to identify and rectify sources of error in prefield forest or nonforest determinations. In the case of the ACP, two of the three plots that were classified as nonforest but were actually forest had been clearcut before the previous field visit. Field crews had erroneously determined the plots to be nonforest when they should have recorded them as nonstocked forest plots. When prefield observers considered the previous field data, they assumed that the field observations were correct despite the evidence of timber harvest on aerial imagery. The

lesson from this analysis is that individual pieces of information, including previous inventory data, are insufficient to eliminate plots from the inventory unless the information is interpreted in the context of all available information.

Discussion

The prefield protocols employed by IW-FIA have decreased inventory costs by minimizing unnecessary field visits to nonforest plots, thereby focusing field efforts on forested plots. Recent innovations have increased the efficiency of the prefield process. Incorporating periodic inventory data within a prefield database and acquiring high quality imagery have both increased the amount and quality of information available to prefield observers. Tools based on GIS technology have also decreased the time required to complete the prefield process. Comparisons of field data with prefield determinations based on these procedures show that the overall prefield process is a cost-effective and reliable method of focusing data collection efforts on forested plots.

The prefield methods described here can be applied to any vegetation inventory that would experience cost savings by differentiating forest and nonforest plots. Other forest inventory projects could use these methods in conjunction with more locally appropriate methods, such as the Common Land Unit information used by Liknes and Nelson (2009) in the Northern Research Station FIA unit. Nonforest vegetation inventories, such as the Forest Service Region 1 ACP, could compile their samples based on FIA prefield data. Finally, prefield methods may also streamline data collection efforts as both the definition of forest lands and the designation of tally species evolve. If these definitions change in the future, the periodic inventory data, which includes information about plant species present and site history, will aid in designation of plots as forest or nonforest under the new definitions.

Acknowledgments

The authors thank Roger Boyer of the White Mountain National Forest, formerly of Interior West Forest Inventory and Analysis, for information about data acquisition costs and thank Paul Patterson of the Forest Service Rocky Mountain Research Station for statistical consultation regarding future sample designs for quality assurance of prefield protocols.

Literature Cited

ESRI. 2006a. ArcGIS 9.2. http://www.esri.com/news/arcuser/0807/forest_image.html. (3 January).

ESRI. 2006b. ArcGIS 9.2. http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm. (22 December).

Liknes, G.C.; Nelson, M.D. 2009. Improving FIA inventory efficiency with common land unit information. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C.; McWilliams, W.J. (eds.). Proceedings, eighth annual forest inventory and analysis symposium, Monterey, CA. Gen. Tech. Rep. WO 79. Washington, DC: U.S. Department of Agriculture, Forest Service: 363-367.

Moisen, G. 2006. Personal communication. FIA Research Forester, U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, 507 25th Street, Ogden, UT 84401.

O'Brien, R. 2006. Personal communication. FIA Deputy Program Manager, U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, 507 25th Street, Ogden, UT 84401.

Reams, G.A.; Smith, W.D.; Hansen, M.H. *et al.* 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. 85 p.

USGS. 2007. National Aerial Photography Program. http://eros.usgs.gov/guides/napp.html. (3 January).

Is There a Better Metric Than Site Index To Indicate the Productivity of Forested Lands?

Maria E. Blanco Martin[1], Michael Hoppus[2], Andrew Lister[2], and James A. Westfall[2]

Abstract.—The Forest Service, U.S. Department of Agriculture's Forest Inventory and Analysis (FIA) program selects site trees for each plot that are used to measure site productivity. The ability of a site to produce wood volume is indicated indirectly by comparing total tree height with tree age. This comparison assumes that the rate of height growth is strongly related to site quality and is insensitive to basal area, species composition, and stand structure. Research indicates that stand age is often difficult to determine, especially in uneven-aged stands. Furthermore, stands with mixed species compositions and less than full stocking cause problems when using site index as a predictor of site growing capacity. Now that the FIA program has thousands of plots in which volume has been remeasured, other metrics for site quality can be evaluated by noting the observed past growth of the trees on the plot and comparing it with the height-age relationship. This study describes the first steps of this effort.

Introduction

The productivity of forests lands is largely defined in terms of site quality as an indicator of the potential of the site to produce wood given a particular species or forest type. Considerable interest has long existed in developing ways to predict potential forest productivity directly without using forest variables (diameter at breast height (d.b.h.), height, and so on), trees, or other vegetation as indicators.

The indirect approaches for estimating site quality (Jones 1969) include site index, vegetation, and environment. Site index is the height to which a tree would grow under forest conditions in a specified number of years, the index age (50 years for inventory plots in Pennsylvania), or the height of a dominant tree at a particular diameter (Bickford *et al.* 1957). Site index determinations require the availability of height-age curves for a particular area and species. Areas without suitable trees may not be accurately classified. Site index is normally associated with even-aged, single species and fully stocked stands. In mixed hardwood forests, such as those often found in Pennsylvania, these conditions are uncommon. Many studies have developed predictive mechanisms in the form of multiple regression equations in which site index is determined from a number of independent site variables (Armson 1977). The FIA program in Pennsylvania currently uses site productivity classes derived from site index, based on fully stocked stands at the age of culmination of mean annual increment, to indicate site quality (Scott and Voorhis 1986). Because of the known difficulties associated with using site index as an indirect metric for site quality, we hypothesize that site productivity classes are not well correlated with actual stand growth. Problems associated with using site index to estimate the potential productivity on FIA plots have been documented before. Only 12 percent of 65 remeasured plots in Vermont retained a constant classification over three surveys. Several of the plots varied by up to three site classes gained or lost (Beattie 1979). The FIA program now offers thousands of remeasured plots per year using the same plot design. This abundance enables the comparison and evaluation of site quality metrics with actual growth rates.

[1] Ph.D. Student, Politechnical University of Madrid, E.T.S.I. Montes, U.D. Estadistica. Ciudad Universitaria S/N, 28040, Madrid, Spain. Phone: +34 306-6401. E-mail: mariaelena.blanco@upm.es.

[2] Research Forester, U.S. Department of Agriculture (USDA) Forest Service, Northern Research Station, Newtown Square, PA 19073–3294.

Objective

The objective of this article is to test two hypotheses: (1) site productivity class measured on FIA plots is unrelated to actual growth and (2) no other metrics attributed to FIA plots work better than site productivity class. Furthermore, if we can reject either of these hypotheses, we would like to use geostatistical methods to map site quality over the landscape. It is important to understand that we are not trying to predict growth per se; instead, we want to find a metric that predicts site quality or potential wood productivity.

Methods

For this study, site productivity class was compared with growth. The environmental approach for assessing site quality was used in locations where factors affecting site quality, such as climatic factors (temperature, precipitation, humidity, and solar radiation) and topographic factors (slope, aspect, elevation, longitude, and latitude), were evaluated. Stocking, stand size, and age were used to stratify the plots to eliminate confounding growth competition factors. Remeasured growth on FIA plots is the response variable that was used as an indicator of site quality.

The study followed these steps:

- Compare plot site productivity class with actual growth.
- Identify other site factors that might accurately predict site quality.
- Develop multiple regression models to map areas of different site quality in Pennsylvania as a result of the integration of these nonbiological factors.

The data used in this research were collected from the Forest Service's Northern Research Station, FIA unit. Under this federally mandated program, sampling is based on an interpenetrating panel design at an intensity of 1 plot per approximately 6,000 acres. Data were collected from the periodic survey conducted in 1989 and compared with annual plots collected from 2000 to 2004 in the State of Pennsylvania. The time interval between plot remeasurement was used to calculate annual growth.

On each sample plot, data were summarized by joining the tree and plot tables, considering only the remeasured plots with trees having a d.b.h. of 5.0 in or larger and plots with no removals or mortality. Condition-level variables were assigned to each plot based on the largest condition area. Finally, the subplot-level table was added to the summary to include topographic factors measured directly by the crew in the field. The number of plots that met the criteria was 558 (11.5 percent). The plot distribution and study areas are shown in figure 1.

FIA data provided the following metrics for each plot: stand size and age, site productivity class, forest type, aspect, eleva-

Figure 1.—*Distribution of forest inventory and analysis plots.*



tion, latitude, longitude, and slope. Temperature, precipitation, solar radiation, and humidity were provided by the DAYMET U.S. Data Center. This resource provides daily surface weather data and climatological summaries of the United States, at a 1-km resolution, from ground-based meteorological stations over a 9-year period (1989 through 1997).

The aspect metric was transformed into a linear variable ranging from 0.00 to 2.00 using the transformation of Beers *et al.* (1966) (i.e., A′ = cos (Amax − A) + 1, where A′ = transformed aspect, Amax = 45°, and A = the measured aspect.

Growth for each plot was analyzed in terms of accretion, considering three components: (1) basal area, (2) sound volume, and (3) gross volume; see equations (1), (2), and (3).

Basal area accretion:

$$ACCBA = \frac{\sum (BA_{ij2} - BA_{ij1})}{r_i} \quad (1)$$

Gross volume accretion:

$$ACCGROS = \frac{\sum (GV_{ij2} - GV_{ij1})}{r_i} \quad (2)$$

Sound volume accretion:

$$ACCSOUND = \frac{\sum (SV_{ij2} - SV_{ij1})}{r_i} \quad (3)$$

where BA_{ijk} = basal area (ft^2) of tree *i* on plot *j* at measurement *k* (K=1,2), GV_{ijk} = gross volume (ft^3/acre) of tree *i* on plot *j* at measurement *k* (K=1,2), SV_{ijk} = sound volume (ft^3/acre) of tree *i* on plot *j* at measurement *k* (K=1,2), and r_i = remeasurement period (year) for plot *i*.

SAS 9.1 was used to summarize the data, perform statistic analyses, and develop models. ArcGIS 9.1 and ArcInfo Workstation were used to generate maps based on the models.

Data Analysis Techniques

The data analysis techniques included the following:

Frequency Analysis

This technique was used to determine how the plots should be grouped into different categories of tree species, stand age, and stocking. The frequency distributions of these variables were evaluated and appropriately sized intervals were selected. Because growth rates can be strongly influenced by differences in these three variables, it is necessary to remove these sources of confounding growth variation as much as possible. Consider that a fully stocked stand of older trees on a poor site will most likely have a greater growth rate than that of a poorly stocked stand of young trees on a good site. Obviously, forest type group is already divided. The variables of age and stocking were grouped as follows: Age class ranges from 1 (youngest) to 4 (oldest) age (0 to 30, 31to 60, 61to 90, and older than 90 years, respectively); stocking class ranges from 1 (lowest) to 3 (highest) percent (0 to 60 percent, 61 to 80 percent, and more than 80 percent, respectively).

Dependence Analysis

Analysis of variance (ANOVA) was carried out to test for differences among the means of the levels of the categorical variables (forest-type group, stand age, stand size, stocking, and site productivity).

The assumptions are that the observations are independent and randomly selected from normal populations with equal variances. These assumptions were assessed by examining plots of the residuals. A statistical comparison of site productivity and growth was made by using ANOVA and scatterplots. Tukey's test was used to determine if any of the tree species, stand age, and stocking groups were statistically different from one another.

Correlation Study

This study, which was completed for quantitative explanatory variables, used Spearman's rank correlation method to discard highly correlated variables. All variables were considered.

Estimating Site Quality by Multifactorial Analysis

Multiple linear regression (MLR) was used to help select the independent variables to predict the quantitative dependent accretion variables. Modeling site quality, from environmental factors using accretion as the quality metric, was carried out using SAS 9.1 to perform MLRs. The following relationship was assumed:

$$SQ = \beta_0 + \beta_1 * X_1 + \ldots\ldots + \beta_n * X_n + \varepsilon$$

where SQ is site quality, expressed in terms of accretion, X_1….. X_n are the vectors corresponding to site variables, $\beta_1 \ldots \beta_n$ represent model coefficients, and ε is the additive error term.

The assumptions were that the expected value of the residuals is 0, or E[e] = 0, which implies that the relationship is linear in the explanatory variables, the errors are distributed with equal variance (homoscedasticity), the errors are independent, the predictors are not correlated, and the errors are normally distributed.

Regarding MLR, the normality of the metrics was checked by using the standardized skewness Z-score test, which examined the lack of symmetry in the data. Of all the variables, it was necessary to transform the variable slope to $\log_{10}$(slope) and the variable aspect to sqrt($\log_{10}$(aspect)) to normalize them. Within the site factor variables, it was necessary to select the ones that could be used to model the site quality variation. The analysis began by examining the variables available in terms of their relationship with accretion in basal area and volume. Simple plots of accretion and growth versus potential predictor variables were analyzed in an attempt to understand the relationship between the variables and growth (e.g., suggesting linear or nonlinear relationships).

Stepwise regression was used because it is considered the best method of automated variable selection (Draper and Smith 1998). This automatic procedure for statistical model selection was used because no underlying theory exists on which to base the selection model. A model that is a result of an automated variable selection process may not be biologically sensible, however, and needs to be evaluated on biological grounds.

The adjusted R-squared selection method was chosen and Akaike's Information Criterion (AIC) (Akaike 1974) statistic was used as the criterion, or the measure, of the goodness of the fit for selecting models. AIC is calculated from

$$AIC = 2\ln(RMSE) + \frac{2c}{n}$$

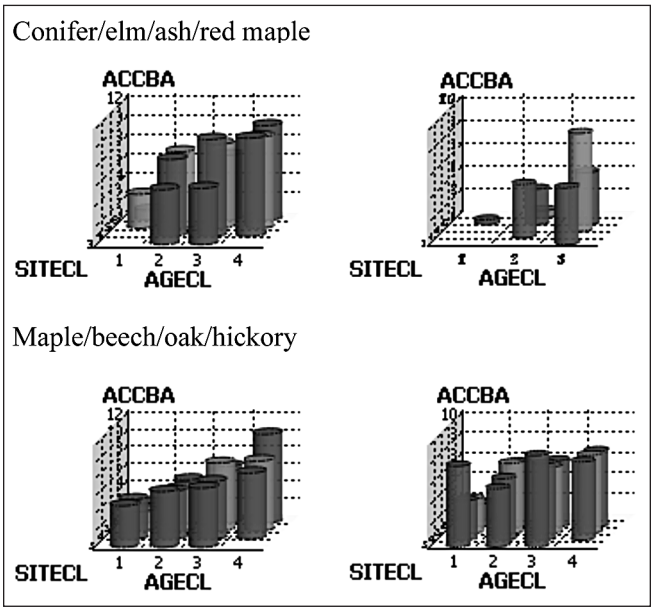
where *RMSE* is the root mean squared error during the estimation period, *c* is the number of estimated coefficients in the fitted model, and *n* is the sample size used to fit the model.

Results

Site productivity classes, based on site index, were plotted against actual growth. In figure 2, mean basal area accretion is presented versus site productivity class and age class for four forest types. The closest site class on the x axis is associated with the highest site productivity class. Notice that the conifers plots show the lowest growth occurring on the best sites. The same pattern of high variability of site productivity classes occurs for sound and gross volume accretion.

According to Tukey's test for ANOVA, growth means were not significantly different (P-value greater than 0.05) among the five productivity classes. Surprisingly, the same observation

Figure 2.—*Mean basal area accretion by forest type.*



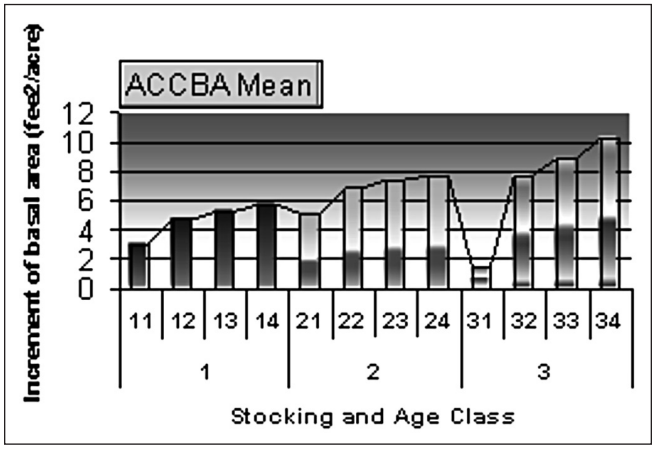
ACCBA = mean basal area accretion. AGECL = age class. SITECL = site class.

was true for forest types. Significant differences (P-value less than 0.05) were found when stocking, size, and age class were analyzed against growth.

Therefore, after analyzing the frequency distributions of both age class and stocking classes, these classes were used to divide the data set of 558 plots into logical compositional growth groups. Figure 3 shows the histogram of increment of basal area for the stocking and age class groups. Within each increase of stocking class, basal area increment rises with the increasing of stand age. Mean sound and gross volume increment follow the same pattern as basal area increment. This result led us to consider 12 different subpopulations by stocking class and stand age to model growth.

For each of the 12 subpopulations and for each dependent variable, simple regression was applied to make models to explain the variability of each one of the accretions. The importance was ranked by the highest R-square obtained. (see tables 1, 2, and 3). Models with the minimum or lowest AIC had R-squares of about 0.3. This observation indicates high variability of the data.

Figure 3.—*Mean basal area increment by stocking class and stand age class.*



The following is an example of a site quality model (using stepwise regression) based on gross volume accretion, created for the subpopulation consisting of stocking class of 3 and age class of 3. To further test the importance of the different factors influencing site quality, we separately compared the correlations between gross accretion and each one of the site factors. The correlation coefficients are listed in table 4. The major factors influencing gross accretion are latitude and slope, with P-values of 0.0063 and 0.0112, respectively.

Table 1.—*Rank of the major factors influencing site quality, by simple regression according to the highest R-square, for mean basal area accretion.*

ACCBA	Variable	Ranking
	Latitude	1
	Longitude	2
	Elevation	3
	Aspect	4
	Temperature	5

Table 2.—*Rank of the major factors influencing site quality, by simple regression according to the highest R-square, for gross volume accretion.*

ACCGROS	Variable	Ranking
	Temperature	1
	Latitude	2
	Precipitation	3
	Longitude	4
	Aspect	5
	Slope	6

Table 3.—*Rank of the major factors influencing site quality, by simple regression according to the highest R-square, for sound volume accretion.*

ACCSOUND	Variable	Ranking
	Latitude	1
	Longitude	2
	Precipitation/Slope	3
	Aspect	4
	Elevation	5

Table 4.—*Pearson correlation coefficients.*

	Independent variable								
Dependent variable	ASPECT	ELEVATE	LAT	LON	SLOPE	P	T	VPD	SRAD
ACCGROS	– 0.15	– 0.04	– 0.34	– 0.08	– 0.32	– 0.07	0.01	0.14	0.02
Prob > |r|	0.23	0.74	0.01	0.51	0.01	0.61	0.94	0.28	0.86

Notes: N = 64

MLR provided the following model:

Dependent variable: ACCGROS; independent variables: SLOPE, LAT, VPD, and T; R-square = 39.9711 percent; R-squared = 35.6054 percent; standard error of estimate = 36.93; mean absolute error = 26.5597; Durbin-Watson statistic = 2.10416; P = 0.6069; Lag 1 residual autocorrelation = -0.0532975.

The output (see table 5) shows the results of fitting an MLR model to describe the relationship between ACCGROS and the independent variables SLOPE, LAT, T, and VPD. The equation of the fitted model is as follows:

ACCGROS = 99.0402 - 0.756874*SLOPE - 0.00015621*LAT - 17.0289*T + 0.60851*VPD

Because the P-value in the ANOVA table (table 6) is less than 0.05, a statistically significant relationship exists between the variables at the 95.0-percent confidence level.

The R-square statistic indicates that the model, as fitted, explains 40 percent of the variability in ACCGROS. The adjusted R-squared statistic, which is more suitable for comparing models with different numbers of independent variables, is 35 percent. The standard error of the estimate shows the standard deviation of the residuals to be 36.93. The mean absolute error of 26.5 is the average value of the residuals.

Table 5.—*Multiple regression fits.*

Parameter	Estimate	Std error	T- statistic	P-value
CONSTANT	99.0402	99.8446	0.991943	0.3256
SLOPE	– 0.756874	0.237175	– 3.1912	0.0023
LAT	– 0.0001562	4.895E–05	– 3.19122	0.0023
VPD	0.60851	0.204799	2.97126	0.0044
T	– 17.0289	9.45356	– 1.80132	0.0771

LAT = latitude. T = temperature. VPD = vapor pressure deficit

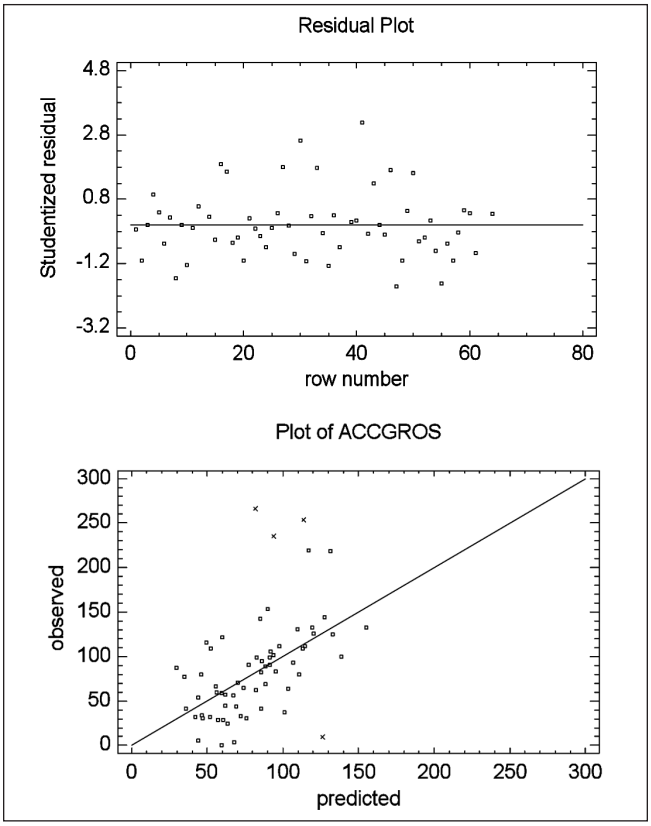
Table 6.—*Analysis of variance table.*

Source	Sum of squares	Degrees of freedom mean square	F-ratio	P-value
Model	49946.7	12486.7	9.16	<0.0001
Residual	75010.3	1363.82		
Total (Correlation)	124957.59			

In determining whether the model can be simplified, we chose a 90-percent confidence level as the basis for retaining predictor variables. The highest P-value on the independent variables is 0.0771, belonging to T. Because the P-value is not greater or equal to 0.1, that term is still statistically significant at the 90.0-percent or higher confidence level. Predicted versus observed data and residual plots are shown in figure 4. Residuals follow a normal distribution, are randomly distributed in figure 4, and have a mean of 0.

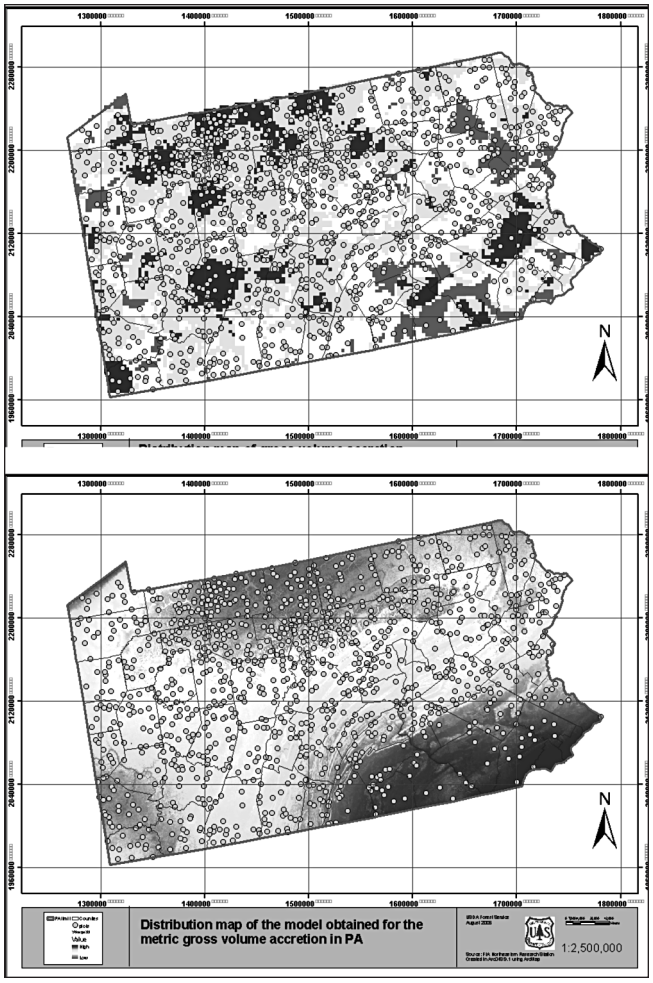
A map of site quality was produced for Pennsylvania by applying the model to the values of the predictor variables, representing each 250-m pixel, in an array of pixels covering the State (fig. 5). This map was compared with a map of actual growth (gross accretion (ft^3/acre/year) for the stocking and age class block 33 (stocking > 80 percent and age from 61 to 90 years). This map was produced by using a GIS moving window that statistically summarized groups of plot data and displayed the

Figure 4.—*Residual plots and predicted versus observed data for the model.*



ACCGROS = gross volume accretion.

