Distributing FIA Information onto Segmented Landsat Thematic Mapper Images Stratified with Industrial Ground Data

Tripp Lowe[1], Chris J. Cieszewski[2], Michal Zasada[3,4], and Jarek Zawadzki[3,5]

Abstract.—The ability to evaluate the ecological and economical effects of proposed modifications to Georgia's best management practices is an important issue in the State. We have incorporated tabular FIA data with Landsat Thematic Mapper satellite images and other spatial data to model Georgia's forested land and assess the area, volume, age, and site quality of two stream and road buffer regimes. Each regime included different buffer widths for perennial and intermittent streams and slope classes. We discuss here the technical details of this work.

With ever-increasing public awareness of the welfare of natural resources, land managers are under great pressure to manage our forestlands with publicly acceptable stewardship. One of the prevalent issues is maintaining and improving the quality of our streams, public lakes, reservoirs, and wetlands. As a guideline for the forestry community, the Georgia Forestry Commission has compiled a set of "common sense, economical, and effective" methods, called "best management practices" (BMPs) (GFC 1999), designed to reduce nonpoint source pollution and protect the waters. While the BMPs are now voluntary in Georgia, many forest managers in the State believe that the regulations within the riparian zones should be further expanded.

For each change considered, we must be able to evaluate the possible ecological and economical impacts. As a basis for these evaluations, we need to know the amount and type of land that will be influenced by the considered change. Here we describe the use of currently available GIS and remote sensing technologies for evaluating such potential adjustments to the BMPs and their likely impacts on the resource availability. We performed several analyses of Georgia's timberlands in which tabular FIA data were incorporated into stands derived from various types of GIS data.

Objectives

The main objective of this study was to model and estimate the cover types in buffers around water resources and roads. This included spatially explicit analysis of various types of ground measurement, GIS, and other remote sensing data and modeling the distribution of forest inventory within the buffer cover type assignments. These analyses were necessary for estimating the effects of establishing different riparian zone buffers in terms of their total area, volume, age, and site quality at the State level as well as at a FIA-unit conglomerate.

Data

We analyzed various spatial data from several different sources. Vector hydrology and county boundary data sets, and raster National Elevation Datasets (NED) were downloaded from the Georgia GIS Data Clearinghouse located at http://www.gis.state.ga.us/Clearinghouse/clearinghouse.html. The hydrology data set includes rivers, streams, and artificial flow paths through water bodies captured from USGS 7.5-minute topographic maps. The NED data contain elevation information for the county at a resolution of 30 m.

The raster "1998 landcover map of Georgia" (GAP) data set, produced by the Natural Resource Spatial Analysis Laboratory (NARSAL, Institute of Ecology, University of Georgia), was used to quickly assess landcover types throughout the entire State. The GAP data set was generated by classifying Landsat Thematic Mapper (LTM5) images captured in

[1] GIS Analyst, [2]Assistant Professor, [3]Postdoctoral Fellow, respectively, Warnell School of Forest Resources, University of Georgia, Athens, GA.
[4] Assistant Professor at Faculty of Forestry, Warsaw Agricultural University, Poland.
[5] Assistant Professor, Environmental Engineering Department, Warsaw University of Technology, Poland.

the summers and winters of 1997 and 1998 and distinguishes among 18 different general landcover types. We evaluated areas classified as "clearcut/sparse," "deciduous forest," "evergreen forest," "mixed forest," and "forested wetland." All other areas classified as something else were ignored. Also obtained from NARSAL was the "Trout Streams of Georgia" vector data set including all streams in Georgia classified as habitable for trout.

Eight LTM 7 data sets were used as a second method to evaluate land cover throughout the State. The winter scenes were captured in November of 1999 and encompass the entire State except for the extreme northwestern and southern parts. LTM 7 satellite data capture information about the Earth's surface in three visible portions, two infrared, one thermal, and one panchromatic portion of the electromagnetic spectrum. Information from the visible and infrared bands was captured at a 30-m resolution, 60 m in the thermal band, and 15 m in the panchromatic band.

Methods

The analyses conducted in this project can be classified into five separate steps. The LTM images were classified by basal area using industrial ground data in the "image classification" phase. The images in the "image segmentation" phase were converted to homogenous polygons representing either coniferous, deciduous, or mixed timber stands. FIA information was distributed to those polygons using two different methods in phase 3, "FIA information distribution." Riparian and transportation buffers were generated in the "buffer creation" phase. Hydrology buffers were created using two buffer widths that incorporate three slope classes, and transportation buffers were generated using two buffer widths. In the final "data intersection" phase, the LTM-generated polygons and buffers were combined.

