1997

Methodology for Integrating Aerial Photography and LANDSAT TM Imagery for Inventory of Forest Land Cover

Chris W. Bennett
University of Arkansas at Monticello

Robert C. Weih Jr.
University of Arkansas at Monticello

Recommended Citation
Bennett, Chris W. and Weih, Robert C. Jr. (1997) "Methodology for Integrating Aerial Photography and LANDSAT TM Imagery for Inventory of Forest Land Cover," *Journal of the Arkansas Academy of Science*: Vol. 51 , Article 6.
Available at: http://scholarworks.uark.edu/jaas/vol51/iss1/6

A Methodology for Integrating Aerial Photography and LANDSAT TM Imagery for Inventory of Forest Land Cover

Chris W. Bennett
Project Program
Spatial Analysis Laboratory
School of Forest Resources
University of Arkansas
Monticello, AR 71656
(Correspondant)

Robert C. Weih
Director
Spatial Analysis Laboratory
School of Forest Resources
University of Arkansas
Monticello, AR 71656

Abstract

Forest cover for 7.25 million acres (2.93 million hectares) in southeastern Georgia was characterized for the years 1988 and 1994 with the intent of assessing the efficacy of remote sensing procedures for broad scale forest inventory. Landsat-5 Thematic Mapper digital satellite scenes of seven spectral bands were obtained for winter and summer of each year and were analyzed as two separate 14-band multi-temporal images. Images were geo-referenced to the universal transverse mercator (UTM) coordinate system prior to classification. Spectral classification with the ISOCLUSTER algorithm produced 250 categories. Color infrared aerial photographs were mapped to the digital imagery and were used to convert spectral categories to land cover features. For this study, land features of interest were limited to water, marsh, pine forest, hardwood forest, mixed pine/hardwood forest, urban, and where distinguishable, clearcut and agriculture. Accuracy assessment techniques indicated very good consistency.

Introduction

The practice of forest inventory is playing an increasingly important role in resource planning as demands for fiber increase upon finite resources. Traditional field inventory techniques are becoming harder to justify economically. In addition, the accuracy of field inventory is often less than optimal; surveys are usually less than one percent of area. Remote sensing techniques have the potential to make frequent, inexpensive, and useful inferences about very large areas. This manuscript reports on stage one of a two-stage project to characterize forest vegetation based on spectral signatures, i.e., reflectance of solar irradiance in specific narrow wavelengths, as a reasonably accurate map of major forest types. The purpose of this manuscript is to describe two aspects of the procedure: 1) the practical application of integrating aerial photography with industry standard digital image classification techniques for large areas; and 2) a method of assessing the accuracy of the resultant map.

The study area for this project, 7.25 million acres (2.93 million hectares) in southeastern Georgia (Fig. 1), was selected primarily because of the availability of detailed digital inventory data. In addition, the physiography of the region has several important features for remote sensing of forest vegetation.

(1) Negligible topographic variation exists. Shadows, various angles of light incidence, and spatial distortions inherent in high-relief areas are typically problematic in their effect on sunlight reflected from surface features, espe-
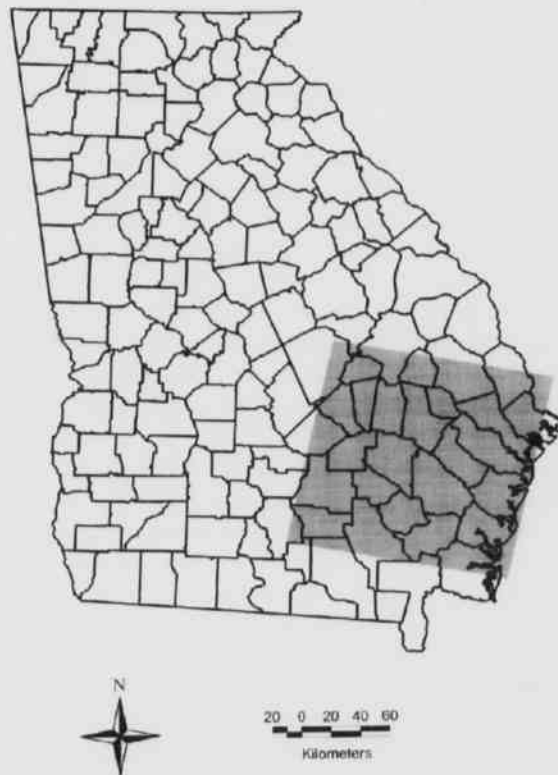

Fig. 1. Study area is in gray. Georgia counties are outlined in black.

cially forest vegetation which is typically highly variable.

