Proof of Concept for an Approach to a Finer Resolution Inventory

Chris J. Cieszewski[1], **Kim Iles**[2], **Roger C. Lowe**[3], **and Michal Zasada**[4]

Abstract.—This report presents a proof of concept for a statistical framework to develop a timely, accurate, and unbiased fiber supply assessment in the State of Georgia, U.S.A. The proposed approach is based on using various data sources and modeling techniques to calibrate satellite image-based statewide stand lists, which provide initial estimates for a State inventory on a common timeline. The system is based on using Georgia ground inventory data from the forest products industry, enhanced by various geographic information system and remote sensing data, and applied with the k-th "nearest neighbor" methods to time-series-stratified satellite imagery. The initial estimates are then scaled regionally to the Forest Inventory and Analysis (FIA) summary totals to eliminate potential bias in the initial estimates. The system enhances the FIA inventory data in four significant ways. First, it removes the need for the specific FIA plot coordinates; although, the coordinates, if available, would probably enhance the analysis. Second, it provides a current common timeline of inventory estimates based on the Landsat Thematic Mapper imagery for the given year and season. Third, it provides currently accurate high-resolution area estimates. Last, it uses various auxiliary data available from private and public sources in the State and can easily take advantage of other data as they become available.

The American Forest and Paper Association's Second Blue Ribbon Panel (BRP) on the Forest Inventory and Analysis (FIA) program called for developing and implementing a plan to conduct a national inventory to be coordinated with State foresters, Federal land management agencies, forest industry, nongovernmental organizations, and others. In response to the second BRP's recommendations, the Agricultural Research, Extension, and Education Reform Act of 1998 (Section 253c) mandated that the U.S. Department of Agriculture (USDA) Forest Service conduct forest inventories in all States at the 20 percent annual rate of sample plots. The forest community is expected to obtain timelier and more accurate estimates of timber inventories and changes in fiber supply due to harvests, urbanization/sub-urbanization, natural disasters, and reforestation programs. Georgia was one of the first southern States to participate in the Southern Annual Forest Inventory System (SAFIS) in partnership with the USDA Forest Service FIA program. Currently, the Georgia Forestry Commission (GFC) provides 11 full-time equivalent positions, trucks, per diem travel, and other supplies to collect FIA data in Georgia.

Full implementation of an annual system by FIA requires reliable forested-area estimates, and standardized operating procedures to maximize benefits from informational resources such as satellite data.

Without locally accurate area estimates, the informational value of an annual sample is greatly reduced. One option for timely and accurate area estimation is to use remotely sensed spatial data such as Landsat Thematic Mapper (TM) satellite imagery, which has supplied reliable land cover area estimates in many parts of the country (e.g., Evans 1994, Rack 2001, Scrivani *et al.* 2001, Wynne *et al*. 2000).

Georgia is the third fastest growing State in the United States, although 72 percent of its land is forest cover with 9.55 million hectares in commercial forests, more than any other State (Smith *et al.* 2002). With more than two-thirds of that forest owned by approximately 630,000 private landowners and forestry contributing more than $30 billion to Georgia's economy each

---

[1] Associate Professor, D.B. Warnell School of Forest Resources, University of Georgia, Athens, GA 30602 USA; Corresponding Author. Phone: 706–542–8169; fax: 706–542–8356; Web site: www.growthandyield.com/chris.

[2] Consultant on Forest Inventory, Kim Iles and Associates Inc., 412 Valley Place, Nanaimo BC, V9R 6A6, Canada.

[3] GIS Analyst.

[4] Postdoctoral Fellow, D.B. Warnell School of Forest Resources, University of Georgia, Athens, GA 30602 USA; Assistant Professor, Department of Dendrometry and Forest Productivity, Faculty of Forestry, Warsaw Agricultural University, Nowoursynowska 159, 02-776 Warsaw, Poland.

year (Cieszewski *et al.* 2000), the results of this project will be useful to many forest managers and the State's economy.