Image Classification

As a first-stage sample of landcover type, the GAP data set was used to locate coniferous and nonconiferous lands. Areas that were not classified by GAP as evergreen forests were masked out and ignored in subsequent analysis. The evergreen areas in the LTM images were classified by basal area using ground-measured industrial data recorded in 1998 (Ruefenacht *et al.* 2002) and the Euclidian spectral distance (Ruefenacht *et al.* 2002) from the "low basal area" signature (eq. 1):

$$\sqrt{\sum_{i=1}^{6}(S_i - A_i)^2}$$

where,

S = the seed pixel's LTM value

A = the adjacent pixel's LTM value

i = the LTM band (excluding the thermal band 6)

Due to the limited number of large, low basal area stands in the industry data set, we derived the "low basal area" signature by first generating a "high basal area" signature by sampling LTM pixels within those stand boundaries that had a high basal area. The average pixel value for bands 1-5 and 7 were calculated to yield the high basal area signature. Assuming that the pixels that are spectrally the furthest away from this signature represent low basal area stands, we calculated the Euclidian spectral distance from the high basal area signature to all other cells. The cells with the largest spectral distance were sampled and averaged to create the low basal area signature. Once again, the Euclidian spectral distance was calculated for all evergreen cells in the data set, but this time it was generated using the low basal area signature.

Separate regression models were generated and applied to the two scenes for which ground data were available by regressing stand basal area on the average Euclidian spectral distance for the stand. See table 1 for the regression model results.

We applied these models to the adjacent LTM scenes using their areas of overlap. Using the LTM-estimated basal area values, the high basal area regions were located on the adjacent, unprocessed scenes. Following the steps described above, the Euclidian spectral distances from the low basal area signature were calculated. Random sample points were generated in the unprocessed scene's area of overlap with the modeled scene.

Table 1.—*Basal area–spectral distance regression results*

LTM Scene	Adjusted R^2	Standard Error
Path 19, Row 37	0.82	19.20
Path 19, Row 38	0.80	21.11

Table 2.—*Adjacent basal area-spectral distance regression results*

LTM Scene	Parent Scene	Adjusted R^2	Standard Error
Path 19, Row 36	Path 19, Row 36	0.82	19.20
Path 18, Row 36	Path 19, Row 36	0.80	21.11
Path 18, Row 37	Path 19, Row 37	0.88	22.49
Path 18, Row 38	Path 19, Row 38	0.89	16.23
Path 17, Row 37	Path 18, Row 37	0.93	10.67
Path 17, Row 38	Path 18, Row 38	0.86	13.01

The LTM-estimated basal area from the modeled scene and the spectral distance from the unprocessed scene were recorded at each point. We combined samples into one-unit spectral distance classes where the basal area values in each sample in the class were averaged because of the large amount of noise in the LTM data set. The averaged sample data were then used to derive similar basal area—Euclidian spectral distance regression models. These steps were then repeated for all the unprocessed scenes adjacent to those LTM scenes. The results of the regression models applied to adjacent scenes are listed in table 2.

Image Segmentation

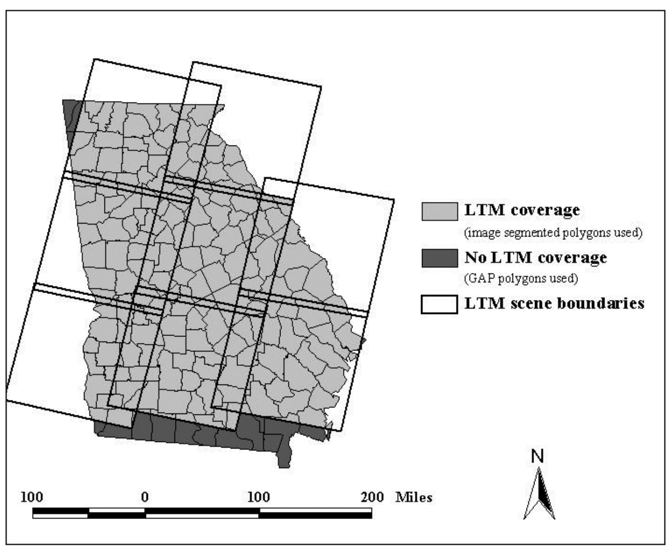
We used the GAP landcover data set as a first-stage sample of landcover type. We located on them the timbered and nontimbered areas, and ignored regions not classified as clearcut/sparse, deciduous forest, evergreen forest, mixed forest, or forested wetland. Then, we generated separate "evergreen," "deciduous," and "mixed" data sets by masking out all areas in the LTM scene that were not:

a. GAP-classified as *evergreen forest* for the evergreen LTM data set
b. GAP-classified as *deciduous forest* or *forested wetland* for the deciduous LTM data set
c. GAP-classified as *mixed forest* or *clearcut/sparse* for the mixed LTM data set.

As the second-stage sample of landcover type, the *evergreen*, *deciduous*, and *mixed* LTM data sets were then converted to homogenous polygons using an image segmentation module for ERDAS Imagine that was developed by the USDA Forest Service Remote Sensing Applications Center (Ruefenacht *et al.* 2002). This module iteratively compares the Euclidian spectral distance (Equation 1) between the first pixel in the image, the seed pixel, and the adjacent cells. If the Euclidian spectral distance between the two cells is less than or equal to the threshold value specified by the user, the pixel is assigned to the same region as the seed. The region is finished when there are no more cells adjacent to a member in the group that satisfy the threshold criteria. When the region is closed, the next seed pixel is selected and the process is repeated.

When vectorized, these image-segmented scenes produce a data set with over one million polygons, too many for ArcInfo to process in one data set. To reduce the data to a manageable size, the image-segmented LTM scenes were clipped using the county boundary. All processes described from this point on were applied on the county data sets. The image-segmented LTM data sets were then vectorized using ArcInfo's *GRIDPOLY* command. Finally, the *evergreen*, *deciduous*, and *mixed* LTM-generated polygon data sets were combined and cleaned

Figure 1.—*LTM scene boundaries.*



up by merging all polygons less than 10 acres with a larger adjacent polygon using the ArcInfo *ELIMINATE* command. GAP polygon data were substituted for those regions with no LTM coverage (fig. 1).

FIA Information Distribution

FIA information was distributed throughout the LTM-generated polygon data sets using two different methods. The first method iteratively assigned FIA information to the LTM-generated polygons using proximity to the FIA plot, current FIA plot type, and LTM-generated polygon type. The FIA data were distributed to the polygons until the sum of those polygons' acreage was within 99 percent of the FIA plot's per acre expansion factor or until there were no more polygons to fill. The data distribution criteria were applied to each FIA point before the next was applied. The criteria were as follows:

1) Assign FIA information to polygons that are within 1,000 meters of the plot location and of the same general type: evergreen, deciduous, or mixed.
2) Assign FIA information to polygons that are within 12,800 meters of the plot location and of the same general type: evergreen, deciduous, or mixed.
3) Assign information from evergreen FIA plots to evergreen or mixed polygons within 6,040 meters.
4) Assign information from deciduous FIA plots to deciduous or mixed polygons within 6,040 meters.
5) Assign information from mixed FIA plots to mixed or deciduous polygons within 6,040 meters.
6) Assign information from any FIA plot to polygons of any type within 6,040 meters.
7) Assign information from any FIA plot that has not been fully distributed to polygons of any type that have not been fully filled.

The second method iteratively assigned FIA information to the LTM-generated polygons using LTM-estimated pine basal area, polygon size, FIA volume per acre, and FIA peracre expansion factor. The evergreen polygons were first assigned the average (for the areas that fell within the stand) LTM-estimated basal area and ranked from highest to lowest; deciduous and mixed polygons were ranked according to acreage from highest to lowest, and the FIA plots were ranked from highest to lowest according to volume per acre. The FIA data were distributed to the polygons in proportions relative to polygon acreage and FIA plot per acre expansion factor. The data distribution criteria were as follows:

1) Assign information from evergreen FIA plots to evergreen LTM-generated polygons and then to mixed and hardwood polygons if more area is needed.
2) Assign information from deciduous FIA plots to deciduous LTM-generated polygons and then to any unassigned mixed and hardwood polygons if more area is needed.
3) Assign information from mixed FIA plots to mixed LTM-generated polygons and then to any unassigned deciduous or evergreen polygons.