(2) Large areas of managed pine forests are common. Because southern pines, particularly *Pinus taeda* and *P. elliotii*, figure prominently in North American supply of wood fiber, it is important that spectral signatures of various stages of pine plantations be adequately captured.

(3) Georgia ranks third among southeastern states in volume of sawtimber produced (FIA Homepage, 1997). Characterization of forest resources for a highly productive region of the southeastern United States has a greater value than for a less productive region.

Southeastern Georgia physiography is typified by slightly rolling hills to flat topography. Sandy soils in the area, though well drained, are usually low in fertility and this is a driving factor in the effort to improve wood fiber yields through genetic selection and intensive cultivation (Keefer, 1994). Agricultural crop production competes with forestry for land use in the western portion of the study area, but poorer soils in the east favor pines.

Because forest vegetation extent was of primary importance in this phase of the project, a limited set of land class features were selected for final categorization. These features are water, marsh, pine (*Pinus* spp.) forest, hardwood (deciduous species) forest, mixed pine/hardwood forest, urban/roads, and where possible to distinguish, clearcut and agriculture.

Materials and Methods

Data and Software.--Four dates of Landsat-5 Thematic Mapper (TM) digital imagery were selected based on their extent of significant cloud cover during dormant and peak growing seasons. The following image dates were used: February 26 and June 17, 1988; January 9 and May 17, 1994. All seven available spectral bands were acquired; they correspond to three visible bands, one near-infrared, two mid-infrared, and one thermal band (Lillesand and Kiefer, 1987).

Geometric correction of the images was performed on Arc/Info version 7.0 (from ESRI, Inc.) Image analysis was performed with ERDAS IMAGINE, version 8.2 (from ERDAS, Inc.). Both software packages operated on a Sun SPARCstation 20 workstation running Solaris 2.4. IMAGINE utilizes the ISOCLUSTER unsupervised classification algorithm. In addition, the software provides a very convenient graphical interface which facilitates the coding of spectral features. All digital maps were projected in the Universal Transverse Mercator coordinate system, Zone 17.

Aerial photographs were obtained from the U.S. Geological Survey National Aerial Photography Program (NAPP) for a subset of the study area, approximately 17%. Photographs were selected to correspond with the dates of the Landsat imagery. Due to the periodic latency in the flight schedule, color infrared photographs were available only for 1988. Only black and white panchromatic photographs were available for 1994.

Mitigation of Cloud Effects.--For this study, "significant" cloud cover denotes degradation of ground surface-reflected sunlight such that demarcation of ground feature categories is not feasible. Fortunately, only the summer-1994 image had any appreciable cloud cover. However, it was indeed significant; approximately 15% of the image was treated as cloudy.

Detection of the extent of clouds was facilitated by the use of the thermal infrared image band (commonly referred to as band 6). The cooler temperatures of the clouds contrasted distinctly with the ground surface features. Cloud temperatures registered as low pixel values in band 6. By plotting a histogram of these values, a breakpoint could be determined that represented the change from cloud to cool ground features such as water. Removing pixels having values lower than this threshold removed most of the cloud effect. Visual inspection of the remaining image indicated that the clouds were adequately removed. However, this left a portion of the image that could not be classified with winter and summer TM scenes together. To mitigate this problem, enclouded ground areas were extracted from the unclouded winter image and classified separately. The two separate classified images were later combined into a single image to complete the analysis.