Figure 5.—*Map comparison of actual growth (gross volume accretion and site quality based on a model of gross volume growth predicted by slope, latitude, humidity, and temperature). In order of increasing growth, the colors are red, yellow, turquoise, and blue. Plot locations are approximate.*



results spatially. As mentioned previously, a site quality map does not necessarily correspond to a map of actual growth. It is interesting to note that the area of greatest growth potential, southeastern Pennsylvania, is considered a productive agricultural area and is known for productive forest sites.

Conclusions and Discussion

Site quality is not easy to assess using the response variable of growth over about 10 years. The factors of site quality, site history, and the vegetation itself are interacting and interdependent, making it difficult to assign cause-and-effect relationships. This study found no straightforward measure of site quality to be entirely satisfactory.

Lessons learned and ponderings:

- Predicting site quality is more difficult than predicting growth because the effects of stocking and stand age have to be eliminated.
- Our results are consistent with other studies (Beattie 1979) that have found that site quality is not easy to assess because site factors and trees themselves are interacting and interdependent.
- It does appear that the site productivity classes of the plots have a poor correlation to actual growth; however, the hypothesis that states that no relationship exists between FIA site productivity classes and growth deserves to be tested over a greater time period than 10 years before making any definite conclusions.
- Soil type was eliminated in the AVOVA test because of technical difficulties with the GIS layer and concern about the soil map accuracy. Soil type is documented to be a major predictor of site quality and must be included in any follow-on study.
- The model and corresponding map need a proper accuracy evaluation despite the fact that 40 percent of the variability of growth was explained in well-stocked stands of advanced age provides some hope that site quality can be predicted.

Literature Cited

Akaike, H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 19(6): 716-723.

Armson, K.A. 1977. Forest soils: properties and processes. Toronto, Canada: University of Toronto Press. 390 p.

Beattie, M.H. 1979. An evaluation of the U.S. Forest Service system for predicting the potential productivity of northern hardwoods in Vermont. Burlington, VT: Graduate College of the University of Vermont. 108 p.

Beers, T.W.; Dress, P.E.; Wensel, L.C. 1966. Aspect transformation in site productivity research. Journal of Forestry. 64(10): 691-692.

Bickford, C.A.; Baker, F.S.; Wilson, F.B. 1957. Stocking, normality, and measurement of stand density. Journal of Forestry. 55: 99-104.

Draper, N.R.; Smith, H. 1998. Applied regression analysis. 3rd ed. New York: John Wiley. 706 p.

Jones, J. 1969. Review and comparison of site evaluation methods. Res. Pap. RM-51. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station. 27 p.

Scott, C.T.; Voorhis, N.G. 1986. Northeastern forest survey site index equations and site productivity classes. Northern Journal of Applied Forestry. 3(4): 144-148.

Improving Forest Inventory and Analysis Efficiency With Common Land Unit Information

Greg C. Liknes[1] and Mark D. Nelson[2]

Abstract.—The Forest Service, U.S. Department of Agriculture's (USDA's) Northern Research Station Forest Inventory and Analysis program (NRS-FIA) examines inventory locations on digital aerial imagery to determine if the land use at each plot location meets the FIA definition of forest and thereby becomes a field visit site. This manual image-interpretation effort requires a significant amount of staff time yet accrues substantial financial savings. Using the same imagery currently used by NRS-FIA, the Service Center Agencies of the USDA have initiated work on a digital data set detailing agricultural areas, including a land class designation. These data are known as Common Land Units (CLUs). In this study, CLU data for the State of Minnesota were acquired from USDA's Farm Service Agency. NRS-FIA manual image-interpretation information for inventory locations was compared with the land class information in the CLU data set. Thirty-five percent of Minnesota FIA plots were identified as nonforest using the CLU data set. A little less than 1 percent of plots fell within nonforest CLU polygons yet were observed during field visits to be forested. Thus, the CLU method for identifying nonforest plots may introduce a small underestimation bias into the inventory. Because NRS-FIA plots previously observed as forested will be revisited regardless of an updated land use status, however, this bias would not occur. Thus, the CLU data set shows the potential to eliminate manual, prefield image interpretation of plots occurring within agricultural areas delineated by CLU boundaries.

Introduction

The Forest Service's Northern Research Station Forest Inventory and Analysis program (NRS-FIA) examines inventory locations on digital aerial imagery to determine if land use at plot locations meets the FIA definition of forest. If the criteria are satisfied, the location becomes a field visit site. Not visiting nonforest plots accrues substantial financial savings, particularly in States in which agriculture dominates the landscape. For example, of approximately 33,000 inventory plots image interpreted for North Dakota, South Dakota, Nebraska, and Kansas during the 5-year first annual FIA inventory, fewer than 4 percent were forested.

Ideally, land use status at plot locations is reexamined during every inventory cycle using updated imagery to detect changes that have occurred since the previous inventory. Although previously forested plots will be revisited in the subsequent cycle regardless of land use change, previously nonforested plots may change to forest land but not be identified as forested plots unless more current image products are interpreted. Omitting forested plots from field visits may result in biased estimates. This prefield effort requires a significant amount of image-interpretation time. The objective of this study was to determine to what extent image-interpretation work could be reduced by using other sources of land use information to determine whether FIA plots are forested.

Data

FIA Plots

The first annual FIA inventory in Minnesota included 16,383 plots (1999–2003). The geographic coordinates of each plot were precisely identified in one of two ways, depending

[1] Research Physical Scientist, U.S. Department of Agriculture (USDA) Forest Service, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108. E-mail: gliknes@fs.fed.us.

[2] Research Forester, USDA Forest Service, Northern Research Station, 1992 Folwell Avenue, St. Paul, MN 55108. E-mail: mdnelson@fs.fed.us.

on whether the plot was forest or nonforest. Locations for nonforest plots were obtained by transferring hardcopy photo pinpricks via Geographic Information System (GIS) software and heads-up digitizing to high-resolution, orthorectified digital aerial imagery. Locations for forested plots were recorded via a Global Positioning System receiver, and these locations were passed through a set of quality checks to ensure they were reasonably accurate.

In addition, the land use code, determined by image interpretation, was queried for each plot. Each image-interpreted land use has an associated field visit decision. For example, plots that appear to fall on or very near forest land receive a field visit while plots appearing to fall on pasture or rangeland do not. All previously forested plots are revisited, regardless of current image-interpretation results. In addition, a fraction of field-visited plots actually were determined to have nonforest land use. This land use information allowed us to compare image-interpretation–based prefield decisions (i.e., whether a plot requires a field visit) with decisions made using an alternative source of land use data.

Common Land Units

The Service Center Agencies of the U.S. Department of Agriculture (USDA), including Farm Service Agency (FSA), Natural Resources Conservation Service, and Rural Development, have initiated work on a digital data set detailing agricultural areas. These data are known as Common Land Units (CLUs) and are defined as the smallest units of land that have the following properties: a permanent contiguous boundary, common land cover and land management, a common owner, and common producer association. Areas were digitized from National Agriculture Imagery Program (NAIP) images, the same imagery currently used by NRS-FIA to make determinations of land use. CLU data for a large portion of the country have been completed, and updates will occur on a continual basis. More information on the Common Land Unit data set can be found on FSA's Web site.[3]

CLU spatial data (polygons) are readily available to USDA agencies, but CLU attribute data are restricted due to landowner privacy concerns. In this study, CLU spatial data were acquired for the State of Minnesota along with a subset of CLU attributes relating to land class designation for each CLU polygon. Land classes include Urban, Cropland, Rangeland, Forest, Water Body, Barren, Tundra, Mined Land, and Other Agricultural Land. No CLU attributes pertaining to landowner information were obtained. Figure 1 shows an example of the NAIP imagery used to create the CLU data set and CLU polygons with their associated land classes for an agricultural area in Minnesota. The CLU data set for Minnesota includes more than 1.2 million polygons.

Figure 1.—*The image at the top is a grayscale representation of a National Agriculture Imagery Program aerial image taken over an agricultural area in Minnesota. The bottom image shows the associated Common Land Unit (CLU) land classes derived from that image. The hatched polygons represent cropland, the gray polygons represent the Other Agricultural Land class, and the narrow linear white features corresponding to roadways are not attributed in CLU.*



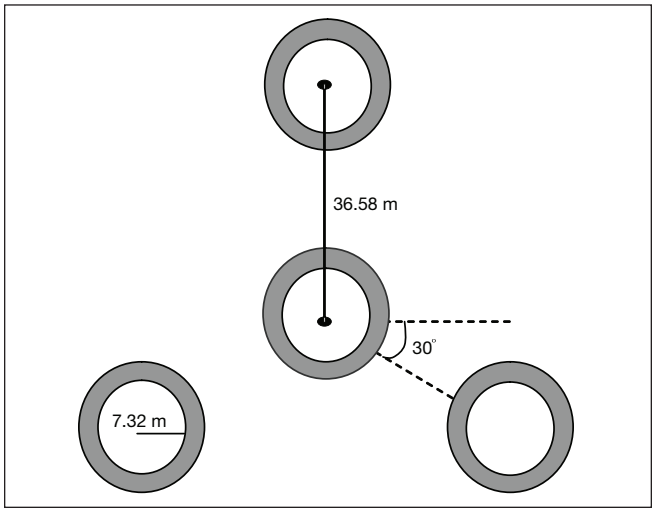
[3] http://www.fsa.usda.gov.

Methods

Using GIS software, plot coordinates (latitude, longitude—North American Datum of 1983) were reprojected to match the projection of the CLU data set (Universal Transverse Mercator, Zone 15N,—North American Datum of 1983). Polygons were constructed to represent FIA's plot design (see the inner circles in fig. 2). Next, 3-m buffers were constructed around each subplot to account for the reported uncertainty in CLU horizontal position (see the gray band surrounding each subplot in fig. 2). A subset of the CLUs was then created which included only polygons in definite nonforest land classes (i.e., polygons with missing or unknown land classes were excluded). This CLU subset was processed using standard GIS functionality to remove small "sliver" polygons.

GIS software was used to select FIA plots (subplots with buffers) that lie completely within nonforest CLU boundaries. This selected set was then labeled as nonforest. Conversely, plots that lie only partially within or completely outside nonforest CLU boundaries were not assigned a label. Because areas outside the nonforest CLU polygons can be either forest or nonforest, plots lying in these areas require additional information to make a forest/nonforest determination. FIA plots labeled as nonforest using the CLU method were then assessed in terms of the image-interpreted land use and/or the land use determined during the field visit.

Figure 2.—*Forest Inventory and Analysis plot design with 3-m buffer (outer gray circles) around each subplot (inner circles).*



Results

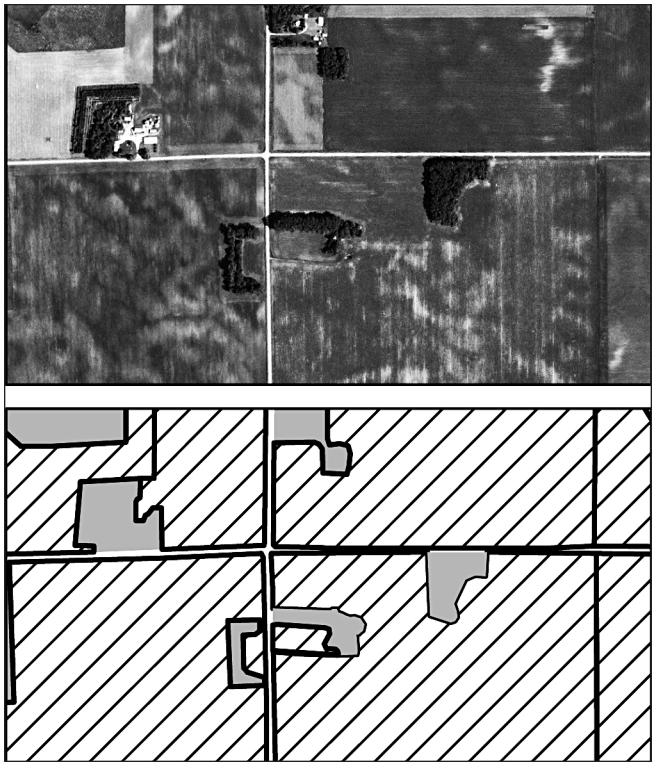
Statewide, the CLU method identified 5,741 out of 16,383 Minnesota FIA plots (35 percent) as nonforest. Of those 5,741 plots, FIA's image-interpretation method also identified 5,551 nonforest plots. A field visit was deemed necessary for the remaining 190 plots identified via image interpretation. Of those 190 plots, 67 were observed to be nonforest during the field visit. The remaining 123 plots are cases in which the CLU method would have led to forested plots being called nonforest. If a single-intensity FIA sample is assumed and each plot represents approximately 2,422 ha, these "missed" plots represent a potential underestimation of Minnesota's forest land area by 298,000 ha, or about 0.95 percent of Minnesota's 6,568,416 ha of forest land (Miles 2001).

The 123 forested FIA plots incorrectly identified by the CLU data set as nonforest are distributed across the various CLU land classes in the following manner: Urban, 1 plot; Cropland, 6 plots; Rangeland, 29 plots; Water Body, 1 plot; and Other Agricultural Lands, 86 plots. Independent image interpretation of Other Agricultural Land polygons revealed riparian trees adjacent to cropland and large tree plantings in proximity to farmsteads in this land class (see the example in figure 3). In some instances, these trees met the FIA definition of forest and therefore required a field visit.

Of the remaining 10,642 plots that the CLU method could not identify as nonforest, a field visit was not required for 5,381, as determined by FIA image interpretation. These plots can be viewed as lost opportunities for the CLU method to reduce image-interpretation work, but would not introduce bias into FIA estimates.

The CLU method shows the most potential savings to FIA in heavily agricultural counties in the southern and western regions of Minnesota. In one county, CLU data identified 89 percent of all FIA plots as nonforest. By extrapolation, the States of North Dakota, South Dakota, Nebraska, and Kansas could show similar results because of the predominance of agricultural (nonforest) land use in those States.

Figure 3.—*A National Agriculture Imagery Program image (top) showing a group of trees (just right of center) that meet the Forest Inventory and Analysis definition of forest. The image at the bottom shows the associated Common Land Unit "Other Agricultural Lands" polygons in gray, and Cropland is represented by the hatched areas.*



Conclusions

The CLU method for identifying nonforest plots has the potential to eliminate a large amount of manual image interpretation, particularly in heavily agricultural areas. Despite this potential benefit, the CLU method incorrectly identified 123 out of 16,383 forested plots as nonforest in the State of Minnesota. Although this figure represents less than 1 percent, it represents the possibility for a biased underestimation of forest land area if those plots are excluded from field visits. It should be noted that the goal of photo interpretation in the second annual inventory cycle will be to identify plots that were nonforest in the previous inventory that have become forested or partially forested. In this scenario, the 123 plots with CLU-FIA disagreement would have received a field visit despite the CLU determination because of their previous status as forested. Therefore, in a second annual inventory, the real concern is how often the CLU method misses plots that have converted from nonforest to forest land use. Additional study is required to make this determination.

As stated previously, 86 of the 123 errant nonforest determinations using the CLU method occurred in the Other Agricultural Lands class, which comprises 635,000 ha in Minnesota. Removing this class from the nonforest CLU subset would increase the number of plots that cannot be assessed using the CLU method but would also reduce the potential for underestimation of forest area in the inventory.

A recommendation for increasing the number of plots that can be identified as nonforest using the CLU data set is to supplement the analysis with road information from another source. Roads are not part of the CLU data set, and they frequently separate cropland fields in agricultural landscapes (fig. 1). Plots that either touch or straddle a road were not identified as nonforest using the method in this study, even if the areas adjacent to the road were completely nonforest. Adding roads to the CLU data set would allow the method in this study to correctly identify plots in this scenario as nonforest. In addition, a large portion of the CLU polygons have a land class of "None" or the land class is missing (more than 2.5 million ha). Updates to the CLU data set may eliminate some of these issues and could greatly improve results obtained using the method described in this study.

Acknowledgments

The authors thank Scott Kapphahn and Lisa MacDonald of the U.S. Department of Agriculture for their assistance in providing CLU data.

Literature Cited

Miles, P.D. 2001. Forest inventory mapmaker users guide. Gen. Tech. Rep. NC-221. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station. http://www.ncrs2.fs.fed.us/4801/fiadb/available_data.html. (6 November 2006).

Phase 2 and Phase 3 Presentation Grids

Joseph M. McCollum[1] and Jamie K. Cochran[2]

Abstract.—Many forest inventory and analysis (FIA) analysts, other researchers, and FIA Spatial Data Services personnel have expressed their desire to use the FIA Phase 2 (P2) and Phase 3 (P3), and Forest Health Monitoring (FHM) grids in presentations and other analytical reports. Such uses have been prohibited due to the necessity of keeping the actual P2, P3, and FHM grids confidential. The Environmental Protection Agency's Environmental Monitoring and Assessment Program (EMAP) grid is often touted as a secure alternative to the FHM grid for legacy FHM plots, but the current FIA P3 sampling frame differs from its FHM predecessor. This article offers an approach involving a weighted bipartite matching algorithm using an alternative to fuzzing and swapping coordinates for FIA P2 and P3 plots to assign plots to EMAP grid points by county and shows preliminary results that keep the actual P2, P3, and FHM grids confidential.

Introduction

In a three-phase sampling design implemented by the Forest Inventory and Analysis (FIA) program, Phase 1 (P1) provides the initial estimate for forest area in double sampling, Phase 2 (P2) is the set of plots from which forest inventory data is collected, and Phase 3 (P3) is a subset of Phase 2 plots (1/16th of P2; details may be found in McCollum and Cochran (2005) from which forest health data is collected. (USDA Forest Service 2005).

For landowner security reasons, FIA data is subject to the Food Security Act of 1985, which states that no one may "disclose such information to the public, unless such information has been transformed into a statistical or aggregate form that does not allow the identification of the person who supplied particular information," (7 U.S.C. 2276). Because FIA plot locations are selected using FIA's hexagon grid sampling array, neither FIA plot location coordinates nor are FIA hexagon locations released to the public. In McCollum and Cochran (2005), for example, several proposed maps showing portions of the P2 and P3 grids were rightfully withheld from the final FIA proceedings publication to comply with this precaution of not releasing hexagon locations. Hence, only maps containing hexagons and nondescript State or county lines were shown, which makes it impossible to link a point on a map to its actual ground location. Since McCollum and Cochran (2005) was presented at the fifth annual Forest Inventory and Analysis Symposium in 2003, FIA has adopted a strategy for releasing plot data, outlined in Lister *et al.* (2005), that is informally known throughout the FIA program as "fuzzing and swapping." It is also referred to by McRoberts *et al.* (2005), as "perturbing and swapping," which involves perturbing the coordinates of the plots so that they are untraceable to any certain location on the ground. In areas where legacy plots are rare, the geometry of the P2 grid is preserved because not more than one plot is spatially proximate to any given grid point. In areas where legacy plots are abundant, the geometry of the P2 grid is not preserved. In neither case is the geometry of the P3 grid preserved. Preserving the geometry of the grid leads to generally equal-sized cells.

The Environmental Protection Agency (EPA) Environmental Monitoring and Assessment Program (EMAP) sampling design incorporates a systematic triangular grid to ensure random selection and appropriate spatial distribution of samples (Overton *et al.* 1990, White *et al.* 1992). EMAP cells are equal in size to the initial FHM cells which were 27 times the size of P2 cells (McRoberts *et al.* 2002). However, when the Forest Health Monitoring program was implemented, the FHM grid was shifted with respect to the EMAP grid.

[1] Information Technology Specialist, U.S. Department of Agriculture (USDA) Forest Service, Forest Inventory and Analysis Unit, 4700 Old Kingston Pike, Knoxville, TN 37923.

[2] Supervisory Forester, USDA Forest Service, Forest Inventory and Analysis Unit, 4700 Old Kingston Pike, Knoxville, TN 37923.

When EMAP cells are decomposed by a factor of 27, then cells resembling P2 cells are produced but are offset. In many cases, what the analyst desires is to give the reader an idea of the sample design of FIA and FHM.

A related objective is to summarize FIA plot attributes within hexagon cells and spatially portray these summaries (e.g., choropleth maps of forest proportion, forest volume). The objective of this study is to create a spatially dispersed grid adequate for public presentation purposes that simulates the FIA sample design.

Methods

McCollum *et al.* (2008) show how to decompose the P2 grid to produce a P1 grid used in double sampling for stratification when strata weights are estimated. In similar fashion, we decomposed the EMAP grid by a factor of 27 to produce P2-sized cells. We then used these EMAP-derived grid points to emulate FIA's P2 grid. This process involved creating a threefold decomposition of a hexagon, where it was necessary to collect the center as well as the nodes, and then constructing Thiessen polygons of those points. The resultant cells are hexagons that are one-third the size of the original hexagons. This process was repeated three times to construct hexagons that are 1/27 the size of the original hexagons. Because the EMAP grid is in the public domain, it is not necessary to keep the distance and direction of this shift confidential.

Creating the P3 grid presents some tribulations. At the initial establishment of FIA's three-phase sampling design, 1 FHM plot per 27 FIA P2 plots was present, reflecting the 27-factor enhancement of the FHM grid to create FIA's P2 grid. Thereafter, the former FHM sample intensity was supplanted by a P3 sample intensity of 1 P3 plot per 16 P2 plots. A grid was created by referencing the P3 plots as they were initially assigned by the developer of the P3 plot locations using the P2 hexagon number. Plots with coordinates previously believed to be true later were later discovered to not be inside their intended P2 hexagon. In one FIA unit (Southern Research Station-FIA), P3 plots were not reassigned unless they fell outside their intended P3 hexagons into another P3 hexagon that was already populated or a P3 hexagon that should not be populated.

In States that historically had FHM plots, the resulting ratio ended with 1 P3 plot per 16.2 P2 plots. In the remaining States, the ratio was 1 P3 plot per 16.0 P2 plots. Details may be found in McCollum and Cochran (2005). In the historic FHM States, P2 cells were spatially combined into groups of nine, and five-ninths of the groups were actually filled with a P3 plot. In the remaining States, P2 cells were combined into groups of 16, and each cell was filled with a P3 plot.

The procedure for allocating plots in the new States is more consistent with the Goodchild criteria, cited by Carr *et al.* (1997), Clarke (2002), Kimerling *et al.* (1999), and Song *et al.* (2002). One of the Goodchild criteria is that points be equidistant. Under the rule where the States had previously measured FHM plots. this criterion is not met. For example, there is a short distance from one point to the next in three directions, and a long distance from that same point to the next point in three other directions. Another criterion stated by Goodchild (1994) is that each area contains only one point. Where cells are defined as the historic FHM cells, some cells contain as many as three points while others contain only one and some are empty. Alternatively, if cells are defined as the Thiessen expansion of the existing points, then at least three other Goodchild (1994) criteria are violated: areas are not equal in size or shape and edges are not equal in length. A fourth view is that the network of P3 plot locations is not a grid at all but just a network of suggested plot locations. Because this is a presentation grid, however, and presentation points do not represent actual data points, and because even to this day no standard method exists for imputing data from P3 plots to all P2 plots, we will present one selection rule rather than two, such being a 16-fold composition of the P2 grid.

Because the FHM grid is shifted with respect to the EMAP grid, the FHM-derived FIA grid also is shifted with respect to the EMAP grid. In general, State and county lines are not conformable with hexagons; slightly different numbers of grid points are present at both the P2 and P3 levels, as shown in table 1.