Buffer Creation

Following the criteria set forth in Georgia's BMPs, the hydrography, GAP, and NED data sets were used to generate riparian zone buffers that incorporated two buffer widths and three slope classifications. A total water mask was generated by first buffering the hydrology data set. Streams and rivers classified as perennial and those streams classified as trout streams in NARSAL's trout data set were buffered on each side by 15 feet, creating a 30-foot primary stream mask. The intermittent streams were buffered on each side by 12.5 feet, creating a 15-foot secondary stream mask. Those bodies of water contained in the GAP data set, all cells classified as "open water" or "coastal marsh," were vectorized and merged with the existing stream water masks to create the total water mask polygon data set. All areas within this mask were considered to be water.

Percent slope data sets were generated from the elevations contained in the NED data. Following the BMP guidelines for slope class, they were reclassified into three classes: 1) slight (< 20 percent), 2) moderate (20 – 40 percent), and 3) steep (> 40 percent). The slope data were then vectorized and incorporated into the water masks by intersection (fig. 2). This process produced a data set containing the polygons in the original water mask cut into smaller pieces where it and the slope data meet.

Transportation buffers were generated using similar methods. Cells classified as "transportation" in the GAP data set were buffered on both sides by 40 feet and 100 feet, excluding the "transportation" cells.

Figure 2.—*Stream buffer LTM polygon intersection.*

Data Intersection

The LTM-generated polygons were incorporated into the hydrology and transportation buffers by intersection (fig. 2). This process produced data sets similar to those created when the hydrology masks were intersected with the reclassified slope data. The final intersected data contain the buffer polygons cut into smaller pieces where they and the LTM-generated polygons meet. Attribute information from both data sets is stored for each data set, as well.

Results

Landsat Thematic Mapper satellite images and other spatial and tabular data were combined to assess the area, volume, age, and site quality of Georgia's riparian zone timberlands. LTM-derived Euclidian spectral distance–basal area models were calibrated using industrial ground data and applied statewide using LTM scene overlap. Homogenous evergreen, deciduous, and mixed forest polygons generated from the LTM images were populated with FIA tabular data using two different methods. The first method was based on the proximity to the FIA plot. Polygons, of the same type as the FIA plot, were assigned data, starting with the closest and proceeding away until the FIA per acre expansion factor-based criteria were met. The other method assigned FIA data to the polygons using ranking criteria. Evergreen polygons were ranked according to LTM-estimated basal area, hardwood and mixed forest polygons

Table 3.—*Intersection data set descriptions*

Data Set Name	Description
Primary hydro/proximity	Primary hydrology buffer intersected with landcover containing FIA information assigned using the proximity criteria
Primary hydro/rank	Primary hydrology buffer intersected with landcover containing FIA information assigned using the ranking criteria
Secondary hydro/proximity	Secondary hydrology buffer intersected with landcover containing FIA information assigned using the proximity criteria
Secondary hydro/rank	Secondary hydrology buffer intersected with landcover containing FIA information assigned using the ranking criteria
100-foot road/proximity	100-foot road buffers intersected with landcover containing FIA information assigned using the proximity criteria
100-foot road/rank	100-foot road buffers intersected with landcover containing FIA information assigned using the ranking criteria
40-foot road/proximity	40-foot road buffers intersected with landcover containing FIA information assigned using the proximity criteria
40-foot road/rank	40-foot road buffers intersected with landcover containing FIA information assigned using the ranking criteria

according to acreage, and FIA data by plot volume per acre measures. The highest ranked polygons were assigned information from the highest ranked FIA plots (of the same type), and so on. Riparian zone buffers were generated using two different buffer distances incorporating three different slope classes. Similar transportation buffers were generated, buffering GAP-classified roads by 40 feet and 100 feet. LTM-generated polygons and riparian and road buffers were incorporated by intersection (fig. 2). Eight intersection data sets were generated (table 3), each containing the buffer polygons split where they and the landcover polygons meet.

Literature Cited

Georgia Forestry Commission. 1999. Georgia's best management practices for forestry. Georgia Forestry Commission.

Ruefenacht, B.; Vanderzanden, D.; Morrison, M.; Golden, M. 2002. New technique for segmenting images. RSAC Program Manual. U.S. Department of Agriculture, Forest Service, Remote Sensing Application Center. 14 p.