Georeferencing.--Landsat images may be purchased in several degrees of geometric correction. Although the image vendor, Space Imaging EOSAT, will provide geometric correction to several map projections, the images often do not match features such as roads and waterways as effectively as when they are fit to a specific area. Therefore, geometric correction was the first step involved in preparing the images for analysis. In addition, the TM images were received with pixel resolutions of 25 x 25 m, and during the georeferencing operation they were converted to 30 x 30 m resolution through a process called nearest-neighbor resampling. This resampling process works by using the value of the pixel in the original image nearest to the relocated pixel coordinate as the new value. Compared to other methods that average nearby values, the nearest neighbor approach results in a blockier appearance but preserves the original sensor values without creating new values from averaging.

The method of georeferencing the four images used in this study was to match the first image to a standard base network (e.g., roads), then match the remaining images to the first corrected image. This image-to-image registration reduces potential errors that might be incurred if all images were referenced to the base network. The reason is that the base network is not part of the analysis procedure. Rather, it provides a standard for the first image, which then becomes the standard for the remaining images. If all images were

Chris W. Bennett and Robert C. Weih

registered to the base network then, supposing a maximum error of one pixel, all images could be displaced from the base network equilaterally and the maximum error between all images would be two pixels. However, if all remaining images are registered to the first image, the maximum error is one pixel. In fact, localized errors can be substantially more than one pixel due to inaccuracies of the base network itself. In this study, the U.S. Geological Survey's 1:100,000 scale Digital Line Graph road network was used as the base network. Image to image registration was performed in IMAGINE.

Arc/Info provides a convenient interface for selecting corresponding points on the satellite image and a digital dataset, such as a road network. When a suitable number of correlated points have been selected, parameters for a first-order polynomial transformation are computed. The resulting equation is applied to affect three image characters-translation, scaling and rotation--that cause the image to match the known dataset, roads in this case. This highly interactive process can take several days to perform properly. Ideally, each pixel in each of the four images should correspond to the same spot on the ground. Of course, not all pixels will line up exactly. For this study, on average, image pixels were within one-half pixel width of their true location, i.e., RMS error in meters was 13.3 in the east-west direction and 13.1 in the north-south direction.

Classification.--Multispectral images are arrays of digital numbers (DN) representing surface reflectance of light in various wavelength bands. For example, the Landsat TM datasets are composed of seven coregistered images called bands. Each pixel location (e.g., row 1, column 1) contains a number from 0 to 255 (eight binary bits) representing reflectance of a particular narrow-band range of the light spectrum for a 25 x 25 m surface area. One feature of satellite imagery that is important to consider in image classification is the concept of a sensor's instantaneous field of view (IFOV). The IFOV is the surface area composing the light received by one sensor element or pixel. Note that IFOV defines the resolution of the image. The digital number representing the reflectance of a given pixel in the image is a weighted average of the respective reflectance of all features in the IFOV. Ideally, all pixels should fall completely upon a single feature, but in fact, many pixels capture light from several features, resulting in what is known as a mixed pixel. Mixed pixels contribute significantly to confusion between features when classifying digital imagery.

All spectral classifiers attempt to define spectral signatures of surface features. Spectral signatures are represented by the collection of DNs for different wavelengths associated with a surface feature. These values sometimes may be averaged or otherwise ciphered, and they sometimes may incorporate covariance between bands or between features. Generally, the signatures are considered to be multidimensional quantities.

Among standard spectral classifiers, two basic methods exist (Wilkie and Finn, 1996); supervised and unsupervised. The important difference is in the way spectral signatures are developed. Supervised methods require prior knowledge of the study area inasmuch as a human analyst selects known portions of the images from which to develop spectral signatures. In this case, a polygon may be digitized around a feature class causing, for example, 10 pixel locations to be acquired for that signature, resulting in 10 seven-dimension values. These polygons are called training areas because they are used to "train" the classification algorithm regarding that feature. By collecting several representative training statistics, the analyst hopes to characterize all such features over the extent of the image. Confidence values could then be applied to each pixel in the image according to its fit with the "known" spectral signature.