Table 1.—*Southern Research Station State counts of Forest Inventory and Analysis Phase 2 and Phase 3 plots from sampling and presentation grids.*

	Sampling		Presentation		Difference	
	Phase 2	Phase 3	Phase 2	Phase 3	Phase 2	Phase 3
Alabama	5,655	351	5,650	355	5	−4
Arkansas	5,732	361	5,735	361	−3	0
Florida	7,090	446	7,099	441	−9	5
Georgia	6,406	392	6,401	398	5	−6
Kentucky	4,359	276	4,358	273	1	3
Lousiana	5,591	352	5,584	352	7	0
Mississippi	5,221	327	5,223	327	−2	0
North Carolina	5,800	357	5,798	363	2	−6
Oklahoma	7,535	473	7,533	470	2	3
South Carolina	3,453	211	3,445	217	8	−6
Tennesee	4,538	282	4,545	285	−7	−3
Texas	28,955	1,807	28,955	1812	0	−5
Virginia	4,613	283	4,610	283	3	0
Southern Research Station	94,948	5,918	94,936	5,937	12	-19

The presentation grid has already attracted a market. Because perturbed and swapped locations were not yet available for all of Florida, the presentation grid points simulated approximate plot locations rather well in Brown (2007). These presentation grids meet analysts' needs for portraying FIA's P2 and P3 sample design without releasing FIA's actual sampling grids. The next logical step after creating a presentation grid is to assign plots to grid points such that an analyst can portray actual plot intensity and corresponding attributes. This process required the use of the "assignment problem" algorithm, discussed by Edmonds (1965) and Gabow (1976) and then implemented by Rothberg (1985) in the public domain programs MATCH and WMATCH, as well as the "minimum cost network flow problem" discussed by Kennington and Helgason (1980) and implemented by them in the public domain program NETFLO in 1988.

Case Study

To see if we could develop methods for assigning plots to grid points, we started with Georgia because it has the smallest counties of any State in the South, and thus finding a solution would be the most difficult: a higher proportion of plots would cross county boundaries when plots are assigned to presentation grid points that are offset from the original sampling grid points.

The first step involved assigning P3 plots to P3 cells. Desirable features of any perturbing and swapping routine are that plots remain in their counties and that source plots remain close to their target locations. Ideally, the target locations would be ecologically similar as well in terms of land use, forest type, and productivity class (Lister *et al.* 2005). Other than attempting to match census water status, we did not consider these attributes.

Operations research analysts know this problem as the "assignment problem," or the "bipartite (weighted) matching problem." This problem is often described as assigning agents to jobs, and, if the cost of assigning agent i to job j is C_{ij}, what is the optimal assignment?

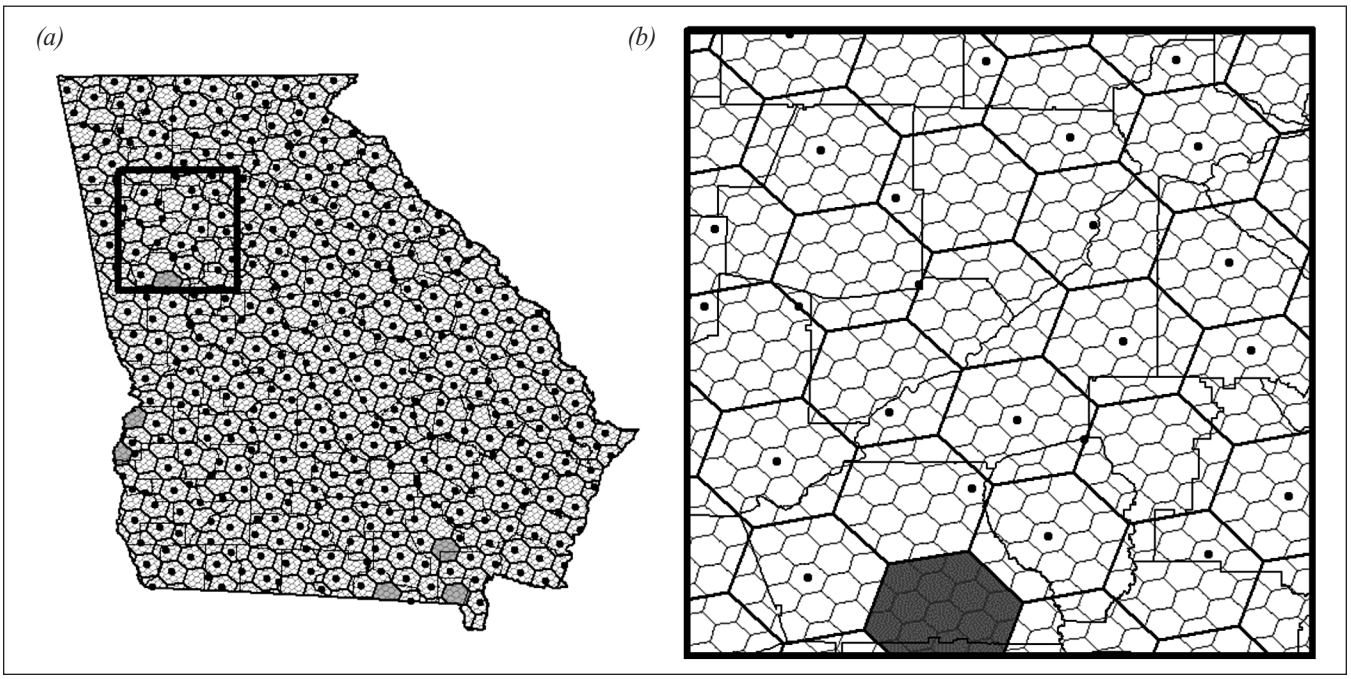
This problem was worked on by Edmonds (1965) and then Gabow (1976) and then operationalized by Rothberg (1985), who wrote C programs called WMATCH for weighted matching and MATCH for unweighted matching. The WMATCH program was used in several stages of our assignment because it was necessary to assign plots to perturbed locations under certain constraints, namely keeping plots in the same county, in unique cells, and reasonably close to their perturbed locations.

Thus, the authors constructed a mathematical graph wherein one type of node was the plots and the other type was the P3 cells. Edges were constructed if the county of the source plot matched the county of the target cell or the county of any remnant polygon within the target cell. The cost was related to the distance between the source plot and the target center, plus a small penalty if census water status did not match, plus a random component designed to prevent unswapping. This graph was input into the WMATCH program, which produced a solution shown in figure 1a, with a detailed view shown in figure 1b. The six P3 cells that could not be matched for Georgia in table 1 are highlighted in figure 1a. In figure 1b, one can see that each P3 cell is composed of 16 P2 cells; actually, 13 whole P2 cells and 6 half cells are present.

The next step involved assigning P2 plots into hexagons that were entirely within one county. Performing this match reduces the dimension of the problem. The results are shown in figure 2a, with a detailed view shown in figure 2b. As in figure 1, colored hexagons are unassigned, and assigned grid points are shown by black dots. Unassigned grid points are shown with empty circles.

A number of unassigned hexes and unassigned plots remain. They could be assigned by trial and error within correct counties, but a better method is available. This method involves consideration of the "minimum cost network flow" problem. This problem involves nodes of supply (where the initial value of the node is the number of available units, recorded as a positive number—one could think of it as the number of agents at this location), nodes of demand (where the initial value of the node is the number of units which are desired, recorded as a negative number—one could think of it as the number of jobs at this location), and possibly intermediate nodes whose initial value is 0. The nodes are connected by edges, which carry a minimum number of units (possibly 0) and a maximum number of units (possibly unbounded), at a specified cost per unit flow. The problem is to transfer units such that the final value of each node is 0. It may be necessary to introduce dummy nodes into the graph. This problem was discussed by Kennington and Helgason (1980), who later wrote a Fortran program called NETFLO (Kennington and Helgason 1988) that was released to the public domain.

Figure 1.—*Phase 3 assignment for using the presentation grids for Georgia and (b) a detailed view.*

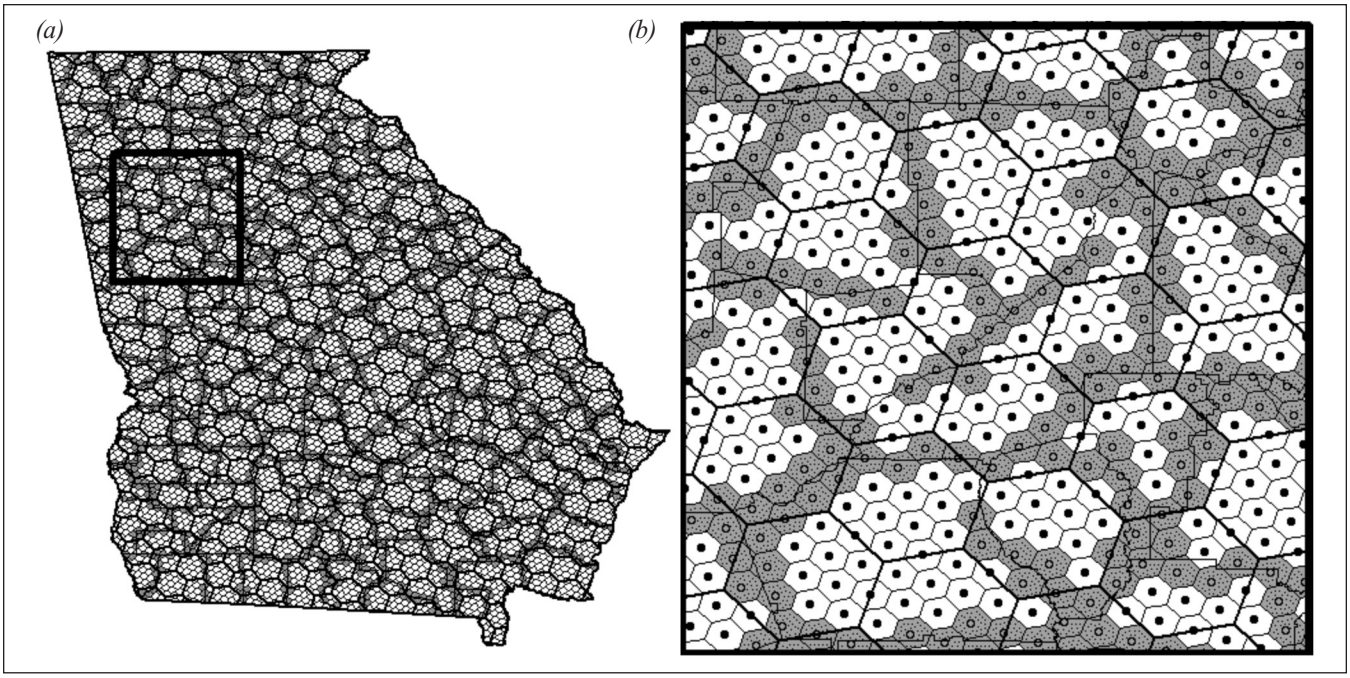


Initially, hexagons belong to the county where their center resides, but a desirable feature of any perturbing and swapping scheme is that the perturbed and swapped location be in the same county as the plot. These features are achieved for plot perturbing and swapping under FIA's current methodology. In terms of the problem, each county was a node, and the initial value was the number of plots not yet assigned minus the number of hexagons not yet assigned. An artificial extra county was introduced that bordered every other county. Its initial value was the number of plots minus the number of target centers; otherwise, edges were constructed between nodes if one county bordered another. The minimum flow was 0, and the maximum flow was the number of hexagon centers along that county border that could be flipped. With the extra county, the maximum flow was 1 if extra plots were present, no more than one could come from any one county, and if defects were present; again, no more than one could be placed in any one county.

The cost of flipping a suggested plot location from a hexagon center to a remnant polygon was inversely related to the size of the remnant polygon plus a severe penalty if the hexagon was shared by three or more counties. The programmatic bookkeeping of flipping a hexagon from county A to county C was much more difficult if it was possible that the hexagon had already been flipped to county B. Thus, such hexagons were not flipped unless it was absolutely necessary, and it was not.

Once it was determined which hexagons should be flipped, it was possible to assign the remaining plots to the remaining plots with WMATCH. This step left five plots unassigned. These plots were assigned to hexagons that had already been assigned. The authors remind readers that, in the days of the 100-second rule (Guldin *et al.* 2006), it was common for multiple plots to be assigned to one location; a Geographic Information System analyst would know to jitter the points (add or subtract a small amount to points that were assigned the same coordinate) before actually making a map. In the event that fewer source points than target points are present, temporary plots could be added that would carry no acres or would not contain any trees but would be merely a forest/nonforest call that would most closely match the actual acres surveyed to be forest with the number of acres that appear to be forest on the map. For example, if a county were more than 50 percent forested and it had 99 plots but 100 hexagons, the extra hexagon would be colored forest if the county itself had 49 or fewer forested plots; if not, it would be colored nonforest.

Figure 2.—*First stage of Phase 2 assignment using the presentation grids for (a) Georgia and (b) a detailed view.*



Discussion and Conclusions

We believe that this method simulates the actual sample design rather well, and, in view of the fact that analysts want to explain the hexagon-based sample design, this grid enables that possibility. We further claim a grid such as this one comes closest to enabling the release of microdata while nearly maximizing the uncertainty of plot locations. The circumscribed radius of a P2 hex is 3,040 m (1.89 mi). Thus, the maximum distance between the true location of a P2 plot and its assigned presentation grid point is 3,040 m (1.89 mi) for those P2 plots occurring within the same presentation hexagons as their assigned presentation grid points. Deviations may be even greater for P2 plots assigned to presentation grid points of presentation hexagons for which the P2 plot locations are outside. These deviations—effectively, perturbing distances—are substantially greater than currently allowed under FIA policy: plot locations may be perturbed up to 1 mi, but many perturbed locations are within 0.5 mi of true plot location coordinates. Indeed, target locations are not tied to actual plot locations at all, unlike with the current method. The P3 signal has been filtered out of the target locations.

Because target locations are not tied to source locations, unfuzzing the data becomes much harder. The proposed method provides an alternative that would be technically feasible if FIA policies were to change. Because target locations are not tied to source locations, however, the presentation grid approach may be less useful than existing public plot location coordinates for remote sensing applications and for geospatial analyses that require the association of FIA plots with specific spatial features. On the other hand, each plot is in its own hexagon, and, if the presentation grid is released, cell-based modeling is enabled.

Potentially, we could consider P1 data when assigning plots to cells. Thus, forested plots would be assigned where forest is most likely to occur.

This method leads to equally sized cells. One could consider a weighted Thiessen method, with the weights based on the current expansion factor (Miles *et al.* 2001) of the plot. Such a method would not guarantee that the polygon area would be equal to the expansion factor. One can imagine trying to construct a jigsaw puzzle of the State such that the size of each piece is equal to the expansion factor of each plot, and the perturbed plot location is on the piece. Ideally, the perturbed plot location should not change over time, although the expansion factors will. The authors submit that maximally dispersed points would allow for maximum flexibility in constructing that puzzle.

Literature Cited

Brown, M.J. 2007. Florida's forests–2005 update. Asheville, NC: U.S. Department of Agriculture Resour. Bull. SRS-118. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 39 p.

Carr, D.B.; Kahn, R.; Sahr, K.; Olsen, A.R. 1997. ISEA discrete global grids. Statistical computing & graphics newsletter. 8(2): 31-39.

Clarke, K.C. 2002. Criteria and measures for the comparison of global geocoding systems. In: Goodchild, M.; Kimerling, A.J., eds. Discrete global grids. Web book. Santa Barbara, CA: University of California, Santa Barbara, National Center for Geographic Information & Analysis. http://www.ncgia.ucsb.edu/globalgrids-book. (22 August 2006).

Edmonds, J. 1965. Paths, trees, and flowers. Canadian Journal of Mathematics. 17: 449-467.

Gabow, H.N. 1976. An efficient implementation of Edmonds' algorithm for maximum matching on graphs. Journal of the ACM. 23(2): 221-234.

Guldin, R.W.; King, S.L.; Scott, C.T. 2006. Vision for the future of FIA: paean to progress, possibilities, and partners. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C.; McWilliams, W.H., eds. Proceedings, sixth annual forest inventory and analysis symposium. Gen. Tech. Rep. WO-70. Washington, DC: U.S. Department of Agriculture, Forest Service. 126 p.

Kennington, J.L.; Helgason, R.V. 1980. Algorithms for network programming. New York: John Wiley. 291 p.

Kennington, J.L.; Helgason, R.V. 1988. The NETFLO program for solving minimum cost network flow problem. ftp://dimacs.rutgers.edu/pub/netflow/mincost/solver-1. (1 November 2006).

Kimerling, A.J.; Sahr, K.; White, D.; Song, L. 1999. Comparing geometrical properties of global grids. Cartography and Geographic Information Systems. 26(4): 271-288.

Lister, A.; Scott, C.; King, S. *et al.* 2005. Strategies for preserving owner privacy in the national information management system of the USDA Forest Service's Forest Inventory and Analysis unit. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C.; McWilliams, W.H.; Cieszewski, C.J., eds. Proceedings, fourth annual forest inventory and analysis symposium. Gen. Tech. Rep. NC-252. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station: 163-166.

McCollum, J.M.; Cochran, J.K. 2005. Forest inventory and analysis and forest health monitoring: piecing the quilt. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C.; McWilliams, W.H., eds. Proceedings, fifth annual forest inventory and analysis symposium. Gen. Tech. Rep. WO-69. Washington, DC: U.S. Department of Agriculture, Forest Service. 222 p.

McCollum, J.M.; Cochran, J.K.; Rose, A. 2008. Discrete global grid for photointerpretation. Res. Pap. SRS-44. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 7 p.

McRoberts, R.E.; Holden, G.R.; Nelson, M.D. *et al.* 2005. Estimating and circumventing the effects of perturbing and swapping inventory plot locations. Journal of Forestry. 103(6): 275-279.

McRoberts, R.E.; Schmidt, T.L.; O'Neill, K.P. *et al.* 2002. Forest inventories in the United States. Environmetrics newsletter. 8(1): 14-17.

Miles, P.D.; Brand, G.J.; Alerich, C.L. *et al.* 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.

Overton, W.S.; White, D.; Stevens, D.L. 1990. Design report for EMAP. EPA/600/3-91/053. Corvallis, OR: Environmental Research Laboratory. 52 p.

Rothberg, E. 1985. The WMATCH program for finding a maximum-weight matching for undirected graphs. ftp://dimacs.rutgers.edu/pub/netflow/matching/weighted/solver-1. (1 September 2006).

Song, L.; Kimerling, A.J.; Sahr, K. 2002. Developing an equal area global grid by small circle subdivision. In: Goodchild, M.; Kimerling, A.J., eds. Discrete global grids. Web book. Santa Barbara, CA: University of California, Santa Barbara, National Center for Geographic Information & Analysis. http://www.ncgia.ucsb.edu/globalgrids-book. (22 August 2006).

U.S. Department of Agriculture (USDA) Forest Service. 2005. Program structure for forest inventory and monitoring. FIA fact sheet series. http://www.fia.fs.fed.us/library/fact-sheets/overview/Pgm_Structure_FS.pdf. (1 December 2006).

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(1): 5-22.

New Functions and Programs in Hypermap Software Development for Internet-Based Displaying of FIA Data

Chris J. Cieszewski[1], Roger C. Lowe[2], Shangbin Liu[3], Ingvar Elle[4], and Daniel Markewitz[5]

Abstract.—This article describes updates on the development of various applications for the Hypermaps applications and the newest versions of the forest inventory data display tool, Interactive Fast Online Reports and Maps (InFORM). The development of InFORM applications is cosponsored by the University of Georgia, Warnell School of Forestry and Natural Resources; the Forest Service, U.S. Department of Agriculture's Forest Inventory and Analysis (FIA) unit; Georgia Forestry Commission; and National Commission on Science for Sustainable Forestry. InFORM provides an engaging and highly interactive interface, displays spatial information rapidly, obtains basic forest inventory summary data, and creates charts and tables describing the FIA inventory. Recent updates also make InFORM highly relevant for developing markets in carbon trading. InFORM fills the gap in public communication between the layman and the Geographic Information System professional by giving entry-level online access of spatial inventory data to users who are unable to navigate through other high-end, but less friendly, applications. In addition to being used as an Internet-based system, InFORM can also be used as a set of native applications on workstations and laptops.

Background

The development of the hypermaps applications at the University of Georgia, Warnell School of Forestry and Natural Resources originated as a means of meeting the internal needs of the Fiber Supply Assessment project at the Warnell School. The first version of the web-based application was essentially a reengineered mirror of the existing Forest Service Southern Forest Inventory and Analysis (FIA) online database. The reengineered aspects of that application mainly focused on increasing application robustness and speed. Ironically, these goals were achieved by decreasing the level of program sophistication within the application, removing its dynamic updating functions, and replacing its binary executables with static HTML pages. The work was very successful and, soon after the reengineered Web pages became publicly available, they practically replaced the original Web site, which is now no longer available. Following this success, the Fiber Supply Assessment team moved on to further altering and expanding the functionality of the online FIA inventory information retrieval system by focusing on keeping online access fast and reliable. Since the beginning of this work, a series of different applications for online access of inventory data have been developed. Current applications fall into two main groups. The first group is the Uni-Select Hypermaps, which is a direct derivation of the data viewing philosophy that was represented by the original Southern FIA online database retrieval system. Applications in this group allow viewing of inventory data summaries for individual counties, while the most recent version allows for selection of different dates of the surveys. The second group consists of the Multi-Select Interactive Online Reports and Maps, which allows for data summary retrievals for multiple counties, either by arbitrary selection or by various selection

[1] Associate Professor, University of Georgia, Warnell School of Forestry and Natural Resources, Athens, GA 30602.
[2] Geographic Information Systems Analyst, University of Georgia, Warnell School of Forestry and Natural Resources, Athens, GA 30602.
[3] Graduate Student, University of Georgia, Warnell School of Forestry and Natural Resources, Athens, GA 30602.
[4] Research Professional, University of Georgia, Warnell School of Forestry and Natural Resources, Athens, GA 30602.
[5] Associate Professor, University of Georgia, Warnell School of Forestry and Natural Resources, Athens, GA 30602.

tools, such as radiuses or polygons. Two examples of recent developments in these two categories of online FIA retrieval systems are illustrated below. First described are the features of the Uni-Select applications, which allow selection of FIA data at the county level for various years. Second, the Carbon InFORM application, which was developed several years after the Multi-Select program and which uses a different main interface and a much broader set of tools for selecting county-level FIA data, is described. Carbon InFORM is a response to public interest in greenhouse gases and developing markets in C trading; FIA data is made available on a tons of C basis.