Objective

Although Georgia government and forest industry support many forest activities and multiple sources of forestry data exist, the State does not have a timely and accurate high-resolution, spatially explicit forest inventory. The FIA-generated State inventory produced a low-resolution statewide survey with reliable estimates for very large areas, but does not provide accurate local, fine-scale values. In addition, the FIA inventory estimates are not derived for a common timeline and define a moving lagged average of the resource availability that is delayed by 3 to several years.

The objective of our project is to use the FIA and other data to derive timely and accurate high-resolution estimates for Georgia forests every year for the current year.

Proposed Approach

A large-scale inventory is very precise at the overall level but imprecise at the polygon level. More field plots will *not* solve this problem; they may exacerbate the problem by adding cost and delay. The objective of obtaining polygon-level precision, therefore, must be sought without the benefit of any additional fieldwork. The most promising approach to achieving this goal is to estimate every polygon volume or other characteristic and to ensure that these estimates add up to an appropriate total. How to determine this total is a separate topic and is beyond the scope of our study. The British Columbia Vegetation Inventory and a number of private forest companies have employed this approach.

The large-scale inventory maintained by FIA assumes that useful data and more precise results must come from statistical samples. Any inventory's design must, of course, be based on certain properties, such as unbiased data and correct measurements. The same problem arose during the design phase for the British Columbia Vegetation inventory; our project will apply the solution and approach used for that inventory.

An inventory that adds up to the same total as any unbiased estimate, regardless the source of that unbiased estimate, is itself an unbiased estimate. Therefore, we can make an estimate for every polygon on the land base, and then ensure that the sum of these estimates is constrained to add up to the total provided by an unbiased statistical process. The FIA is better equipped than any other organization to provide such an unbiased total for large areas in the United States. The FIA does not have the resources to provide the fine scale resolution of this total in individual polygon values. Other organizations, however, are prepared to make the individual polygon estimates, which can then be constrained by the FIA results.

Polygon estimates can be made using several methods, including "nearest neighbor" estimates, historical data, old inventory data, projected past values, aerial photos, and personal judgment, and any remote sensing technology. Inconsistent or partially available data is not a problem. Currently, the only advantage satellite methods offer is their ability to process large amounts of data, which increases the refresh cycle frequency. Satellite imagery, however, does not provide acceptable accuracy, because of a self-imposed attempt to produce the estimates automatically and because of insufficient resolution. One important advantage of satellite information is its ability to detect large-magnitude change.

Typically, the first objection made to using many estimates of polygon values is that they are biased. Adjusting the estimated parts to an unbiased overall total addresses this problem. A second objection, that such estimates are only available for a portion of the area or provide inconsistent precision, is not a serious constraint.

Table 1 presents a simple example of a small group of polygons that have been changed based on local knowledge of some sort to provide a more precise polygon-level estimate. The initial sample that the total of the polygon is based on was unbiased and provided a set of statistics describing that total. The absence of bias in the initial procedure and the value of any statistics regarding that total or average also apply for the revised polygon values. The difference is the improved polygon level resolution. The process's flexibility and inclusiveness are evident because other groups can contribute to the process; a "ground truth" visit, however, can verify any potential change.

Three significant changes in forest inventory data use and maintenance have occurred during the past few decades:

Table 1.—*A simple example of a small group of polygons that have been changed based on local knowledge to provide a more precise polygon level estimate.*

Polygon	Initial estimate	Final estimate	Further revised estimate	Further revision criteria
a	1,877	2,141.0	6,000	ecological guess
b	1,836	2,094.2	1,000	actual cruise
c	1,941	2,214.0	2,000	field visit
d	717	817.8	500	pure guess
e	1,584	1,806.8	2,200	10% more than c
f	996	1,136.1	600	similar neighbor
g	1,580	1,802.2	500	same as d
h	866	987.8	200	1/10 of c
	11,397	**13,000**	**13,000**	

Unbiased total = 13,000 Simple Correction Ratio = 1.141

- Aerial photography and other methods enable areas and forest value estimates to be made without fieldwork.
- Fast and high capacity databases allow individual polygons to have individual values; strata averages are not needed to store and report data only.
- Geographic information systems now function reliably, after a long and frustrating wait. Field information can be matched with information available from many sources.