Unsupervised methods, on the other hand, attempt to designate signatures based on natural "clumpiness" of the spectral dataset. Consider Fig. 2 which depicts a hypothetical dataset of three bands, green, red, and infrared. Some features may be well defined, whereas others may be barely distinguishable. Still others may not readily fall into any conceivable class. Often, similar classes can be made distinct by including another spectral band. However, a drawback of unsupervised classification is that it produces only spectral classes; it does not produce feature classes. Consequently the analyst must investigate all output classes to determine their relation to feature classes. Furthermore, the burden of assigning confidence values lies completely with the analyst.
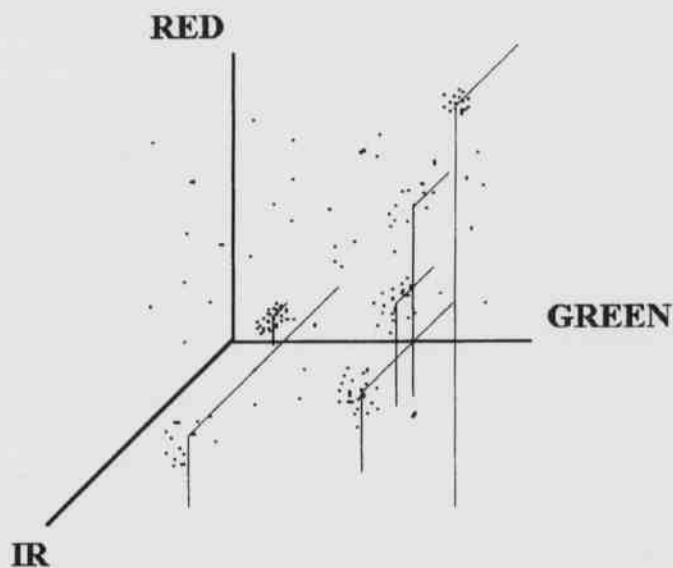

Fig. 2. Hypothetical clusters in three spectral bands. Lines are drawn to midpoints of clusters.

The IMAGINE software includes the ISODATA algorithm for unsupervised classification. A detailed explanation is included here to clarify the need for steps that follow. Execution of the program is controlled by several user-supplied parameters: maximum number of classes (NBINS), maximum number of iterations (MAXITER), and percent convergence threshold (CONV%). Assuming a seven band image, the operation of this algorithm is as follows (ERDAS, 1994).

1) Define NBINS classes evenly dividing the spectral space. Basically this means 255/NBINS for each band. These class values will change during execution of the program. Note that these class values would denote seven-value coordinates in spectral space.

2) Working through each pixel location in the image, compute the distance from each pixel's spectral coordinate to each NBINS value according to the modified Euclidean distance formula

$$D = \sqrt{\Sigma \ (d_i - C_i)^2}$$

where D = distance from segment i; i = band number from 1 to seven; d = digital number of band i; and C = class i from step 1. The pixel is temporarily assigned to the class having the shortest distance to the pixel. This step is repeated for all pixel locations in the image.

3) Collect all pixel values in each tentative class and average the pixels in each band. This becomes the new n-coordinate class value. Repeat this procedure for all classes; then begin step 2 again with the new class values.

4) During the iteration of steps 2 and 3, many pixels will change classes causing the means of the class values to change. The rate of change eventually slows, and the algorithm stops when either the number of classes denoted by CONV% no longer changes, or MAXITER has been reached.

When all the images in this study were transformed to the UTM (zone 17) coordinate system, the two images for each year were combined into a single file, one per year, to facilitate application of the algorithms. Thus, ISOCLUSTER was executed on one 14-band image for 1988 and 1994, respectively, and one seven band image for the enclouded area of the 1994 image. The computational procedure took five days per image on the SPARCstation 20 to create output images composed of 250 values. The enclouded area, representing only 15% of the study area, was processed with NBINS=50 due to time constraints. Recoding time is a function of NBINS, not area.

Obviously, this is many more classes than needed for the final map. However, two or more distinct features will sometimes occupy a class regardless of how many initial divisions are used. For example, swamp and marsh are often subject to this confusion due to the effect of surface water in the understory. This fact is less problematic with more classes than with fewer classes.

One can think of the output image as a reduction of practically an infinite number of potential classes ($2^{8*14}-1$) to NBINS classes, 250 in this case. Consequently, a given class will be scattered throughout the image, associated with whichever ground features happen to have similar reflectance. The feature coding process involves highlighting each class individually, verifying the associated ground feature, and recoding the numeric class to a descriptive class. The process of recoding a class requires the analyst to view as many of the given class's associated ground features as possible. This is why aerial photographs are so valuable. Actual ground visitation at a similar level of detail would be prohibitively expensive. However, some ground visitation is important to resolve ambiguities and spot check estimates from photos.