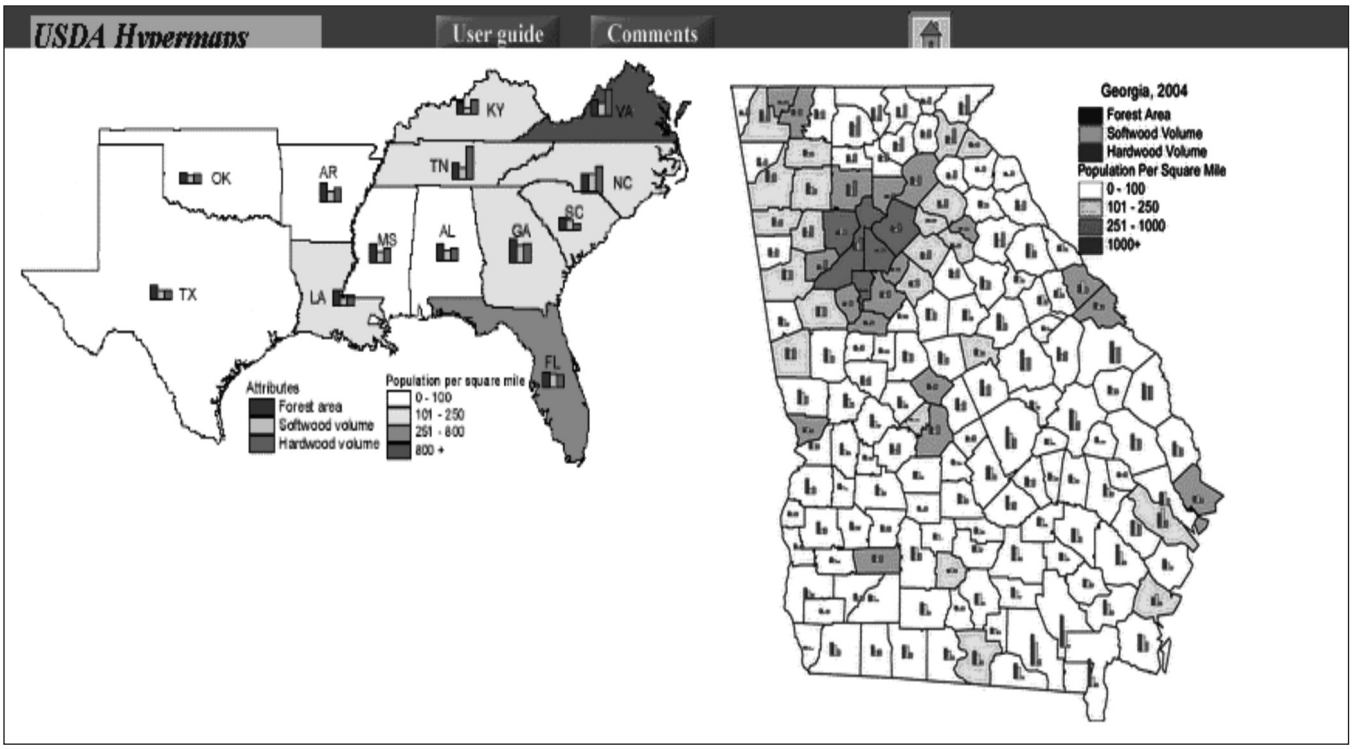
Uni-Select Hypermaps Project

The current Uni-Select Hypermaps allow for FIA data summary retrievals for various years of FIA surveys. The applications are available for all 13 Southern Region States and provide the following functionalities:

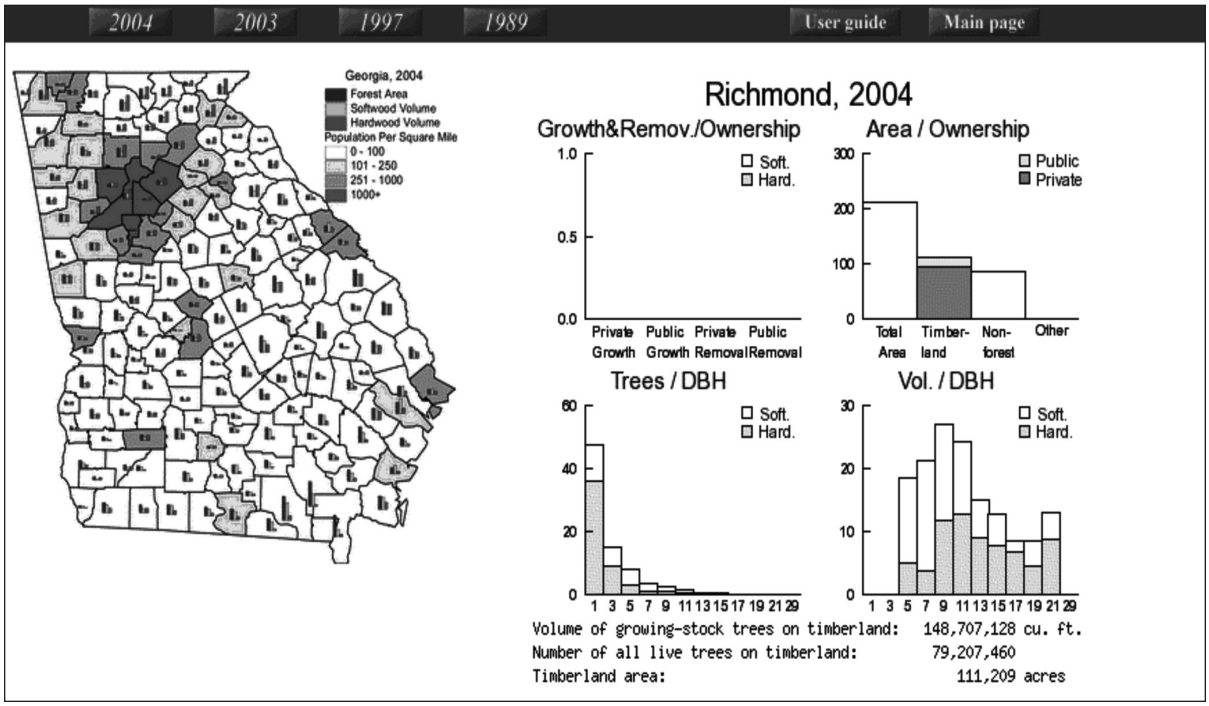
1. Selecting different counties visually by mouse-over function using each State map.
2. Creating forestry-data charts indicating wood availability by properties such as growth and removal by ownership, area by ownership, tree number by diameter class, and volume by diameter class.
3. Generating the 26 standard tables of FIA Mapmaker.

Unlike the FIA Mapmaker, which dynamically generates forest statistics based on FIA permanent sample plot data, the hypermaps provide a fast display of forest statistics at the county-scale level by accessing static pages that are pre-processed and stored in a MySQL database. The following figures explain the general usage of the software by illustrating the necessary steps in using the software. Computer screen shots are used to show the visual menus and illustrate the different steps in using the program.

Step 1. On the left-side navigation map below, move the cursor over the State to see its overlay and related information and click. For each State, a different map overlay containing the county-level information, such as forest area, hardwood and softwood volume, and population density, appears on the right.

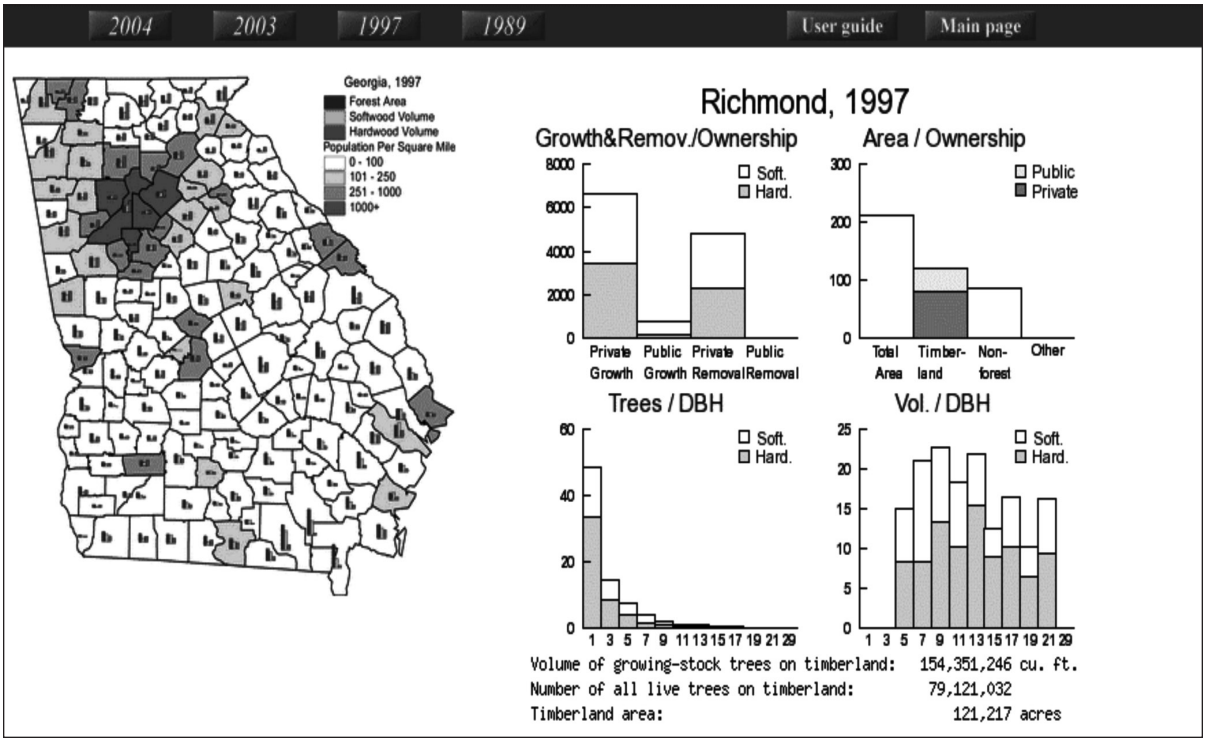


Step 2. In the image below, the default map overlay (the most recent report year) shows on the left. On the right of the screen, thumbnail images of forestry charts will appear for any county over which the cursor is placed. To have the thumbnails appear in a full-page view, click on the county.

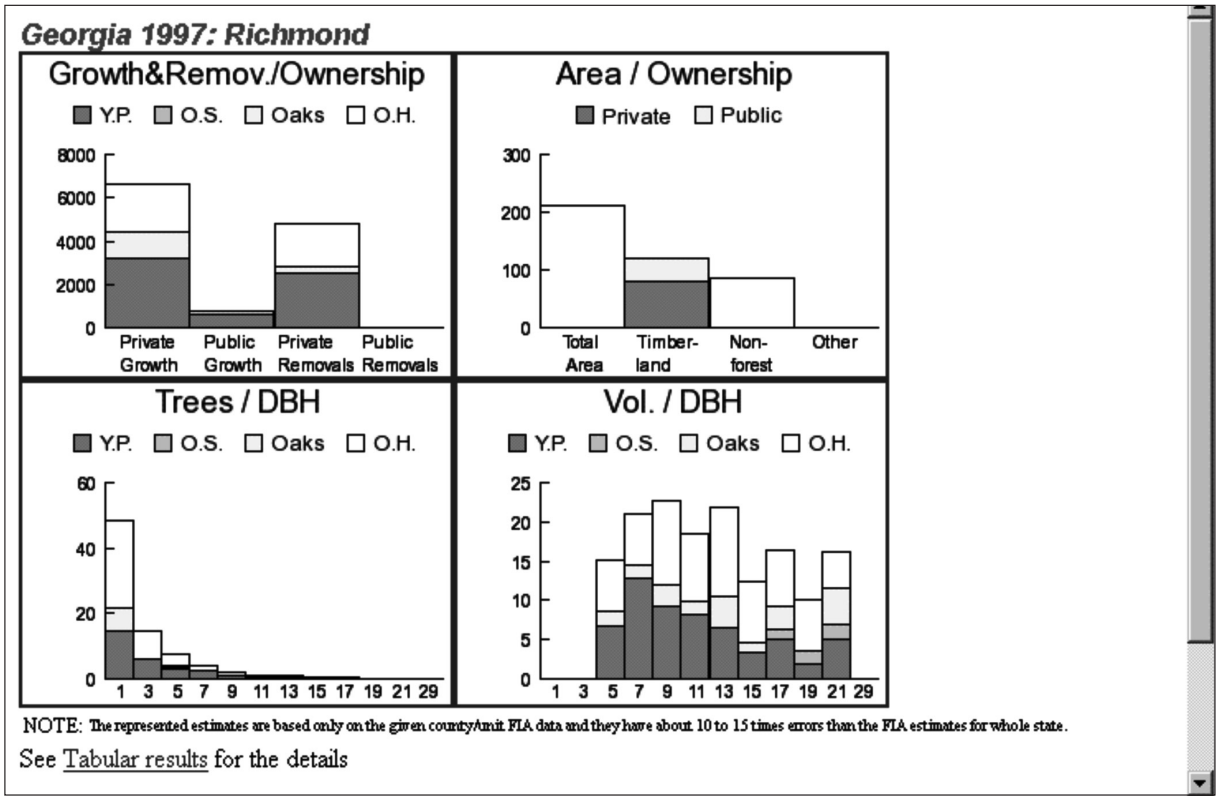


To view a different report year, move the mouse to the top of the screen, click the desired year, and then click on the desired county.

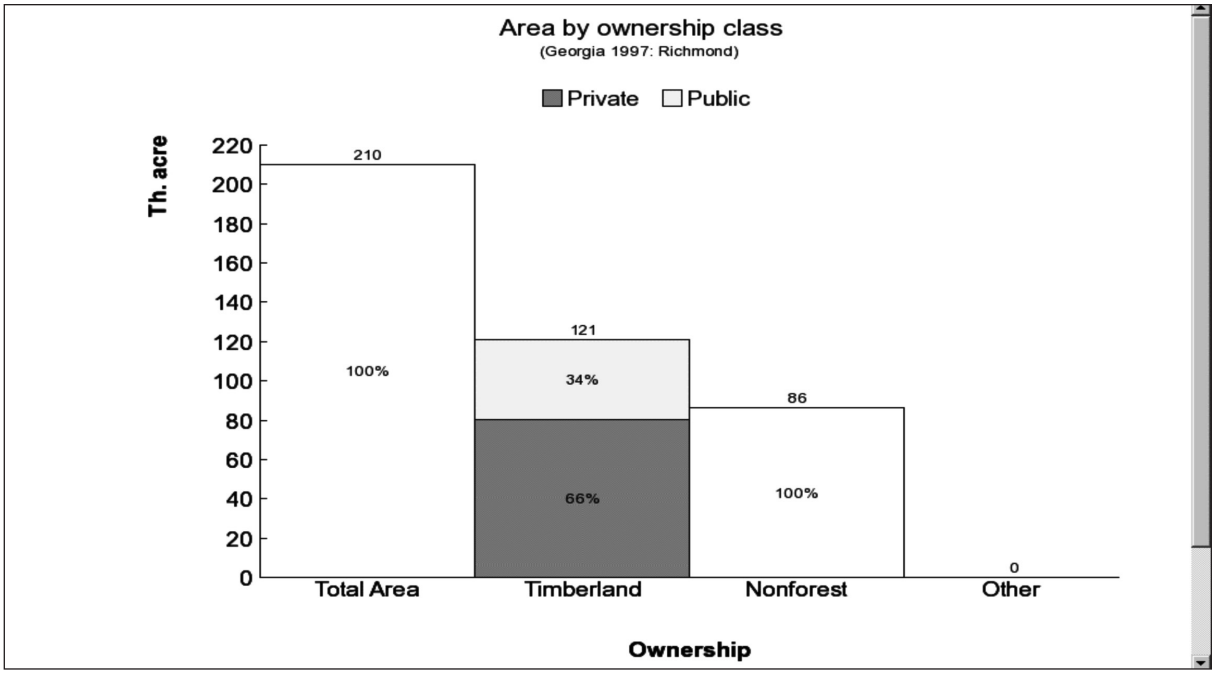
Step 3. After clicking on the desired county, a four-panel display of data charts will show (see image below).



Step 4. Click any chart from the four-panel display to view the full-size version of the chart. (See the Area/Ownership example below.)



Step 5. In the four-panel display screen (shown in Step 3), click the Tabular results hyperlink to display the tabular information of the county.



Click the Table of Contents hyperlink to unfold a list of all tables for that particular county. (Click the Table of Contents hyperlink again to close the county list.)

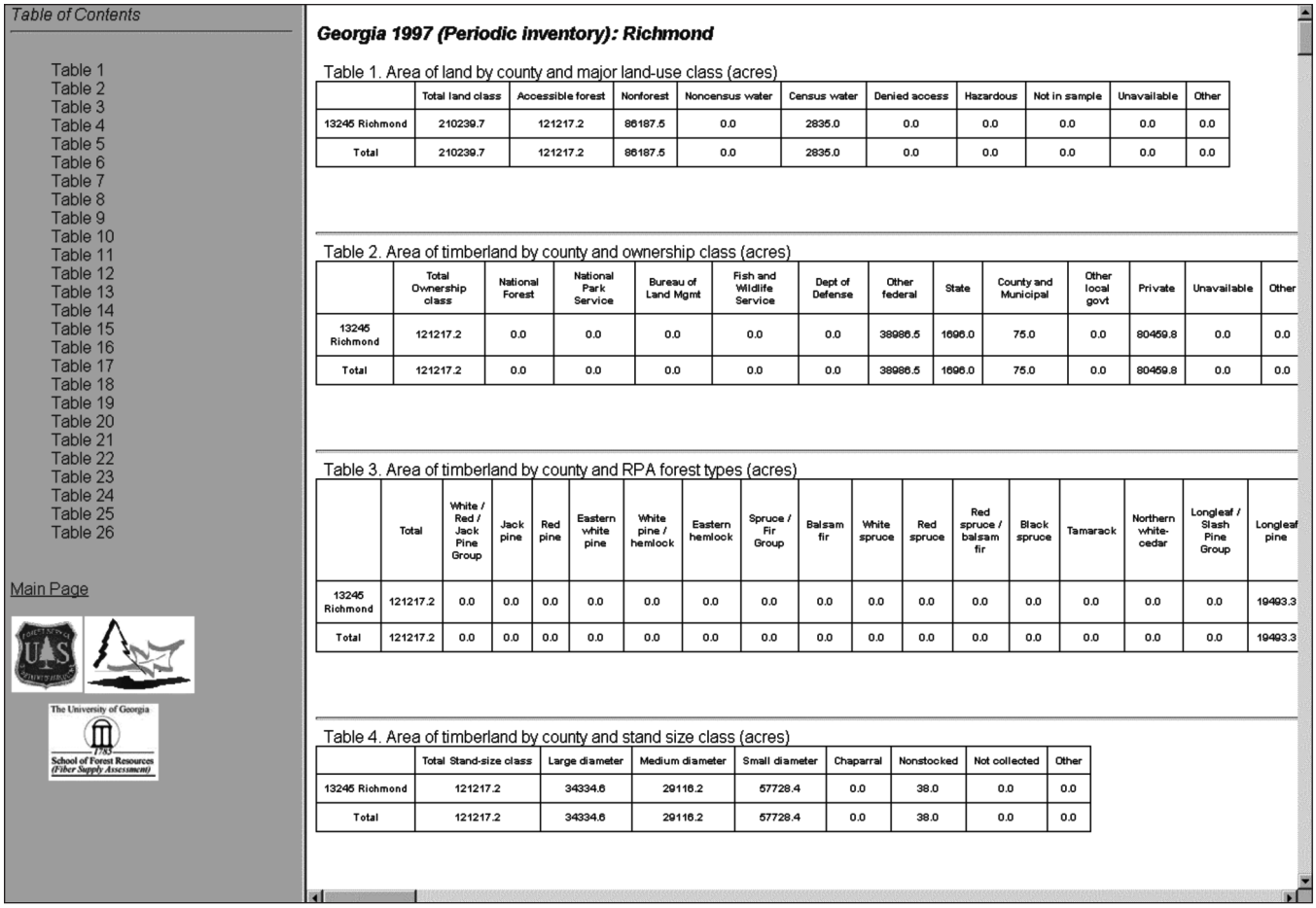


Table of Contents

Table 1
Table 2
Table 3
Table 4
Table 5
Table 6
Table 7
Table 8
Table 9
Table 10
Table 11
Table 12
Table 13
Table 14
Table 15
Table 16
Table 17
Table 18
Table 19
Table 20
Table 21
Table 22
Table 23
Table 24
Table 25
Table 26

Main Page

The University of Georgia
1785
School of Forest Resources
(Fiber Supply Assessment)

Georgia 1997 (Periodic inventory): Richmond

Table 1. Area of land by county and major land-use class (acres)

	Total land class	Accessible forest	Nonforest	Noncensus water	Census water	Denied access	Hazardous	Not in sample	Unavailable	Other
13245 Richmond	210239.7	121217.2	86187.5	0.0	2835.0	0.0	0.0	0.0	0.0	0.0
Total	210239.7	121217.2	86187.5	0.0	2835.0	0.0	0.0	0.0	0.0	0.0

Table 2. Area of timberland by county and ownership class (acres)

	Total Ownership class	National Forest	National Park Service	Bureau of Land Mgmt	Fish and Wildlife Service	Dept of Defense	Other federal	State	County and Municipal	Other local govt	Private	Unavailable	Other
13245 Richmond	121217.2	0.0	0.0	0.0	0.0	0.0	38986.5	1696.0	75.0	0.0	80459.8	0.0	0.0
Total	121217.2	0.0	0.0	0.0	0.0	0.0	38986.5	1696.0	75.0	0.0	80459.8	0.0	0.0

Table 3. Area of timberland by county and RPA forest types (acres)

	Total	White / Red / Jack Pine Group	Jack pine	Red pine	Eastern white pine	White pine / hemlock	Eastern hemlock	Spruce / Fir Group	Balsam fir	White spruce	Red spruce	Red spruce / balsam fir	Black spruce	Tamarack	Northern white-cedar	Longleaf / Slash Pine Group	Longleaf pine
13245 Richmond	121217.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	19493.3
Total	121217.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	19493.3

Table 4. Area of timberland by county and stand size class (acres)

	Total Stand-size class	Large diameter	Medium diameter	Small diameter	Chaparral	Nonstocked	Not collected	Other
13245 Richmond	121217.2	34334.6	29116.2	57728.4	0.0	38.0	0.0	0.0
Total	121217.2	34334.6	29116.2	57728.4	0.0	38.0	0.0	0.0

Clicking a specific table also unfolds the corresponding table description and displays the tabular information.

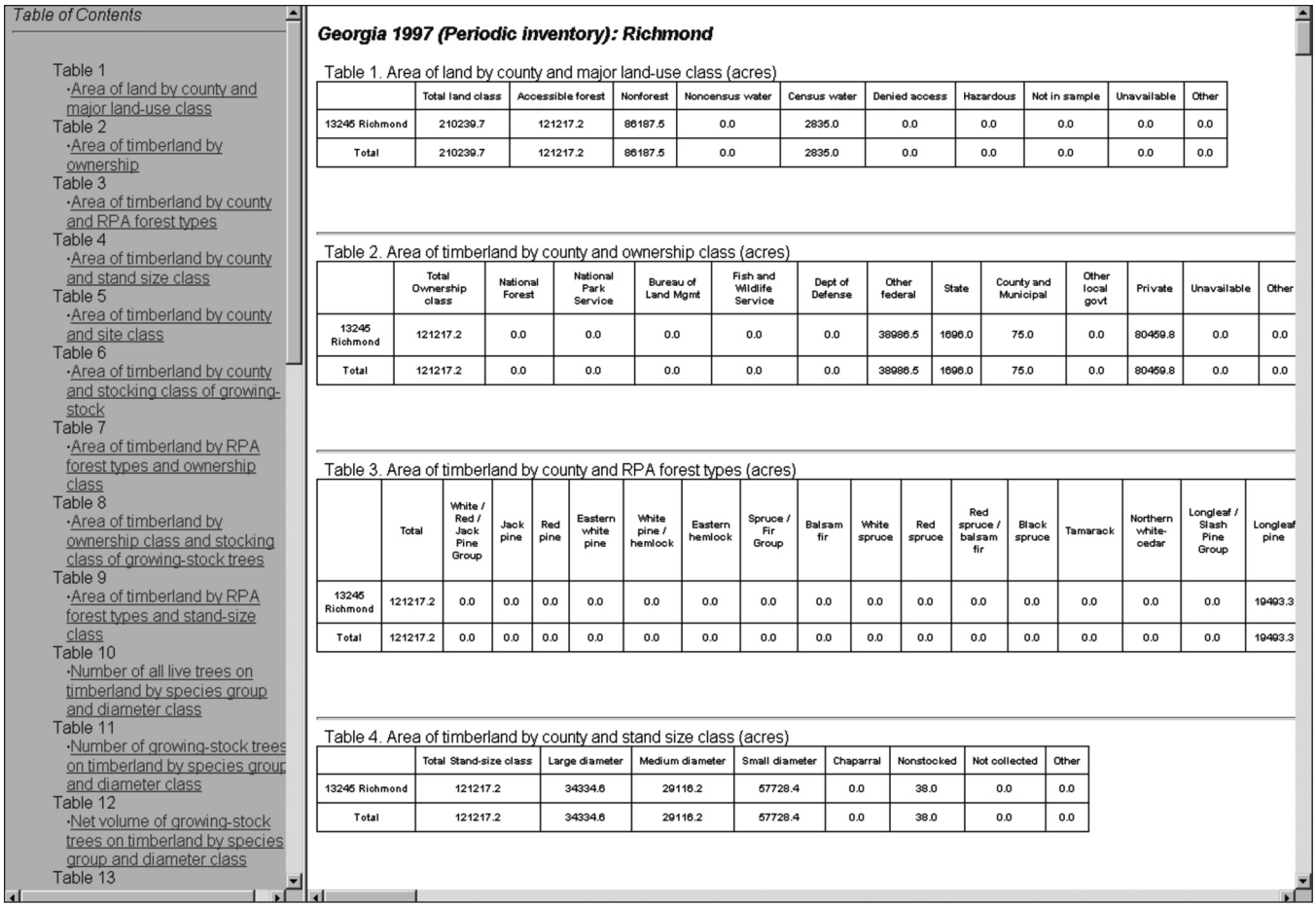


Table of Contents

Table 1
·Area of land by county and major land-use class
Table 2
·Area of timberland by ownership
Table 3
·Area of timberland by county and RPA forest types
Table 4
·Area of timberland by county and stand size class
Table 5
·Area of timberland by county and site class
Table 6
·Area of timberland by county and stocking class of growing-stock
Table 7
·Area of timberland by RPA forest types and ownership class
Table 8
·Area of timberland by ownership class and stocking class of growing-stock trees
Table 9
·Area of timberland by RPA forest types and stand-size class
Table 10
·Number of all live trees on timberland by species group and diameter class
Table 11
·Number of growing-stock trees on timberland by species group and diameter class
Table 12
·Net volume of growing-stock trees on timberland by species group and diameter class
Table 13