The FIA contributed to these changes with the following actions:

- Developed a sample process to cover the entire land base, or at least make the grid extendable to all areas.
- Performed the fieldwork and made a continual effort to improve definitions, consistency, and data quality.

FIA data offers rigorous, statistically valid data with good quality control; other sources may offer only the ability to discriminate on a relative basis, and for only a portion of the overall land base.

The University of Georgia plans to combine these various types of data to create a fine-resolution inventory with location-specific information that is unbiased over some area to which it has been balanced. Because this information can be further refined, many specialist groups may be able to provide insight into improving the distribution of individual values that sum to a specific total.

How can this data be maintained and improved? When a newer or better estimate of the total is available, the individual polygons can be adjusted. Some polygons may be adjusted more than others, depending on how reliable the current estimate may be. Over time, the polygons should be grown or depleted according to the best information available. Although technically any inventory is biased as soon as the stands age, this detail is not expected to cause any serious errors.

The Forest Service has been working on several projects involving imputation and estimation that fit well with our project's approach. One closely associated method is the "most similar neighbor" work by Melinda Moeur, Al Stage, and others in the Forest Service (described in Moeur and Stage 1995).

Other Initial Estimation Aspects

The high-resolution inventory will be compiled in several steps. This section briefly describes the general framework for the unbiased, fine-resolution, spatially explicit estimation. First, to improve the analysis' accuracy, we will prestratify the Landsat TM images using multi-image change tracking.

Second, we will use various available inventory data provided by the forest industry and private forest land owners to develop models that stratify the satellite images to different species groups and volume/basal area classes in the prestratified classes. Although availability of the FIA exact coordinates would provide more reference points to calibrate the k-th nearest neighbor models, it is not imperative because the FIA estimates are used to adjust all high-resolution estimates to the unbiased total or average.

Third, we will remove any bias in the high-resolution estimates by scaling them so that the sums of their volumes or basal areas in each satellite image will be equal to the corresponding sums in the FIA estimates for a corresponding timeframe. For example, the corresponding classification can be applied to scenes from the time of the inventory estimates and, after scaling, the corrected estimates can be forwarded to the current time. One challenge for us is to determine how to do the scaling. The FIA estimates are not for any given time but for an average in a 5-year period. At any time during this period, we can expect removals and growth that are intractable; ignoring the removals and growth, however, can create a bias.

Fourth, the adjusted estimates for the k-th nearest-neighbor-calibrated polygons will be used to compute the current inventory for the given year.

We expect satellite data to help quantify forested resource areas. In addition, using consecutive images over the last 30 years, we will be able to identify when specific areas were cleared and reestablished, so that we will be able to estimate their current ages. Optimal success in this effort requires reliable ground data on a large number of acres at different ages. Some acres will be used for training sites (that is, sites to develop classification algorithms); the remaining acres will be used to test modeled sites (that is, sites to evaluate the efficacy of the classification algorithms). We have industrial and private nonindustrial cooperators willing to provide these data.

Stand structure data from the Piedmont and Coastal Plain regions of Georgia will be obtained from our cooperators and supporters. Study location selections will be based on digital data availability from our large industrial cooperators. Data from neighboring nonindustrial private landowners who volunteer to be partners in this project (through cooperation with GFA) will fill in around these industrial land holdings.

We will generate an urban mask to disregard areas within city limits, such as parks, and the confusing satellite signatures from suburban areas. Using data from U.S. Geological Survey (USGS) paper maps and other available sources that show remote building locations will ensure that most dwellings and other structures are masked.