Georgia 1997 (Periodic inventory): Richmond

Table 1. Area of land by county and major land-use class (acres)

	Total land class	Accessible forest	Nonforest	Noncensus water	Census water	Denied access	Hazardous	Not in sample	Unavailable	Other
13245 Richmond	210239.7	121217.2	86187.5	0.0	2835.0	0.0	0.0	0.0	0.0	0.0
Total	210239.7	121217.2	86187.5	0.0	2835.0	0.0	0.0	0.0	0.0	0.0

Table 2. Area of timberland by county and ownership class (acres)

	Total Ownership class	National Forest	National Park Service	Bureau of Land Mgmt	Fish and Wildlife Service	Dept of Defense	Other federal	State	County and Municipal	Other local govt	Private	Unavailable	Other
13245 Richmond	121217.2	0.0	0.0	0.0	0.0	0.0	38986.5	1696.0	75.0	0.0	80459.8	0.0	0.0
Total	121217.2	0.0	0.0	0.0	0.0	0.0	38986.5	1696.0	75.0	0.0	80459.8	0.0	0.0

Table 3. Area of timberland by county and RPA forest types (acres)

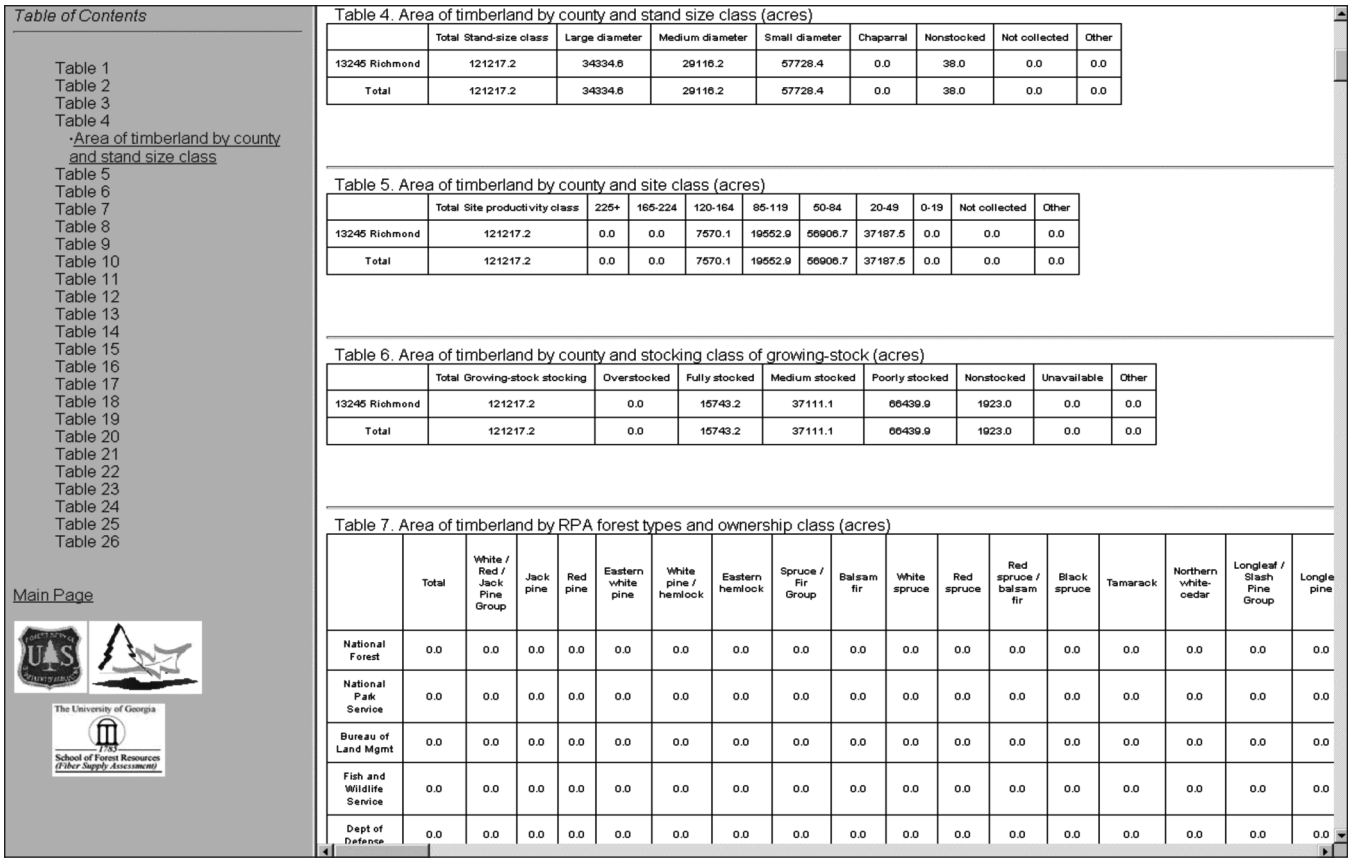
	Total	White / Red / Jack Pine Group	Jack pine	Red pine	Eastern white pine	White pine / hemlock	Eastern hemlock	Spruce / Fir Group	Balsam fir	White spruce	Red spruce	Red spruce / balsam fir	Black spruce	Tamarack	Northern white-cedar	Longleaf / Slash Pine Group	Longleaf pine
13245 Richmond	121217.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	19493.3
Total	121217.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	19493.3

Table 4. Area of timberland by county and stand size class (acres)

	Total Stand-size class	Large diameter	Medium diameter	Small diameter	Chaparral	Nonstocked	Not collected	Other
13245 Richmond	121217.2	34334.6	29116.2	57728.4	0.0	38.0	0.0	0.0
Total	121217.2	34334.6	29116.2	57728.4	0.0	38.0	0.0	0.0

Carbon InFORM: Interactive Fast Online Reports and Maps

This section describes the various functionality of the multicounty-based InFORM online software that displays FIA data summaries, which include additional processed information about forest carbon densities and availability. The Energy Information Administration has recently release guidelines for *Voluntary Reporting of Greenhouse Gases* for which the Forest Service provided input of forest C accounting. Carbon InFORM makes FIA data directly available, on a tons of C basis, to assist interested parties (e.g., States or landowners) in assessing the C balance or C sequestration potential of forest lands. Carbon was estimated from FIA data by converting green tons of biomass to dry tons of biomass by assuming that green tons of biomass is 50 percent water and 50 percent dry biomass. Then carbon is 50 percent of the dry biomass (i.e., carbon = green tons of biomass x 0.25). The Web pages display county-scale forest statistics, which helps those who are unfamiliar with the FIA datasets to quickly evaluate FIA forest inventory summaries relating to carbon, volume, number of trees, land area, and other forestry data. See the Web pages at http://tyan.growthandyield.com/multiselect/volandc/XX/state_tmplate.swf where "XX" stands for any of the Southern Region's State abbreviations (e.g., GA for Georgia). The following is a description of the Carbon InFORM application's main features and how they are used.



Displaying the Data

When the cursor is hovered over a county on the Carbon InFORM map, the county will turn blue and summary statistics for the county will display in the table in the upper right side of the screen (see fig. 1).

The information that displays will depend on what has been chosen in each of the five menus within this table. Each menu allows the choice of displaying for coniferous and deciduous trees the amount of forestland (acres), timberland (acres), volume (cubic feet), growing stock (cubic feet), carbon (tons), merchantable carbon (tons), and number of trees—all at the county level.

Directly below and to the left of this table, the amount of total land (acres) and percentage of forested land for the county is displayed. Also below the table, but to the right, is a colored bar chart representing (from left to right) the relative amounts of the forestry variable values as they appear (from top to bottom) in the table above the chart.

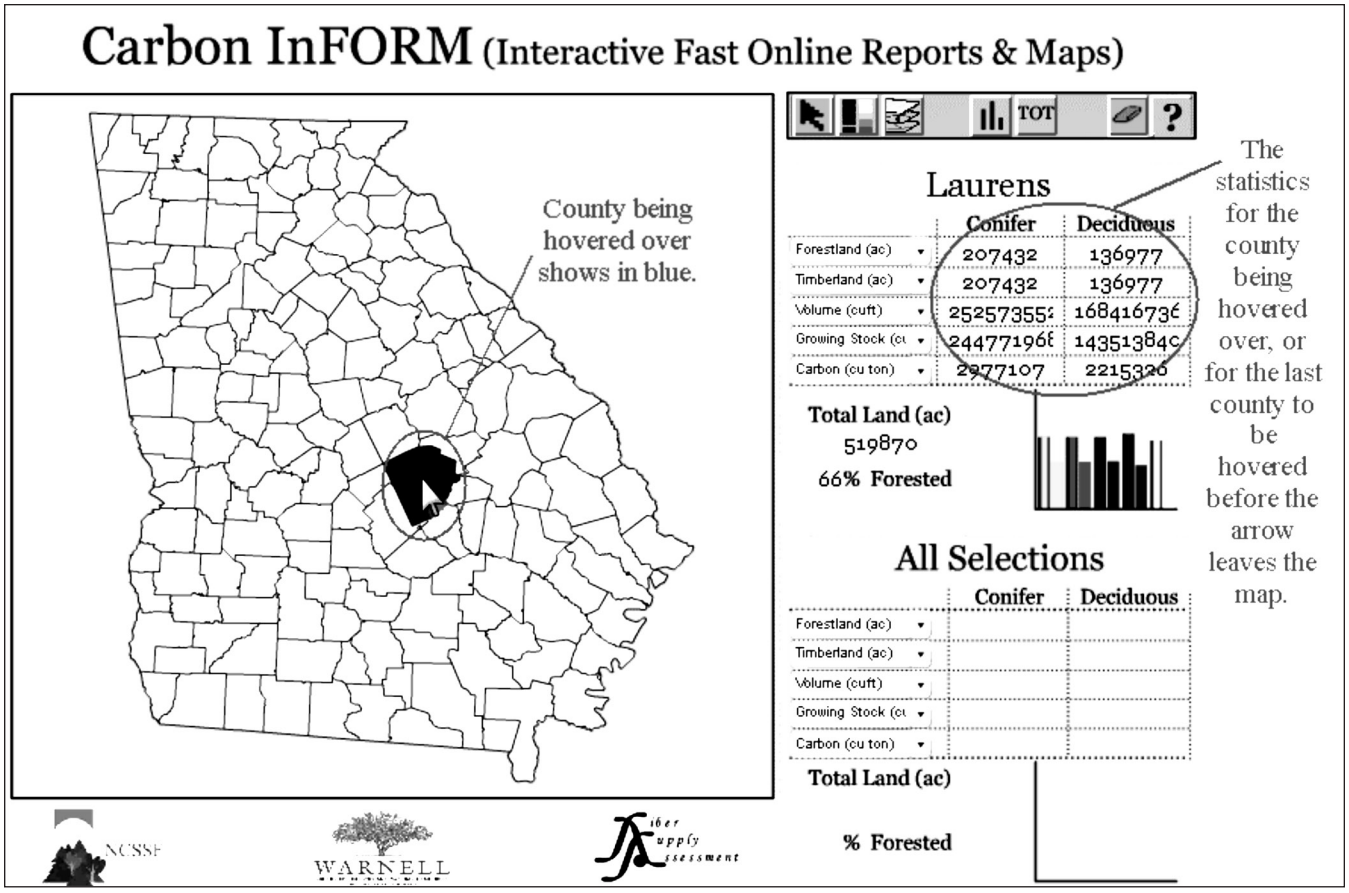
The default for menu values shows totals for the counties. To display them as per-acre values, click the TOT icon in the icon row above the table.

Selection Tools for Carbon InFORM

Displaying Data Through Single and Multiple Selection

When the cursor is hovering over a county, the data display for only that county. When hovering, the cursor must stay over the same county or the data will change. Alternatively, the selection palette (top left screen of fig. 2) contains several tools for making semipermanent selections of one or more counties. Clicking, instead of hovering, a county or counties makes semipermanent selections; once a semipermanent selection is made, the cursor can be moved off the county in order to select other menu options. The clicked county or counties are termed semipermanent because they will remain selected until clicked again. To make a semipermanent selection, click the arrow icon

Figure 1.—*Shown at the top right of the window are statistics for the county in which the cursor is hovering over. Alternatively, the data will display statistics for the county the cursor was last hovered over before the selection arrow was moved off of the map.*



in the icon row (at the top right side of the screen) or the arrow icon on the selection palette. Clicking an individual county within a State page with the arrow tool will turn the county yellow. Data for the selected county will show up in the All Selections window (fig. 2). To select multiple counties, click one county after another; the sum of all these selected counties will then display in the All Selections window (fig. 3).

Charts and Tables Relating to the Selection

Extensive data on the selected county or counties are also available in the form of reports and charts, which can be generated by clicking the Generate Charts icon in the icon row at the top right side of the main InFORM display (fig. 4). Data for the selected counties will appear in the All Selections window and in the charts and reports window (see figs. 2 through 4).

Circle, Rectangle, and Polygon Tools

The next three options on the selection palette allow a user to select single or multiple counties by circle (fig. 5), rectangle (fig. 6), or arbitrary polygon (fig. 7).

Eraser Tools

Two separate eraser icons can be used to clear the selected counties from the map. One eraser is in the icon row at the top right of the InFORM display, and the other is the last tool on the selection palette (fig. 8).

Selecting by Name

After expanding the selection palette by clicking on the gray diamond on the top left side, counties can be selected by county name (fig. 9).

Figure 2.—*After the county as been selected, the data for the county will be displayed in the All Selections section.*

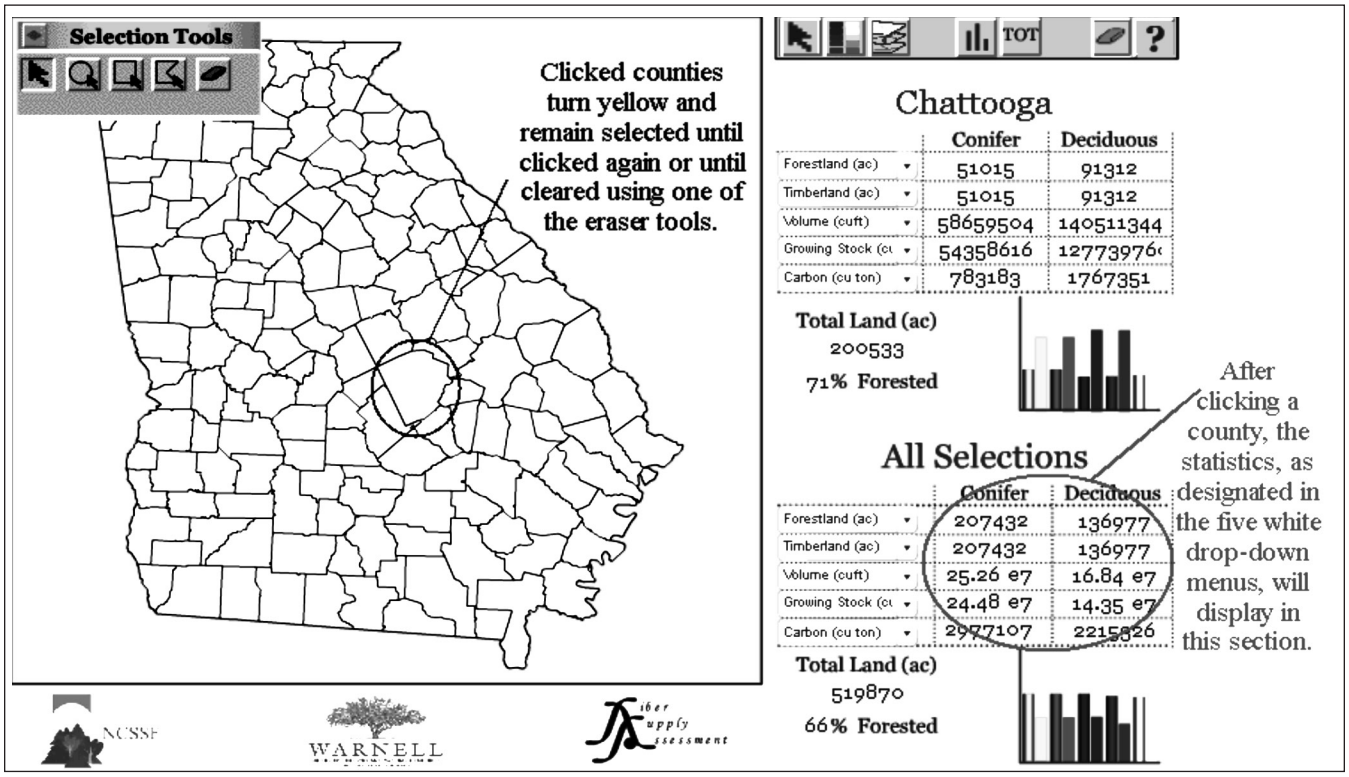


Figure 3.—*For multiple counties selected, the summed values of all the counties will be displayed in the All Selections section by chosen variable.*

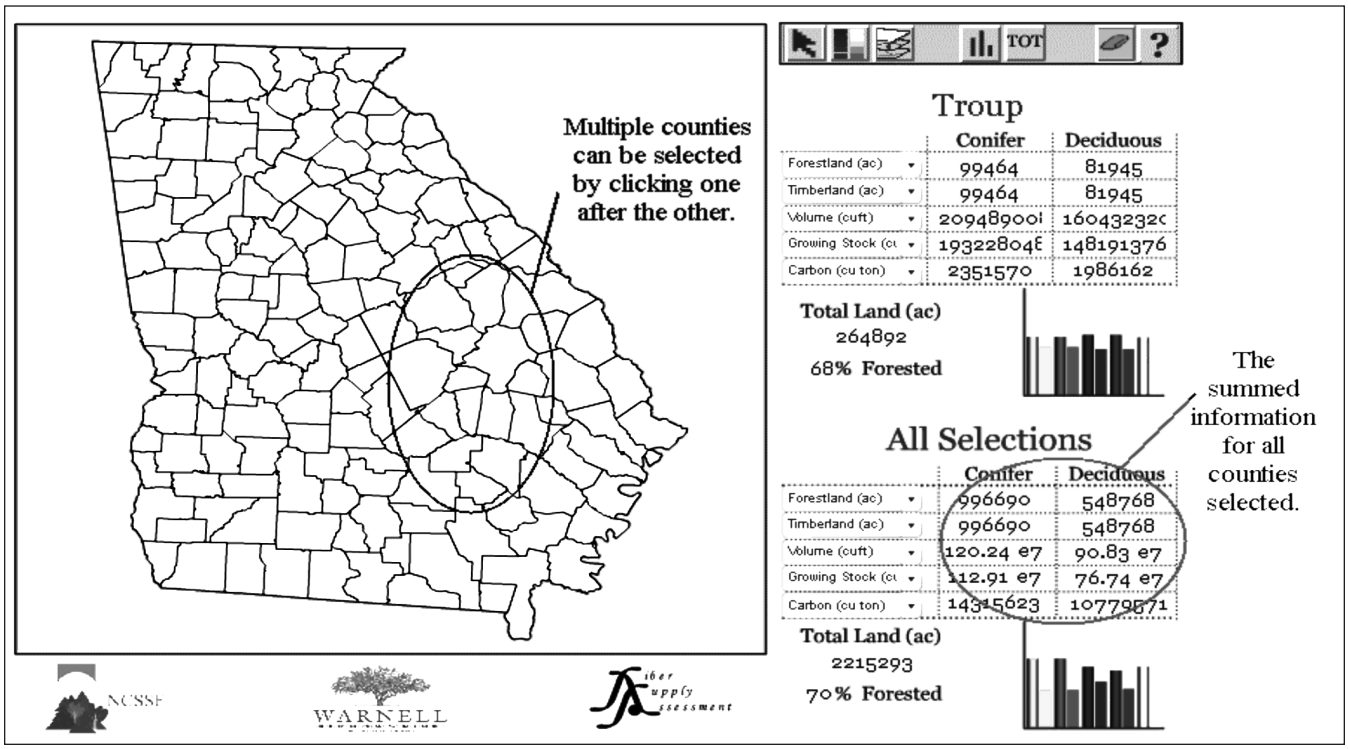


Figure 4.—*Click the Generate Charts icon to display summary charts for the selected counties. Clicking the Tabular results hyperlink below the charts will display 26 FIA tables with data for the selections.*

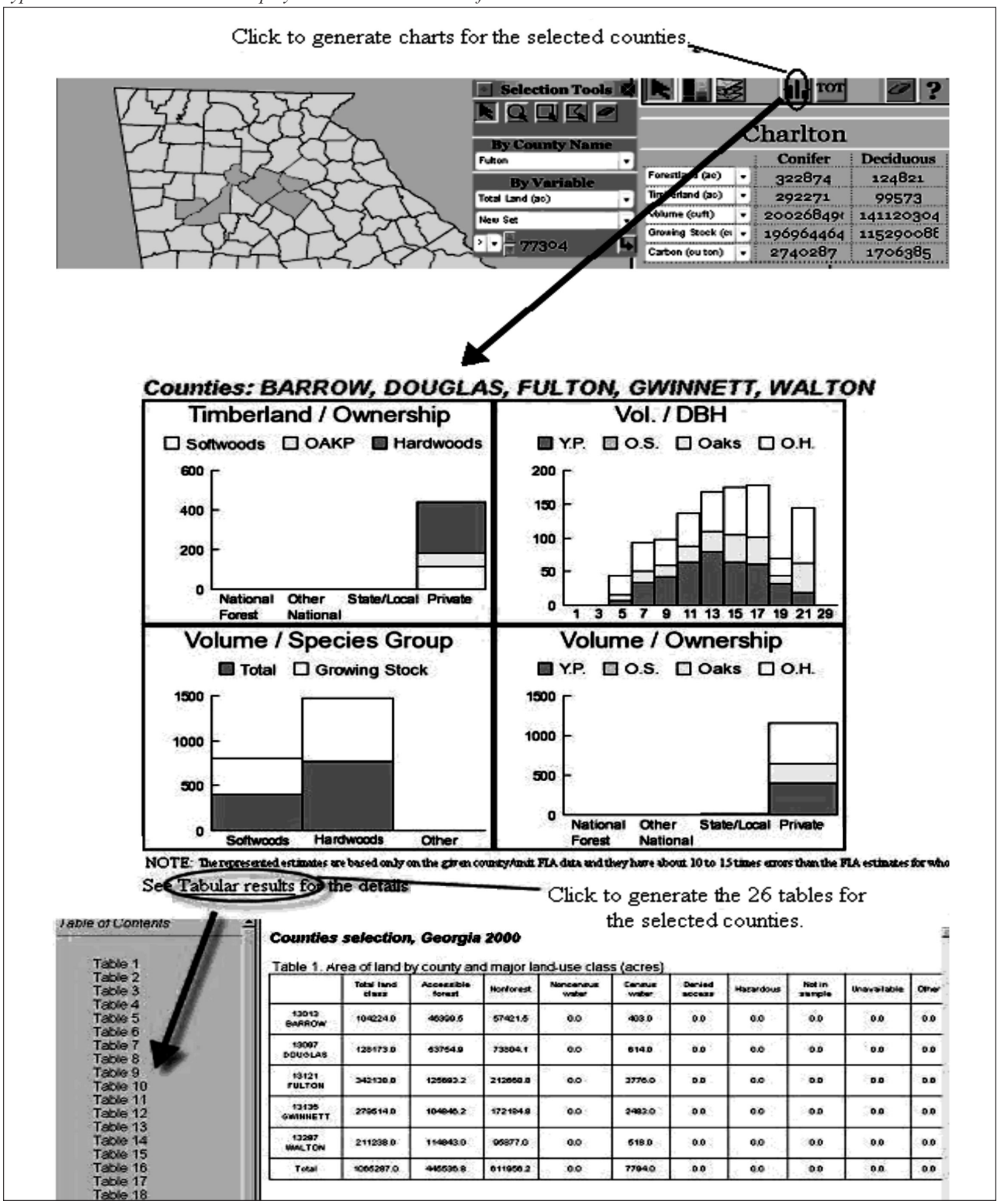


Figure 5.—*Circle tool illustrated.*

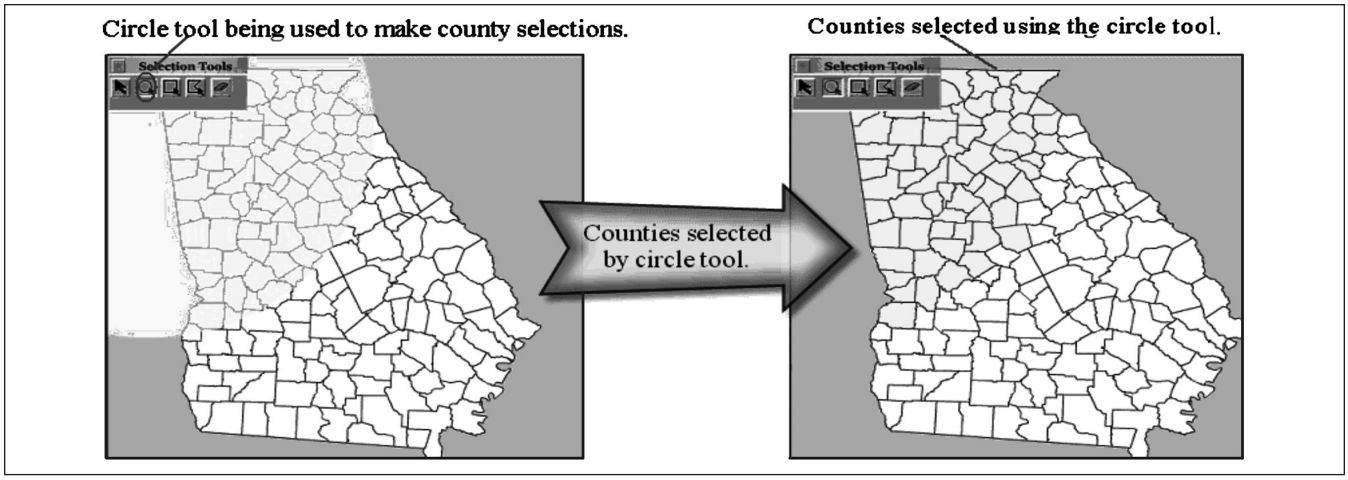


Figure 6.—*Rectangle tool illustrated.*

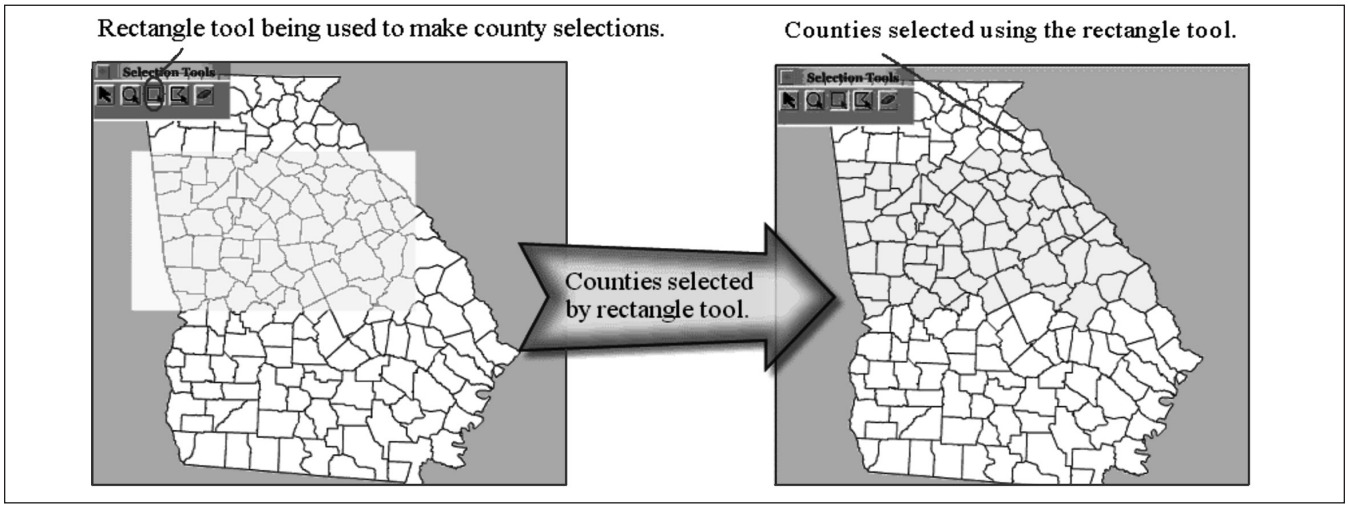


Figure 7.—*Arbitrary polygon tool illustrated.*

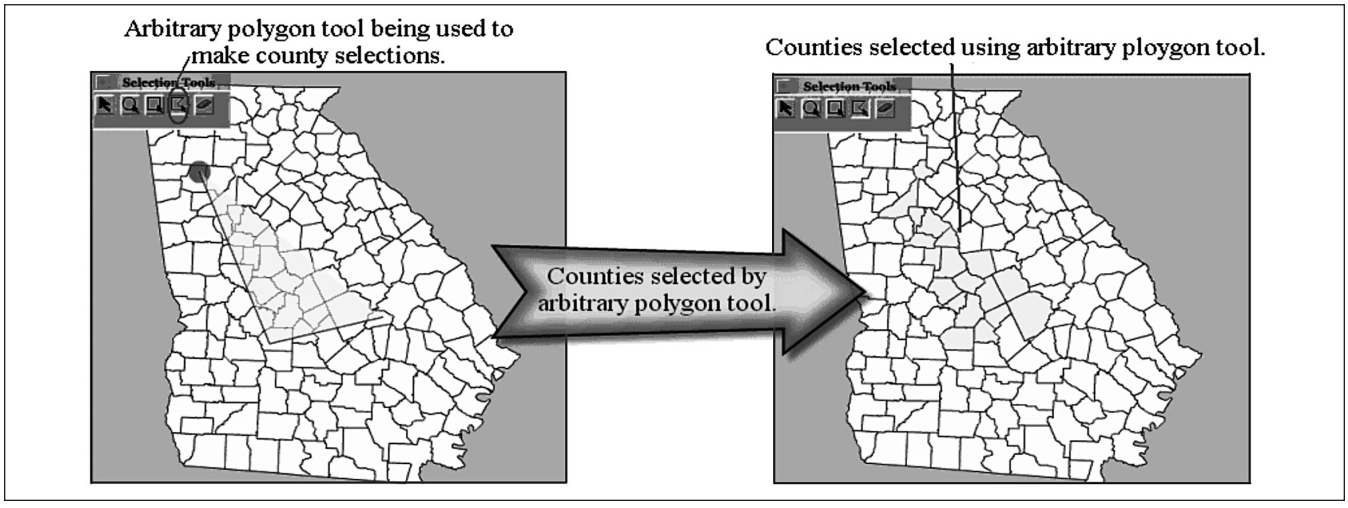


Figure 8.—*Clicking either of the eraser tools will clear all counties selected by any of the selection tools.*

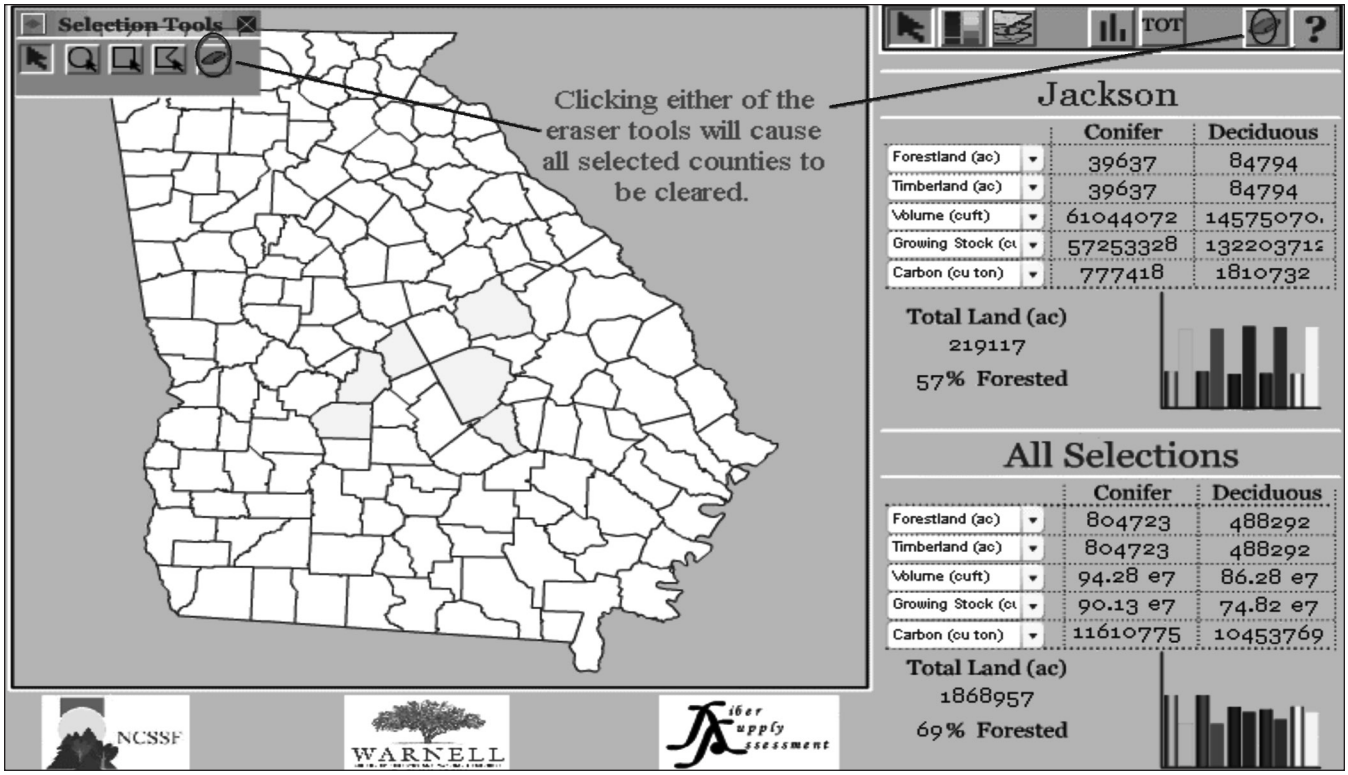
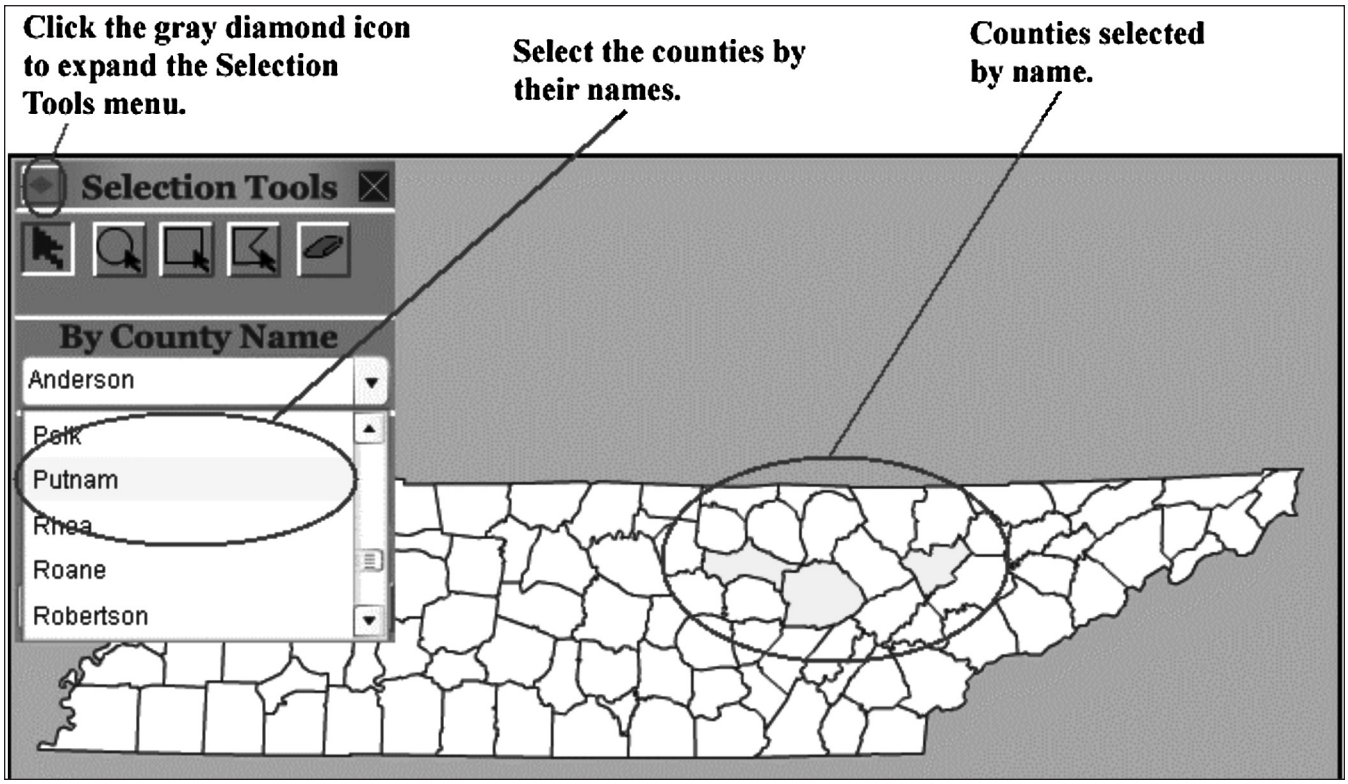


Figure 9.—*Selecting counties by name.*

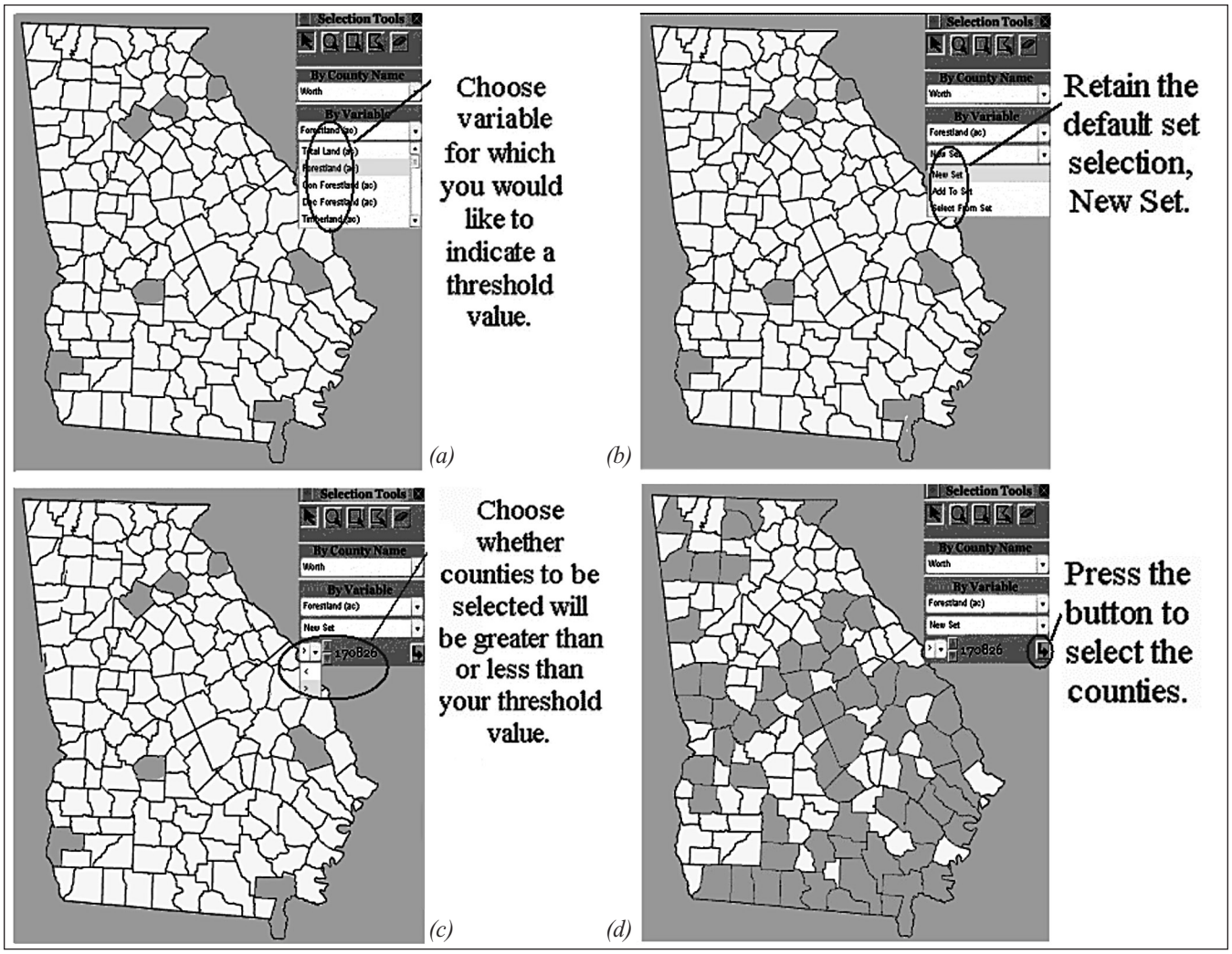


Selecting by Variable

Use the By Variable tool to select counties for which the value of a designated forestry-related variable for each county is greater than (>) or less than (<) a given threshold value. Two pop-up menus and one value range indicator exist for this task. Use the first pop-up menu to choose the variable (the default variable is Total Land) for which to indicate a threshold value (fig. 10a). The second menu allows 3 options: New Set, Add To Set, and Select From Set. These options differ by how selected counties appearing on the map are treated immediately before test invocation.

New Set. New Set, as the name suggests, will ignore any counties selected on the map at the time of test invocation and only test whether the threshold number is > or < that for each and every county in the State. To use New Set, first choose a variable from the first menu in the By Variable section (fig. 10a). Retain the default set selection under New Set (fig. 10b). Select either the > or < symbol to designate whether counties selected will be greater than or less than the threshold value (fig. 10c). Set the exact threshold value by selecting the entire number at the bottom of the expanded Selection menu, and type in the threshold value. Then click the test invocation button (fig. 10d). The map will show any counties in the State meeting the greater-than or less-than threshold value test.

Figure 10.—*Illustrated steps for finding counties with values that are > or < a threshold value for a chosen variable. Note that by choosing New Set (step b), all previously chosen counties will be cleared when the test is run (step d).*



Add To Set. Add To Set (fig. 11) will select from all unselected counties with values either greater than or less than the chosen threshold number but will also keep all previously selected counties, regardless of whether they themselves have values greater than or less than the threshold number. To use Add To Set, select the initial set of counties, choose Add To Set, select the threshold value, designate either > or <, then activate the Add To Set function by clicking the test invocation button. All counties originally selected will appear on the map, plus any previously unselected counties in the State that meet the specified threshold test.

Select From Set. At the time of test invocation, Select From Set selects only from any already selected counties whose values are greater than or less than the displayed threshold value (fig. 12). To use Select From Set, select the initial set of counties, choose Select From Set, type in the threshold value, designate either > or <, and then activate the Select From Set function by clicking the test invocation button. From the originally selected counties, Select From Set will select and display those counties meeting the threshold test.

Classify Tools for Carbon InFORM

The Classify palette in the Carbon InFORM program provides a range of options for constructing a map that illustrates for every county the relative amount of various forestry variables. Three display levels are available for the Classify palette (fig. 13).

Level 1

Clicking the multicolored Classify icon in the icon row (fig. 13) opens the Classify palette to default level 1 with the default variable being Total Land. The relative value of the variable in this field for each county can be represented on the map by a range of 3, 4, or 5 colors (fig. 14). Also, the variable in this field can be changed by clicking directly on it, which will open the Quick Classify (QC) menu (fig. 15), which allows a selection from total, deciduous and coniferous classes for forestland (acres), timberland (acres), volume (cu ft), growing stock (cu ft), carbon (tons), merchantable carbon (tons), and number of trees. To choose a variable, simply click on it in the QC menu. To choose another variable for that row, the existing variable in the menu must be unselected by clicking it again.

Figure 11.—*Selecting counties by variable using Add To Set means that initially selected counties will remain, while counties with values meeting the threshold criteria will be added.*

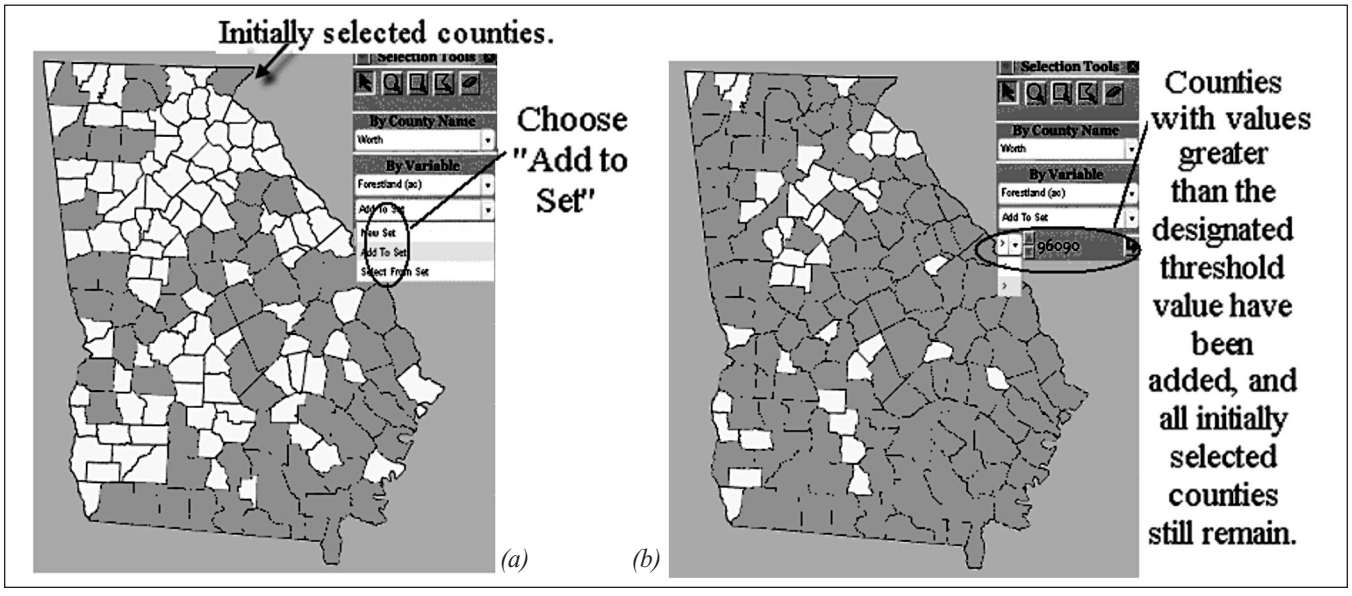


Figure 12.—*In using Select From Set, the only counties tested for the threshold criteria are the ones that are already selected at the time of test invocation.*