Using the field data in conjunction with the TM data, we will determine TM signatures for the forest types of interest. We will evaluate the consistency of these signatures in each satellite image. From the combination of the summer and winter TM data, we will generate a hardwood mask to help prevent us from confusing the signatures of pine at different ages with the signatures of hardwoods and pasturelands. Ancillary elevation and stream data will be used to help separate hardwoods in riparian zones from upland hardwoods and help define buffers along drainages where it may be difficult to distinguish between hardwoods and pines.

For each polygon (delineated area) of data provided by our cooperators and field crews, we will determine the “overriding” land cover class in the TM signatures for that same area. We will then evaluate the accuracy with which our list of the forest-types and stand ages can be classified. We will verify the accuracy of our forest/nonforest polygon classification and forest-type polygon classification based on the match between polygon field class and polygon satellite class. We will then reevaluate the TM signatures and recheck some field locations before we report which forest-types are most commonly confused in TM signatures and why.

The primary analyses will focus on investigating the images’ changes over time, which mark the harvested polygons. The change points will be examined with geostatistical methods, such as variograms and cross-semivariograms (Zawadzki *et al.*, 2005) that define the cross-sectional changes consistently over time, except for periods and locations of disturbances, which this approach will attempt to identify.

Data

In this analysis, we use FIA data, forest ground inventory data obtained from the local forest industry, geographic information system (GIS) data, and the Landsat TM imagery. The FIA data came from the plot FIA database (Hansen *et al.* 1992, Miles *et al.* 2001). The initial estimates of a high-resolution statewide forest inventory will be based on various spatial data available publicly and privately. Some examples of the publicly available data, other than the USDA FIA data, are described below.

- The Georgia GIS Data Clearinghouse (GGDC) (http://www.gis.state.ga.us/Clearinghouse/clearinghouse.html) provides access to numerous county-level GIS data for the entire State.
- Hydrology data in vector format are available at the 1:24,000 scale. These data sets were captured from the USGS 1:24,000-scale topographic quadrangles and include linear features such rivers and streams and polygonal features like lakes and ponds. Most features are attributed by class (e.g., perennial, intermittent) so that major and minor rivers and streams can be determined.
- Road and highway data are available at the GGDC at the 1:12,000 scale. These data were captured from the 1993 digital orthophoto quarter quadrangles. They contain public roads including interstates, State highways, county roads, and city streets. These vector data are well suited to incorporate in various distance-related analyses in which the features are buffered to create polygons for further investigation.
- The GGDC also serves raster data. Digital elevation models (DEMs) are available at the 1:24,000 scale and a 30-meter pixel size.
- DEMs contain elevation information from which slope and aspect data sets can be derived.
- Land cover data are available at the 1:100,000 scale and a 30-meter pixel size. These data, developed using satellite imagery from the late 1980s and the early 1990s, divide the landscape into different classes such as conifer, deciduous, agriculture, and urban. Though dated, they provide a source for stratifying the landscape into broad cover types.
- Aerial photographs, historical and recent, are available from the GGDC in digital format and in paper format from the University of Georgia's Science Library and the GGDC. The GGDC sells two sets of digital aerial photographs:
  - The 1993 black and white digital orthophoto quarter quadrangles (DOQQs) have a 1-meter pixel and are available for the entire State.
  - The 1999 color-infrared photos (1-meter pixel) are available for select counties.
- The University of Georgia's Science Library maintains a large set of historical paper aerial photographs from the early to mid to late 1990s.
- USGS sells recent paper aerial photographs from the 1980s through the current decade (http://edcsns17.cr.usgs.gov/finder/finder_main.pl?dataset_name=NAPP). These data provide a good model verification foundation.
- Satellite imagery is available from the USGS EROS Data Center (http://edc.usgs.gov).
- Landsat Multispectral Scanner (MSS), Landsat 5 TM, and Landsat 7 Enhanced Thematic Mapper Plus (ETM+) satellite data are suitable for these types of landscape studies (Note: Recent malfunctioning equipment for ETM+ has yielded some suspect data in a scene).
- Other satellite imagery available includes ASTER, MODIS, and AVHRR, each of which can be used to discriminate between land cover types.