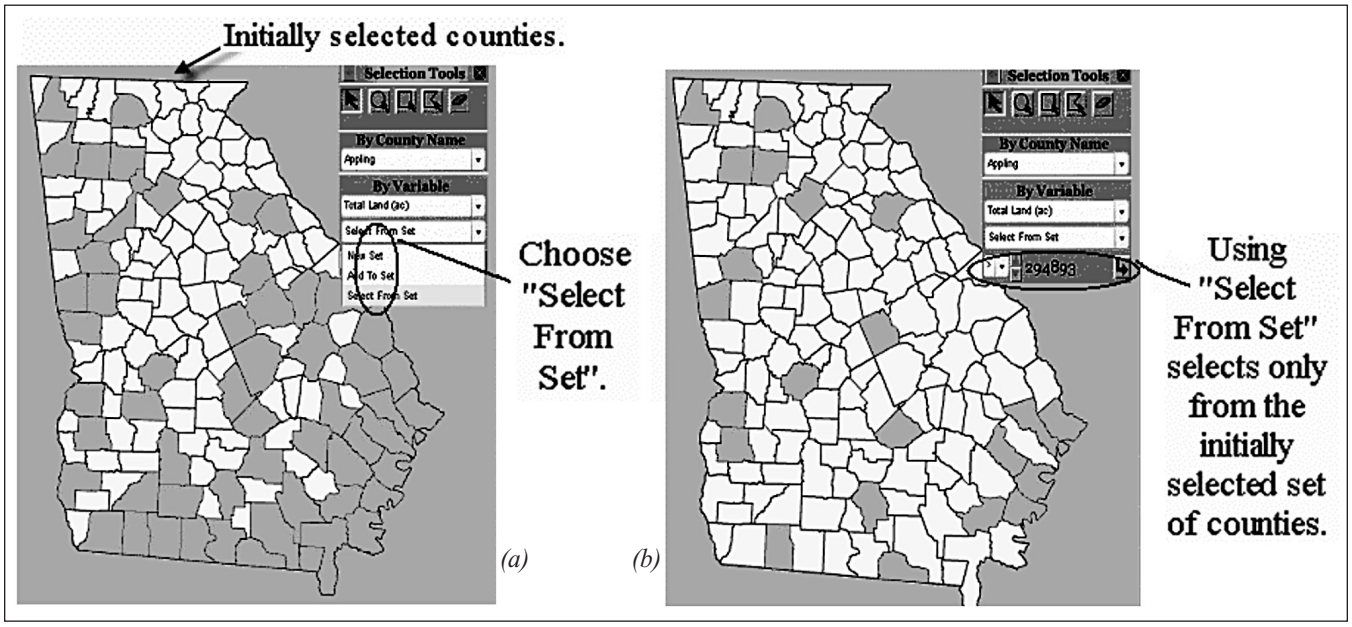
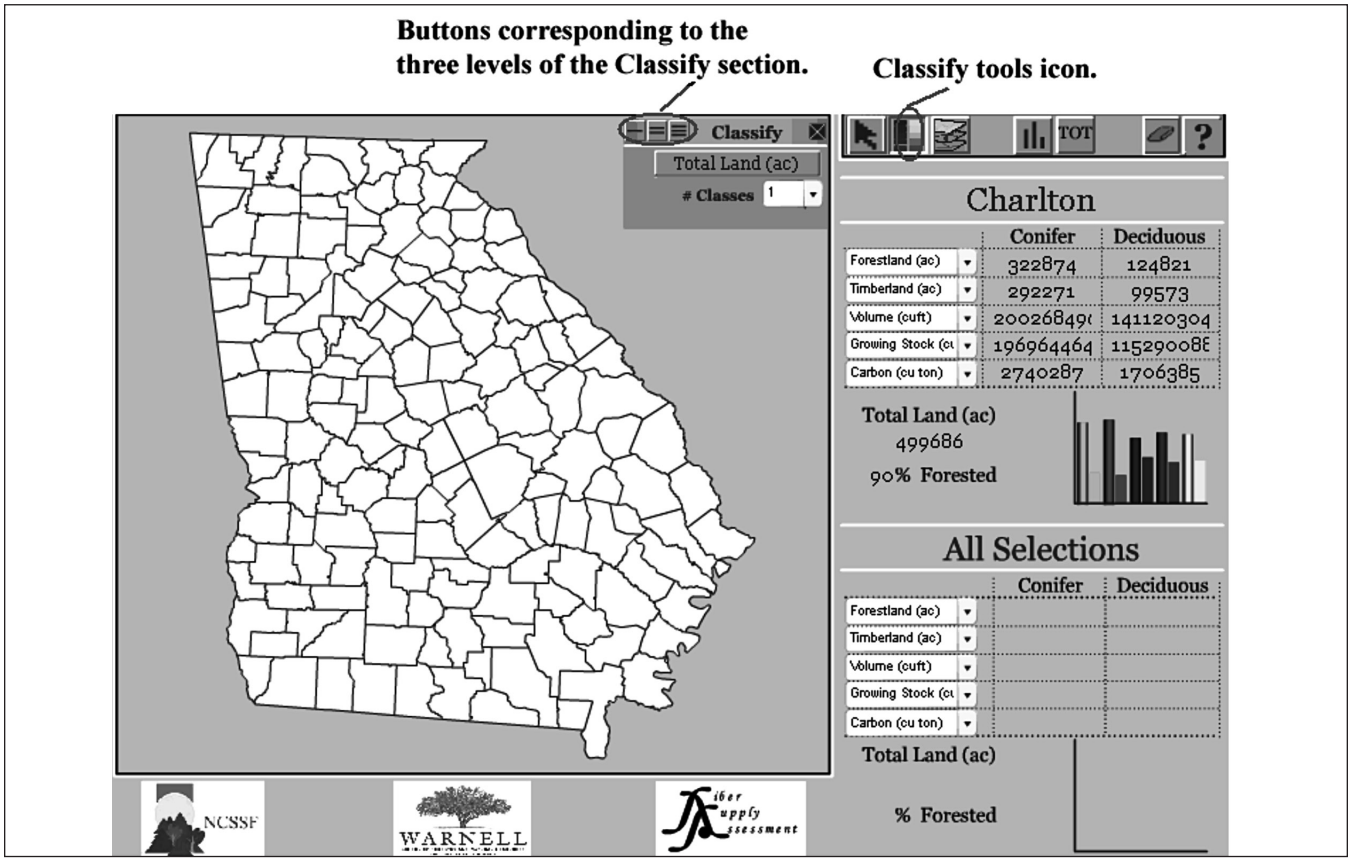


Figure 13.—*The icon for opening the Classify palette is the multicolored one in the icon row at the top right of the screen. The Classify palette contains three buttons corresponding to its three display levels.*



Level 2

Level 2 (fig. 15), which can be opened by either clicking the variable value for row 1 in the Classify palette or by clicking on the double dash symbol in the palette, shows the Maximum (Max), Average (Ave), and Minimum (Min) values for the selected variable. These values will also display for any of the values in the QC menu over which the cursor is hovered.

Figure 14.—*At level 1, Total Land is the selected variable and the different colors illustrate the range of values of the selected variable.*

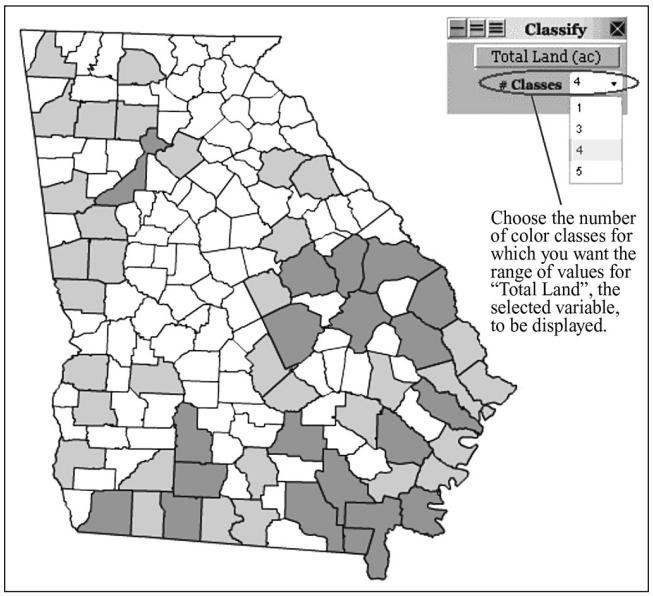
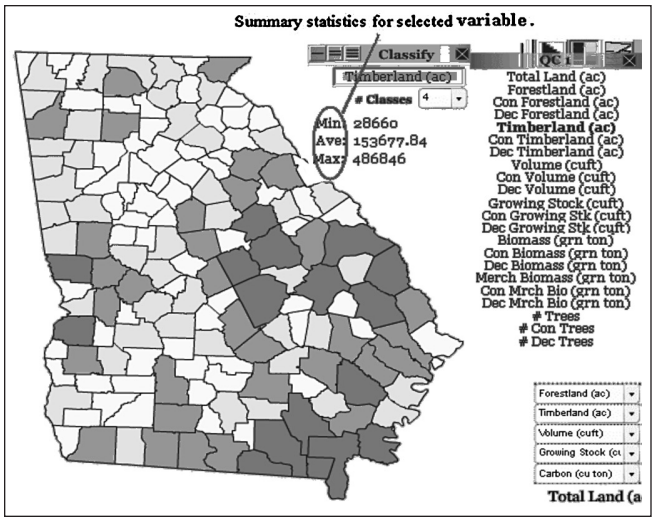


Figure 15.—*When the Classify palette is opened to level 2, summary statistics are shown for Quick Classify row 1.*



Level 3

To represent more than one variable on the map, the user can open level 3 by clicking the triple dash symbol on the Classify palette. In this level, all six QC rows are visible (fig. 16). As mentioned earlier, for the first row, a range of up to five colors can illustrate the relative values of the selected forestry variable for each county on the map (fig. 14). The forestry variable for QC row 2 is represented by graduated circular symbols (fig. 16). Each forestry variable in rows 3 through 6 is represented in the map as a bar chart (fig. 17). For the graduated circle and the bar charts, the color can be customized (fig. 18). For the Classify section, to clear the graduated circles left on the map by using row 2 and the bar charts left by using rows 3 through 6, shift-click on the row to clear the elements. To clear all elements, including selected counties, simply refresh the page using the browser's refresh function.

Layers for Carbon InFORM

Click the gray diamond on the left side of the layers palette to open a menu that allows the user to turn on and off roads, rivers, FIA regions, and physiography and to show locations of major cities. Click the name of the layer to display it; to cancel the selection, click the same name again.

Figure 16.—*When the Classify palette is opened to level 3, and Quick Classify row 2 is visible, graduated circles illustrate forestland values for each county.*

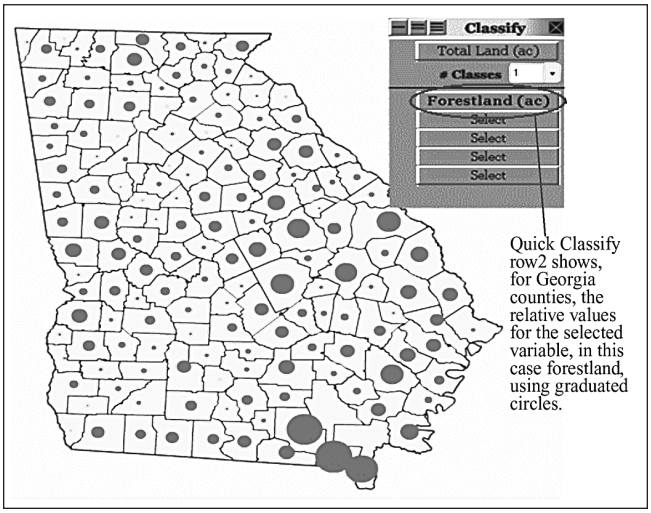
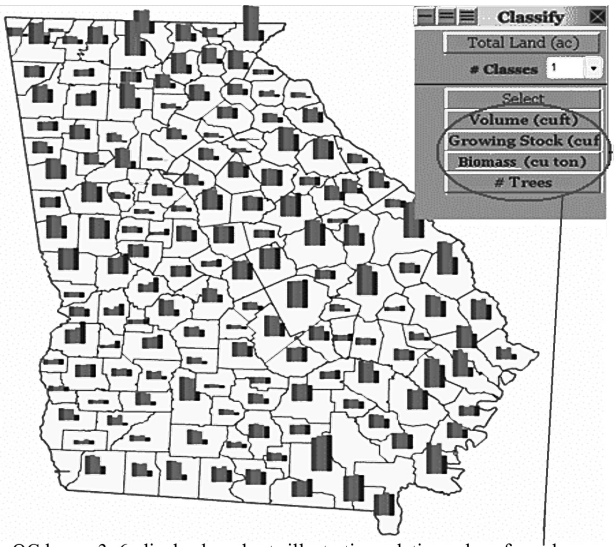
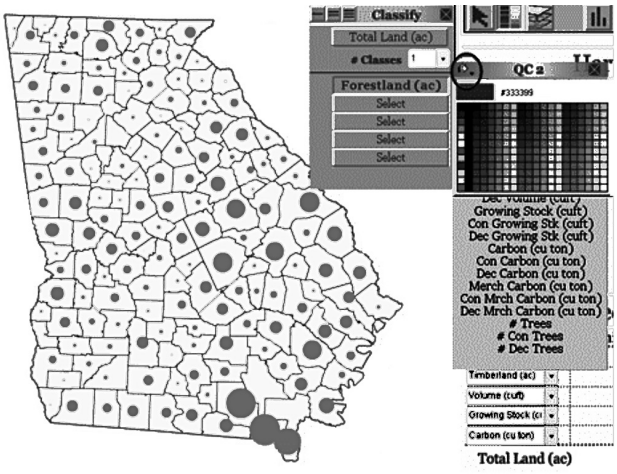


Figure 17.—*When the Classify palette is opened to level 3, and the Quick Classify rows 3 through 6 are shown, bar charts illustrate relative variable values for each county.*



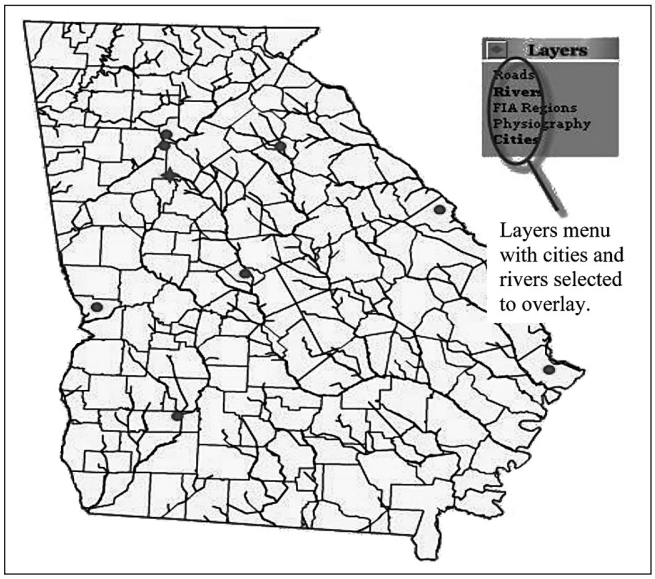
QC layers 3–6 display bar charts illustrating relative values for volume, growing stock, biomass and tree number for Georgia counties.

Figure 18.—*Expand color menu to select color choices for the graduated circles displayed by use of Quick Classify row 2.*



Clicking color tool brings up color palette, which allows choice of colors for graduated circles (row 2), and bar charts (row 3–6) when present.

Figure 19.—*One possible combination that can be displayed using the layers menu.*



Conclusion

The updates described to the Uni-Select and Multi-Select modules continue to increase availability and value of FIA to the general public, private landowners, policymakers, and forest managers. The Uni-Select updates now make multiple years of FIA data available so that trends over time can be easily tracked. Using the Multi-Select platform to create Carbon InFORM demonstrates the utility and flexibility of the FIA data for addressing various public concerns about forest and environmental management. It also provides a mechanism for interested parties to rapidly assess their forest C balance and the potential for future C sequestration.

Next-Generation Simulation and Optimization Platform for Forest Management and Analysis

Antti Mäkinen[1], Jouni Kalliovirta[2], and Jussi Rasinmäki[3]

Abstract.—Late developments in the objectives and the data collection methods of forestry create new challenges and possibilities in forest management planning. Tools in forest management and forest planning systems must be able to make good use of novel data sources, use new models, and solve complex forest planning tasks at different scales. The SIMulation and Optimization (SIMO) platform offers a modular forest planning system that adapts to various forest planning tasks by being flexible, adaptable, and extendable. Two different growth and yield simulators for Finnish conditions have been successfully implemented in the SIMO platform.

Introduction

Forest planning and management in Finland relies on certain well-established procedures that define how the data is collected and how it is processed. The basic unit of forest planning is a forest stand, and the conventional forest management objective has been steady timber procurement. The decision-making process in forest management planning is based on evaluating alternative forest development scenarios with a forest planning system. The forest planning system typically consists of a simulator and some optimization method. The simulator predicts the future development for the individual stands, and the optimization method is used to select the best forest management scenario from the number of alternatives produced by the simulator. Because stand-wise data with stand-level mean attributes have been the main data source for forest planning, the data models in the existing forest planning systems have been designed to use mainly stand-level data with a particular set of attributes. The data model dictates what kind of information the system can use and how it is recorded in the system (Paananen 1994, Tokola *et al.* 1997). Also, the prediction models and the program logic have been designed to process the stand-level data. This strong connection between the customary data collection process and the management planning process constitutes a system that can be efficient but, on the other hand, inflexible if changes need to be made.

The objectives in forest management have also been quite simple and conventional, but the settings and the manner in which forestry is practiced have been changing. Today, ecological and social values have to be taken into account in addition to the economic values (Kangas and Kangas 2005). The tools used in forest planning have to be able to adapt to these developments.

Forest planning data is collected through stand-wise inventory mainly by visual assessment and partly by subjectively positioned sample plots. The problem with this kind of data collection process is that the quality of the data is low and the costs are fairly high because it is a labor-intensive method. Given the often poor quality of the data, it is unrealistic to expect reliable prognoses from the planning system. The technical developments have brought new data sources, however, which can provide more precise and accurate forest data with comparable costs. Examples of these new technologies are remote-sensing methods such as aerial photography and laser scanning with fine spatial resolution, even at the single-tree level (Holmgren *et al.* 2003, Korpela and Tokola 2006, Maltamo *et al.* 2005,

[1] Researcher, University of Helsinki, Department of Forest Resource Management, Latokartanonkaari 7 (PL 27) 00014 University of Helsinki, Finland. E-mail: antti.makinen@helsinki.fi.

[2] Researcher, University of Helsinki, Department of Forest Resource Management, Latokartanonkaari 7 (PL 27) 00014 University of Helsinki, Finland. E-mail: jouni.kalliovirta@helsinki.fi.

[3] Researcher, University of Helsinki, Department of Forest Resource Management, Latokartanonkaari 7 (PL 27) 00014 University of Helsinki, Finland. E-mail: jussi.rasinmaki@helsinki.fi.

Næsset and Økland 2002). New field data collection devices such as the laser-relascope can provide more precise and efficient field plot measurements (Kalliovirta *et al.* 2005). The existing forest planning systems cannot take full advantage of these novel data sets because, in most cases, the data do not converge to the system's data model. Changing the data source from stand-level data to tree-level remotely sensed data (e.g.) can be hard or even impossible if the data model is fixed.

Because the forest planning data have traditionally been quite uniform, it has been possible to do most of the growth and yield calculations with a certain limited set of models. This process has been simple and straightforward; all the planning data is of the same type and format and the planner can generate the growth and yield predictions with the same set of models. Some properties of the planning systems, such as the timber prices and the costs of forest operations, have been parametric and thus controllable by the user. Changing the actual mathematical models or the logical rules that define how the simulation process runs inside the system has been challenging and has required a lot of programming. This change has been an obstacle for users wanting to implement modifications in the existing models (e.g., a local calibration of the growth models).

Many users of forest planning systems, especially larger forest companies, operate globally and a forest planning system that could incorporate local growth and yield models anywhere in the world could be of great advantage. If the growth and yield models and the simulation logic are hard-coded in the system, it can be a very demanding task to implement a new set of models from a different geographic area in the system.

A forest planning system, in which new models could be easily implemented, would be a powerful tool also for research purposes. This adaptable forest planning system would allow researchers to test and evaluate new models as a part of larger forest planning system, which would provide instantly the effects of the new model in the whole planning process.

Existing forest planning systems also almost completely lack the control of the quality of data and of the quality of the forest management plans produced with the system. From the user's point of view, it would be essential to have an estimate of the reliability of the forest management plan, because the decisions on forest management in practice base heavily on the management plans. This estimate becomes even more important because the demand for efficiency in forest planning is increasing and the forest planners do not have the time or the resources to thoroughly evaluate the management plans and their quality. A system that could point out the stands in which the input data was likely to be erroneous would help the planners concentrate their efforts to those problematic stands.

A project for developing a SIMulation and Optimization (SIMO) platform for next-generation forest planning started at the University of Helsinki in 2004.[4] The SIMO planning system is implemented as a software platform consisting of a number of modules that include different functions. These modules together constitute a forest planning system that is flexible, adaptable, and extendable. The term "platform" is used here in the sense that various forest models for different conditions and planning tasks can be implemented in the same platform. One part of the project is to implement a set of growth and yield models for Finnish conditions to produce a functional planning system that could by used by forest planners in Finland.

Requirements for the Next-Generation Planning System

The requirements for the next-generation forest planning system were based on interviews with the different forestry organizations that are the users of the existing forest planning systems. The results of the interviews were reviewed and the different requirements cited by the users were summed up into more concise form. The most notable issue that the users repeatedly brought up was the demand for a system that was transparent and that could be easily modified according to the users' needs.

[4] The SIMulation and Optimisation (SIMO) platform is being developed in the SIMO research project, which is financed by UPM-Kymmene Forest, Tornator Oy, Metsämannut Oy, Metsähallitus, Forestry Development Centre Tapio, and Tekes.

In the process of designing the platform, we needed to consider several different aspects. Different users may use different data sources, such as traditional compartment-wise inventory data consisting of stand-level mean attributes or single-tree–level remotely sensed data, in their planning process. In addition, grid-based data, in which forest cover is divided into cells of given size, has been becoming more popular. The ability to accommodate different types of data in the planning process demands flexibility from the planning system's data model.

The planning system should also be capable of solving planning problems of different types and scales. A forest planning problem can be optimizing an operational small-scale stand harvest schedule in which the planner only has to consider maximizing the forest property's net present value. Then again, the planning problem might be a large-scale ecological conservation plan in which the ecological and social aspects as well as the economical values have to be taken into consideration. Often, the inclusion of new objectives in forest planning requires that new kinds of data or models be incorporated in the planning system. Because we cannot exactly foresee the future needs of forest planning systems, the systems should be designed so that it would be as easy as possible to extend them when a need arises.

The main requirements for the SIMO forest planning platform were that the system's data model be flexible and the system be adaptable to different types of planning problems and extendable for possible future needs. More specific requirements were given as well. One important requirement was that the planning system not work like a "black box" in which most of the functionality is hidden and cannot be manipulated by the user. Instead, the system should be transparent and provide the users with access to the different mechanisms inside the planning system and a means to verify the internal workings of the system.

Forest planning systems are often used as a part of or in connection with more extensive data management systems. Because of this connection, interfacing the planning modules to different databases and user interfaces should be made as easy as possible. Because the data models and formats differ between different systems, the data interface to the planning system should be such that it can handle different types of data.

Forest information systems in general include spatial information on the forest stands. Also, tree-level spatial information can be stored with the new and coming data collection methods, such as single-tree remote sensing. The existing forest planning systems do not use this spatial information, however. With the next-generation forest planning system, it should be possible to use the spatial information in the planning process.

Technical Solutions

Platform Design

The aim of this project was to develop a forest planning platform comprising modules that include all of the functionality needed in a forest planning system. This kind of functionality demands a simulator module for creating the forest growth projections and an optimizing module for selecting the management scenario that conforms to the user's objectives. The modularity of the system is important because the users may want to include only some specific forest planning functions in their information systems. A user interface development is not involved in this project because the modules can be integrated behind existing user interface. Here, the simulator module is given the main focus because it can be considered as the core of the planning system and a lot of decent optimization tools are available.

The structure of the simulator is divided here into four logical components: (1) the data model, (2) the models used in the simulations, (3) the simulation logic, and (4) the application logic. The simulation logic comprises the logical rules that define how the models are used to process the data and the application logic is the actual program code that interprets and executes the simulation rules defined in the simulation logic. To produce a flexible, adaptable, and extendable system, the data, the models, and the simulation logic can be modified by the user, but there should be no need to modify the application logic. The user-modifiable components of the simulator are implemented using Extensible Markup Language (XML) documents (Boag *et al.* 2005). The use of XML is proposed here because it is inherently universal and portable, human readable, and easy to process programmatically (McGrath 2003). The syntax for the

simulator's XML documents is defined in XML schemas. One advantage of using XML documents is that they contain not only the data values but also the interpretation and metadata for the values. A common vocabulary in all of the simulator's XML documents describes the relationships and works as a glue between the different components of the simulator.

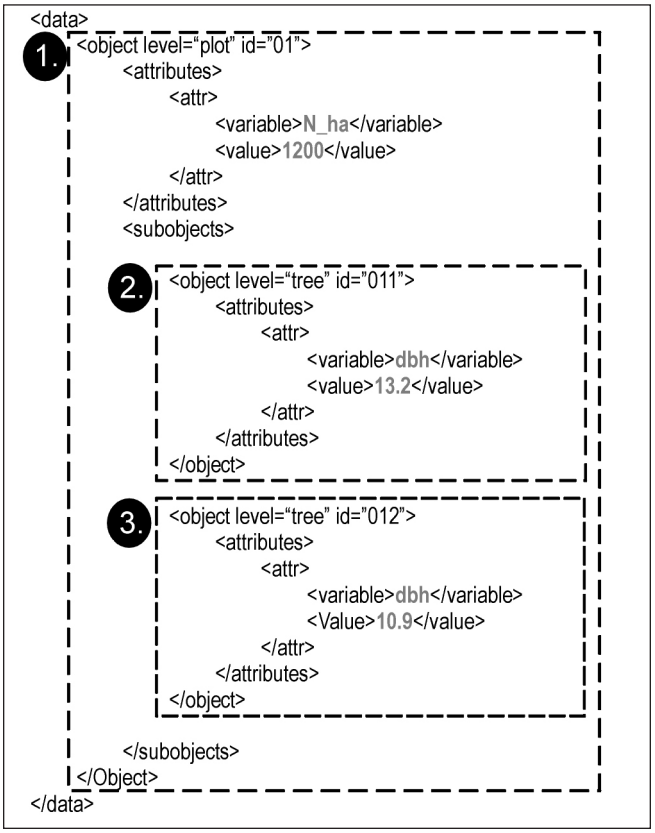
Simulator Module

Data Model. The requirements defined that the simulator's data model should be flexible instead of fixed. The data model proposed here is a hierarchical structure that consists of objects that can have attributes and subobjects. The object-subobject structure is recursive; i.e., the subobjects are objects. The attributes consist of variable-value pairs that define both the interpretation and the value of the attribute. The definition of an object in the data model is abstract and the user can define what the objects actually represent. A typical implementation of this object model could be Finnish national forest inventory plot data with a two-level hierarchy in which a single plot includes a number of trees as subobjects. This kind of hierarchy is pictured in figure 1. Implementing different kinds of data sets on this data model is easy because the objects in the hierarchy can represent almost any real-world object. For the data interface, the use of XML is proposed. Forest data is usually hierarchical and converts into XML format easily, which makes XML a good interface for transporting data between different forest information systems.

Models. The models in the forest planning system describe the natural processes; they also describe the human actions, such as harvests, in the forest. An individual model can be a simple mathematical model describing some stand or tree feature. Predicting the future development of a stand is more complicated and usually involves the use of a number of individual models together. One of the requirements for the next-generation planning system was that implementing new models in the system should be easy. An extendable model base is a set of individual model implementations separated from the application. This kind of separated model base enables straightforward implementation of new models and reuse of old models. Each individual model in the model base consists of three different components: (1) the program code for the model, (2) an XML format description of the model, and (3) an interface between the model and the simulator, generated automatically from the model XML description (fig. 2). The XML description of a model interfaces the separate model implementation to the system and specifies the model's input and output (Kokkonen *et al.* 2003). Currently, models written in C, C++, Fortran, and Python programming languages can be implemented in the model base.

Simulation Logic. The actual simulation is described in the simulation logic, which includes the logical rules on how and in which order the models in the model base are used to process the data. The simulation logic is defined in XML documents as model chains that consist of a hierarchical task structure. The simulation logic is split into tasks that are processed if the condition for the task is satisfied. The conditions in the model chains are logical expressions that can be evaluated against data values. Because the task structure is hierarchical, the tasks can be further split into subtasks. Each task ends up in a single model at the bottom of the task hierarchy. The

Figure 1.—*Hierarchical data in which a plot-level object (1) has tree-level subobjects (2 and 3).*



simulator can be modified by changing the models in the model chains, changing the parameters of the models, or changing the conditions of the tasks. All this can be done by modifying the XML files and without touching the program code. A hierarchical task structure of a model chain is shown in figure 3 in XML format and as a more illustrative tree structure.

As a part the development project, we implemented two different forest simulators for Finnish conditions in the SIMO platform: a tree-level simulator that predicts forest growth by using single-tree growth models on a set of trees generated with diameter distribution models and a stand-level simulator that predicts forest growth with models using stand-level mean attributes. The tree-level simulator is based on the Finnish Forest Research Institute's MELA forest planning system models (Hynynen *et al.* 2002) and the stand-level models are by Vuokila and Väliaho (1980).

Application Logic. The application logic covers the rules that define how the input data, the models, and the simulation logic files are processed by the simulator application. The simulator application parses the XML files containing the simulation logic (i.e., the model chains and the simulation parameters) and applies different models to the data according to the logical rules in the model chains. The simulation application itself does not constitute an actual simulator; it needs all the information about the data structure, the models depicting the natural processes, and the simulation as XML input files. Because the simulation logic is separated from the application logic, the simulator can be modified without touching the actual program code. The simulator application is written in the Python programming language. Users should not be required to do any reprogramming on the simulator's application logic to modify the simulator or even to implement new simulators on the platform. The program code is well documented, however,

Figure 2.—*Three model components: (1) Extensible Markup Language definition, (2) interface between model and simulator, and (3) program code implementation.*

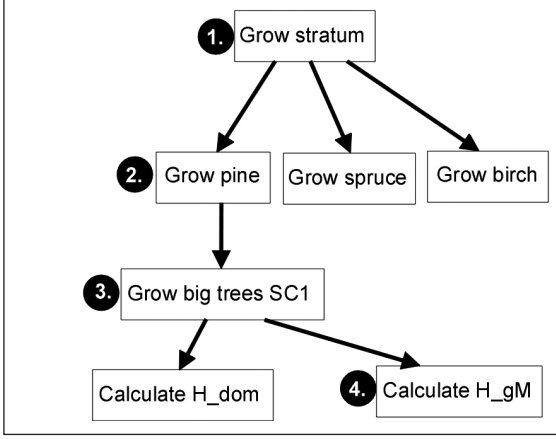
```
1.
<model>
    <name>Mean_height_pine_JKK</name>
    <Implemented_at>StaticstandModels.dll</implemented_at>
    <implemented_in>C</implemented_in>
    <author/>
    <description/>
    <published_in/>
    <species_list>
        <species>1</species>
    </species_list>
    <geographical_coverage/>
    <applies_for/>
    <variables>
        <Variable>
            <name>H_dom</name>
            <level>stratum</level>
            <lower_limit>1.3</lower_limit>
        </variable>
    </variables>
    <result>
        <object>self</object>
        <variables>
            <variable>
                <name>H_gM</name>
            </variable>
        </variables>
    </result>
</model>

2.
def Mean_height_pine_JKK(arg, libc):
    return libc.Mean_height_pine_JKK(
        arg.arg_table[0], byref(arg.nres), arg.mem,
        Byref(arg.warnings), byref(arg.errors),
        arg.error_check_mode, arg.allowed_risk_level, arg.rect_factor)

3.
extern "C" __declspec(dllexport)
int Mean_height_pine_JKK (double H_dom,
                          int *nres, double *modelresult, char *warnings, char *errors,
                          int errorCheckMode, double allowedRiskLevel, double rectFactor) {
}
```

Figure 3.—*A hierarchical task structure in Extensible Markup Language format and as a tree structure.*

```
<model_chain name="Grow stratum" output="on"> 1.
    <task name="Grow pine"> 2.
        <condition>
            <expression operator="eq">
                <variable>
                    <name>SP</name>
                    <level>stratum</level>
                </variable>
                <value>1</value>
            </expression>
        </condition>
        <task name="Grow big trees SC1"> 3.
            <task name="Grow H_dom">
                <model output="on">
                    <name>Dominant_height...</name>
                    <prediction/>
                </model>
            </task>
            <task name="Calculate H_gM"> 4.
                <model output="on">
                    <name>Mean_height_pine_JKK</name>
                    <prediction/>
                </model>
            </task>
        </task>
    </Task>
</model_chain>
```



and structured for the possible developers who want to extend or modify the application logic.

Other Platform Modules

In addition to the simulator module, the platform includes other modules that constitute different kinds of functionality in the planning system. Other important modules that are part of the platform include optimizer, validator, Geographic Information System (GIS), and reporter.

The optimizer module is basically a programmable interface to another program library. It is an "optimization package" that includes linear programming methods and metaheuristic methods for selecting the best alternative management scenario. This interface can be modified for several different optimization packages because different optimization methods suit different planning problems.

The validator module's main function is to validate the input data and all the files constituting the simulation logic. The use of XML documents as the data exchange format as well as the format for the simulation logic files provides a lot of tools for the automatic validation of the structure and format of the files. Validating the structure and format is easy, but validating the content (i.e., the data values) is a different kind of problem and is strongly linked to the quality control of forest planning data.

Because providing tools for using spatial information in the planning process was one of the objectives of the forest planning system, a module that handles spatial data processing is also being implemented in the platform. The GIS module will be similar to the model base of the platform; it will be a set of functions that is separated from the application. The GIS module will be able to compute basic spatial operations, such as neighborhood operations.

A reporter module is needed to process the simulation results into different formats for the user. Using the reporter module, the user can produce XML format data exchange files, text files, and different kinds of charts. The reporter module can be also used to produce data transfer files to some other information system.

Discussion

The development of a next-generation forest planning system was guided by three main objectives: flexibility, adaptability, and extendability. The modular structure of the platform components and the separation of the application logic from the simulation logic were the most important technical solutions in achieving these objectives. The resulting platform can be seen as an open system that can be modified and extended with new components according to current needs without time-consuming reprogramming (Reynolds 2005). The modules produced within this project include all the functionality needed in a forest planning system, except the user interface.

All the user-modifiable components are implemented as XML documents, which are text files and thus are easily modifiable by the user. The simulator's data model can be modified freely as long as the data is hierarchical. The simulation descriptions can also be modified without touching the program code. The only aspect that calls for programming is implementing new models in the model base. Nevertheless, in many situations, users can take ready-made models and implement those in the model base or reuse models already implemented in the model base.

Modifying the XML documents that define the current simulation task, although fairly easy for even a nonprogrammer, requires acquaintance with the syntax. The different structures defined in the XML schema files are designed to be as simple as possible and therefore the syntax the user needs to learn is quite easy. In addition to human readability, XML documents can be easily manipulated programmatically. A graphical user interface for manipulating the simulator's XML documents would make the system even easier to modify. Because it should be possible to integrate the SIMO modules in existing forest information systems, XML seems like a suitable interface for data transfer. Database contents can be transformed into XML format (e.g., with the standard XML query language, XQuery) (Boag *et al.* 2005).

The transparency of the system was one of the requirements of the development process. The separated simulation logic implemented as XML documents provides a simple and

human-readable access to the system and the simulation logic. A separated model base with well-documented model implementations and the XML definitions that include the metadata and the semantics for the models also adds up to the transparency of the system. A well-documented and structured program code is a necessity for an open system that can be continuously developed and extended.

The SIMO modules constitute a general forest planning platform that can be applied to different conditions and various planning tasks. As a means for verifying the adaptability of the SIMO platform, a tree-level simulator and a stand-level simulator were implemented on the platform. The simulation definitions, model chains, and model definitions were constructed for both of the simulators separately from the application development and without touching the program code. The experiences gained from implementing two very different simulators on the same platform are very positive and support the presumption that the SIMO platform constitutes a flexible, adaptable, and extendable forest planning system.

The next likely step in the further development of the SIMO platform is creating graphical user interfaces for running the simulations and for manipulating the XML documents. For a forest planning system to be widely usable, a user-friendly interface is a necessity. Implementing different types of growth and yield models (e.g., models for different vegetation zones or process-based growth models) would provide more valuable experiences on the modifiability of the system.

Literature Cited

Boag, S.; Chamberlin, D.; Fernández, M.F.; Florescu, D.; Robie, J.; Siméon, J., eds. 2005. XQuery 1.0: an XML query language. http://www.w3.org/TR/xquery. (7 January 2007).

Holmgren, J.; Nilsson, M.; Olsson, H. 2003. Estimation of tree height and stem volume on plots using airborne laser scanning. Forest Science. 49(3): 419-428.

Hynynen, J.; Ojansuu, R.; Hökkä, H.; Siipilehto, J.; Salminen, H.; Haapala, P. 2002. Models for predicting stand development in MELA system. Res. Pap. 835. Helsinki, Finland: Finnish Forest Research Institute. 116 p.

Kalliovirta, J.; Laasasenaho, J.; Kangas, A. 2005. Evaluation of the laser-relascope. Forest Ecology and Management. 204(2-3): 181-194.

Kangas, J.; Kangas, A. 2005. Multiple criteria decision support in forest management— fundamentals of the approach, methods applied, and experiences gained. Forest Ecology and Management. 207: 133-143.

Kokkonen, T.; Jolma, A.; Koivusalo, H. 2003. Interfacing environmental simulation models and databases using XML. Environmental Modelling & Software. 18: 463-471.

Korpela, I.; Tokola, T. 2006. Potential of aerial image-based monoscopic and multiview single-tree forest inventory: a simulation approach. Forest Science. 52(2): 136-147.

Maltamo, M.; Mustonen, K.; Hyyppä, J.; Pitkänen, J.; Yu, X. 2005. The accuracy of estimating individual tree variables with airborne laser scanning in a boreal nature reserve. Canadian Journal of Forest Research. 34(9): 1791-1801.

McGrath, R.E. 2003. XML and scientific file formats. Urbana, IL: University of Illinois, Urbana-Champaign, National Center for Supercomputing Applications. 25 p. http://hdf.ncsa.uiuc.edu/projects/nara/XML_and_Binary.pdf. (13 December 2007).

Næsset, E.; Økland, T. 2002. Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. Remote Sensing of Environment. 79(1): 105-115.

Paananen, R. 1994. A data model for a GIS-based forest information system. Res. Pap. 493. Helsinki, Finland: Finnish Forest Research Institute. 78 p.

Reynolds, K.M. 2005. Integrated decision support for sustainable forest management in the United States: Fact or fiction? Computers and Electronics in Agriculture. 49(1): 6-23.

Tokola, T.; Turkia, A.; Sarkeala, J.; Soimasuo, J. 1997. An entity-relationship model for forest inventory. Canadian Journal of Forest Research. 27: 1586-1594.

Vuokila, Y.; Väliaho, H. 1980. Growth and yield models for conifer cultures in Finland. Communicationes Instituti Forestalis Fenniae. 99.2. In Finnish.

SOLE: Enhanced FIA Data Analysis Capabilities

Michael Spinney[1] and Paul Van Deusen[2]

Abstract.—The Southern On Line Estimator (SOLE), is an Internet-based annual forest inventory and analysis (FIA) data analysis tool developed cooperatively by the National Council for Air and Stream Improvement and the Forest Service, U.S. Department of Agriculture's Forest Inventory and Analysis program at the Southern Research Station. Recent development of SOLE has enhanced data analysis capabilities to include diameter-based categorization of variables, mapping based on plot hex, and tree-volume growth projections. The breadth of analysis options provides a powerful system for FIA data mapping and analysis (http://ncasi.uml.edu/SOLE/).

Introduction

Internet-based database analysis tools maximize the accessibility of databases. Developers can specify data retrievals and analysis algorithms that ensure correct data analysis. The nuances of complicated analyses can be "hidden" behind a simple interface. Thoughtful software development empowers users to quickly and easily explore data they may otherwise never use.

The Southern On Line Estimator (SOLE) is an online annual forest inventory and analysis (FIA) data analysis tool that has been operational for 3 years. Users select an area of interest; quantitative and qualitative variables; filters (optional); and a tabular-, chart-, or map-based analysis. The mechanics of using SOLE are available in a previous FIA Symposium paper (Spinney, Van Deusen, and Roesch 2006) and also in the integrated online help files. This paper describes interface enhancements and additional analysis capabilities.

Interface Enhancement: Polygon Data Retrieval

The first interface enhancement provides the ability to select data via user-drawn circles and polygons. To access this feature, the user must first select States(s) on the **US Map** tab and activate the ***Select States then Counties button***. SOLE automatically advances the user to the **Data** tab. By default, the county-level data selection on the ***County Tools*** subtab is activated. Choose the ***Draw Tools*** subtab to use the polygon selection tools. (See figure 1.) The following four Draw Tools are available:

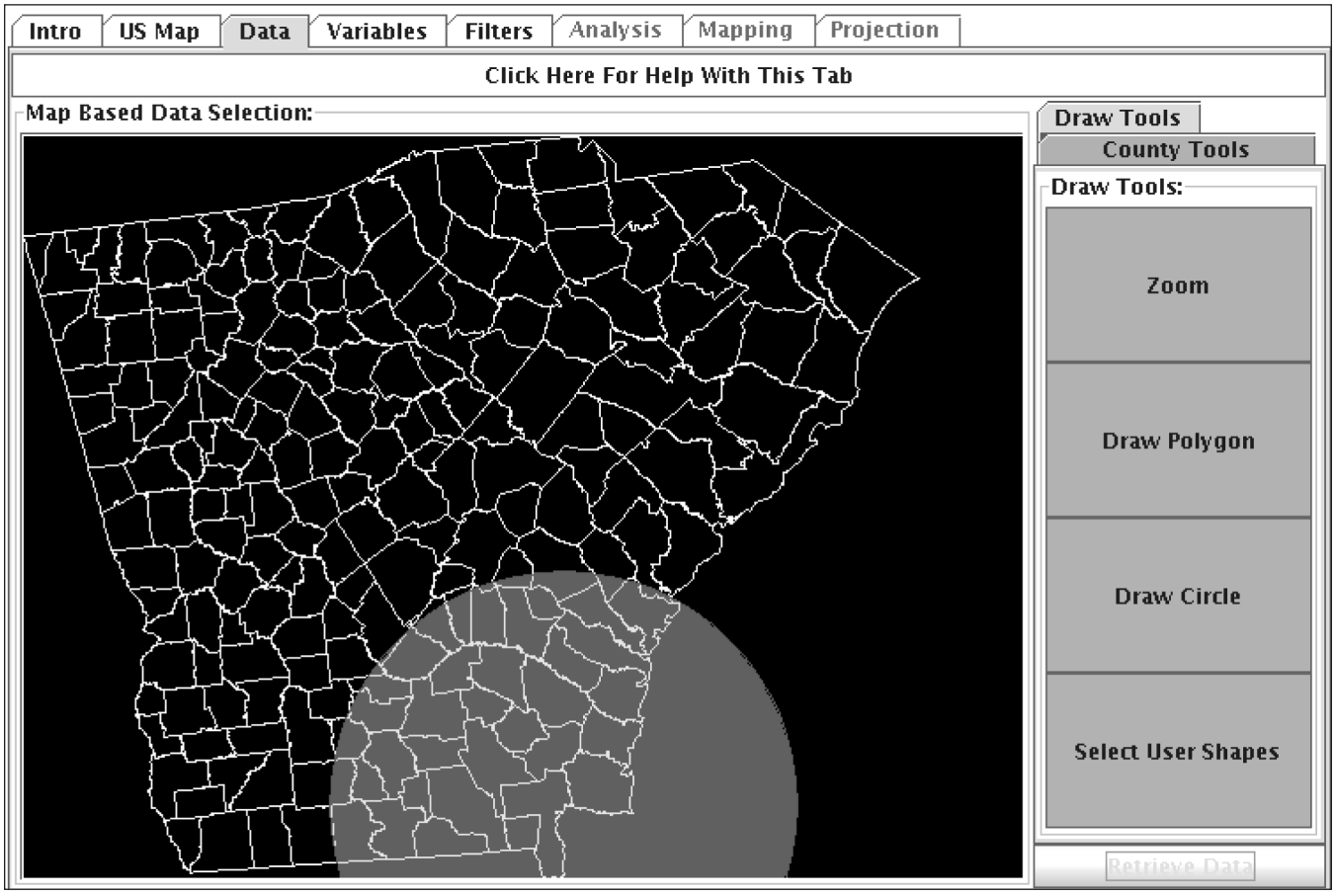
1. Zoom: Left clicking zooms in on the map and right clicking zooms out.
2. Draw Polygon: Left clicking adds sides; right clicking adds last side and closes the polygon. After the shapes are drawn, user-defined shapes will be outlined in red.
3. Draw Circle: A popup window will ask if you want to draw your circle freehand or if you want to manually enter center coordinates and radius. After the shapes are drawn, user-defined shapes will be outlined in red.
4. Select User Shapes: Activating this tool turns the map background gray. Left clicking a user shape will select the shape and turn it green. Note: One or more shapes must be selected before the **Retrieve Data** button is pressed.

New Analysis Capabilities

SOLE now offers the capability to analyze volume and biomass variables in 5-in diameter at breast height (d.b.h.) classes. These classes increase the precision of estimates through indicating the allocation of a variable over the diameter distribution, which also acts as a proxy of wood product class. d.b.h. classes can be used as a qualitative (factor) variable (on **Variables** tab) or as a filter (on **Filters** tab).

[1] Quantitative Forestry Analyst, National Council for Air and Stream Improvement, 600 Suffolk Street, Fifth Floor, Lowell, MA 01854. E-mail: mspinney@ncasi.org.
[2] Principal Research Scientist, National Council for Air and Stream Improvement, 600 Suffolk Street, Fifth Floor, Lowell, MA 01854. Web site: http://ncasi.uml.edu.

Figure 1.—*SOLE's user-defined polygon data retrieval.*



The user can also produce maps based on FIA plot hex, which are the 2,402.6-ha hexagons encompassing each FIA plot that completely tessellate the State. Using this plot hex mapping method increases spatial mapping resolution over the county or State level. The user can choose hex mapping options to map the individual plot hex values for any qualitative or quantitative variable, or produce a loess-smoothed response surface.

SOLE also offers the option to generate a 40-year projection of potential timber yield for accessible forest land in a State. SOLE's projection analysis "grows" FIA inventory plots ahead in time using a hotdeck matching technique (Sande 1983). These management scenarios provide input to a harvest-scheduling program that estimates potential yield trends for accessible forest land. Public land and land close to cities are assumed to be unavailable for timber harvesting, and all other private land is assumed to be accessible.

The Habplan harvest scheduler (http://ncasi.uml.edu/projects/habplan/) was used to implement a landscape management plan over a 40-year planning period. Harvest scheduling, as defined here, requires the State to be divided into polygons with a list of potential management regimes for each polygon. The polygons for this application are FIA plot hexes. In any given year, a plot can be clearcut or left to grow another year. A regime denotes the timing of outputs that will occur if this regime is followed. A schedule is produced by assigning one regime to each polygon. Van Deusen (2001) provides additional information about Habplan's formulation.

The broad spatial scale (from the perspective of operational forest management) of FIA plots is reflected in the simple formulation of potential management action. Within the 40-year planning period, harvest can reoccur provided at least 20 years have elapsed between harvest events. Forest management

is assumed to be sustainable and follow best management practices, and timber growth/removals rates are assumed to remain relatively stable from one year to the next.

Projection analysis results in current and projected estimates of growing stock and potential removals. SOLE's projection analysis option is a work in progress. The harvest scheduling assumptions are relatively simple and the projected removals estimates are intended to indicate a general trend in potential timber harvest. They should not be confused with FIA's growth removals mortality data.

Conclusions

Internet-based FIA analysis tools are essential for proper analysis of FIA data. SOLE provides a simple interface that enables users to obtain customized analytical results. Flexibility in each component of SOLE ensures that SOLE remains highly adaptable to changes in both database structure and user needs. Recent data analysis development has leveraged SOLE's modular structure to enable polygon-based data retrievals, mapping at the plot hex level, projecting growth and removals via harvest scheduling, and classifying volume and biomass estimates by 5-in diameter class.

Acknowledgments

This project was supported by a Forest Service grant through the Southern Research Station.

Literature Cited

Sande, I.G. 1983. Hot-deck imputation procedures. In: Madow, W.G.; Olkin, I., eds. Incomplete data in sample surveys. New York: Academic Press: 334-350. Vol. 3.

Spinney, M.P.; Van Deusen, P.C.; Roesch, F.A. 2006. SOLE: online analysis of southern FIA data. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C.; McWilliams, W.H., eds. Proceedings of the sixth annual forest inventory and analysis symposium. Gen. Tech. Rep. WO-70. Washington, DC: U.S. Department of Agriculture, Forest Service: 75-79.

Van Deusen, P.C. 2001. Scheduling spatial arrangement and harvest simultaneously. Silva Fennica. 35(1): 85-92.