Our industrial partners supplied various GIS data, including boundaries and tabular data that may be the richest source of forest information. The final inventory of Georgia's forest resources will be scaled to be consistent with the FIA inventory regional and subregional statistical summaries (Thompson 1998).

Summary

This report describes a proof of concept to develop a high-resolution inventory based on pooling information from various types and sources of data. Because the data do not originate in a consistent statistical framework, they are likely to initially generate a biased inventory. Therefore, the initial inventory estimates are scaled to make the summary values equal to the summary values of the FIA inventory estimates, which will remove any existing bias in the final estimates of the high-resolution inventory. The proposed approach should allow timelier and more accurate inventory estimate compiling than either the initial remote-sensing-only based inventory or the moving average FIA survey estimates.

Literature Cited

Cieszewski, C.J.; Borders, B.E.; Whiffen, H.; Harrison, W.M. 2000. Forest inventory in Georgia. In: Zawila-Niedzwiecki, T.; Brach M., eds. Proceedings, IUFRO conference on remote sensing and forest monitoring; 1999 June 1–3. Rogów, Poland: Warsaw Agricultural University, Faculty of Forestry.

Evans, D.L. 1994. Forest cover from Landsat Thematic Mapper data for use in the Catahoula Ranger District geographic information system. Gen. Tech. Rep. SO-99. New Orleans, LA: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station. 14 p.

Hansen, M.H.; Frieswyk, T.; Glover, J.F.; Kelly, J.F. 1992. The eastwide forest inventory database: users manual. Gen. Tech. Rep. NC-151. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station. 48 p.

Miles, P.D.; Brand, G.J.; Alerich, C.L.; *et al.* 2001. The forest inventory and analysis database: database description and user's manual, version 1.0, rev. 8 (1 June 2001). St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station (http://www.ncrs2.fs.fed.us/4801/fiadb/fiadb_documentation/fiadb_chapter1.htm).

Moeur, M.; Stage, A.R. 1995. Most similar neighbor: an improved sampling inference procedure for natural resource planning. Forest Science. 41(2): 337–359.

Rack, J. 2001. Forest/nonforest classification of Landsat TM data for annual inventory phase one stratification. In: Reams, G.A.; McRoberts, R.E.; Van Deusen, P.C., eds. Proceedings, 2nd annual forest inventory and analysis symposium; 2000 October 17–18; Salt Lake City, UT. Gen. Tech. Rep. SRS-47. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station: 8–10.

Scrivani, J.A.; Wynne, R.H.; Blinn, C.F.; Musy, R.F. 2001. Phase I forest area estimation using Landsat TM and iterative guided spectral class rejection: assessment of possible training data protocols. In: Reams, G.A.; McRoberts, R.E.; Van Deusen, P.C., eds. Proceedings, 2nd annual forest inventory analysis symposium; 2000 October 17–18; Salt Lake City, UT. Gen. Tech. Rep. SRS-47. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station: 11–14.

Smith, B.W.; Vissage, J.S.; Darr, D.R.; Sheffield, R.M. 2002. Forest resources of the United States, 1997, METRIC UNITS. Gen. Tech. Rep. NC-222. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station. 127 p.

Thompson, M.T. 1998. Forest statistics for Georgia, 1997. Resour. Bull. SRS-36. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 92 p.

Wynne, R.H.; Oderwald, R.G.; Reams, G.A.; Scrivani, J.A. 2000. Optical remote sensing for forest area estimation. Journal of Forestry. 98(5): 31–36.

Zawadzki, J.; Cieszewski, C.J.; Zasada, M. 2005. Use of semi-variances for studies of Landsat TM image textural properties of loblolly pine forests. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C.; McWilliams, W.H.; Cieszewski, C.J., eds. Proceedings, 4th annual forest inventory and analysis symposium; 2002 November 19–21; New Orleans, LA. Gen.Tech. Rep. NC-252. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station. 258 p.