Integrating aerial photography for feature coding.-- Ordinarily, field forays may supplement the use of aerial photography as ancillary information. However, due to the availability, quantity, timeliness, and low cost of color-infrared aerial photography versus the cost of trips to the study area, in this study nearly all feature coding was accomplished through aerial photographs. Some ground reconnaissance was performed prior to classification whereby the authors drove the highways of the study area and spot-checked a hardcopy of the satellite imagery. In addition, for this study, land managers substituted for detailed field forays in "reality checking" the final classifications.

To facilitate the time consuming task of finding on the aerial photograph the spot that corresponds with a point on the computer-displayed map, outlines of the aerial photos were digitized in a form that could be displayed against any of the images. Fortunately, IMAGINE allows a large number of image viewing screens to be displayed simultaneously. In addition, these viewers may be geographically referenced, such that information displayed in one viewer is accessible in other viewers. Therefore, a useful layout incorporates views of summer and winter TM images, the classified image, boundaries of aerial photographs, and any necessary magnifications of these constituents. Another feature of IMAGINE is that any given class can be displayed separately or highlighted. This makes it much easier to find a given pixel on the aerial photographs.

Ultimately, however, the task of verifying each of the classes is time consuming due to potential ambiguity within classes. Mixed pixels create the most difficult ambiguities because the pixel assumes spectral characteristics of several features and this can give it the appearance of something completely different. Note that a relatively high degree of proficiency in interpreting aerial photography is required since it is used so extensively.

Accuracy Assessment.--Errors in interpretation, incomplete data, or confusion of feature classes will result in errors in the final classified map. Therefore, some method of assessing accuracy must be implemented. The general prin-

Chris W. Bennett and Robert C. Weih

ciple of accuracy assessment is to compare estimated feature values for pixels to known feature values for the same pixels. Normally, one cannot test every pixel so a percentage of pixels is sampled. Known feature values can be obtained in a number of ways, e.g., aerial photographs, ground visitation, expert knowledge, or other ancillary data. Aerial photographs can provide good reference data, but some pixels will be difficult to visually co-register.

Congalton (1988) noted several variations on simple random sampling, including stratifying based on land cover and stratifying geometrically, cluster sampling, and systematic unaligned sampling. He concluded, however, that the simple random sampling or stratification by land classes performed best for agriculture and forested areas in that study.

IMAGINE provides an interface for assessing the accuracy of classified images. The approach is similar to that recommended by Congalton (1988) because one can select simple random sampling, stratification based on class value, or equalized random sampling (i.e., each class is assigned an equal number of samples). In addition, one can define a minimum number of points per class, total number of sample points, and some degree of homogeneity of surrounding pixels via a "majority window. " This last issue is implemented by considering the pixels in a window of n x n pixels (where n is greater than two) surrounding the candidate sample. If the values of the pixels are the same for some user-defined threshold, then the candidate pixel is retained. The purpose of such a threshold is to control the number of pixels in areas of high diversity as such areas are more likely to contain misclassified pixels due to the larger number of mixed pixels.

Digital outlines of the photographs were used to eliminate areas of the classified image not actually verifiable. Thus a "clipped" image was created corresponding to the coverage of the aerial photographs and the accuracy assessment was performed on this subset (Figure 3). Table 1 presents the results for the accuracy assessment for the classifications of the 1988 and 1994 images, respectively. Generally, the classes for 1988 and 1994 are within 10 percentage points of each other, except for the water class. Apparently, the majority window restricted the number of possible samples in the already small population of inland water such that fringe pixels had a higher likelihood of being selected. This result warrants future attention.

Discussion

Much theoretically oriented literature exists regarding the accuracy of spectral classification (Fenstermaker, 1994). However, a rigorous analysis of the classification accuracy was not attempted for this report. The purpose of the classification accuracy assessment as presented herein was to provide a meaningful indication as to the efficacy of the resul-
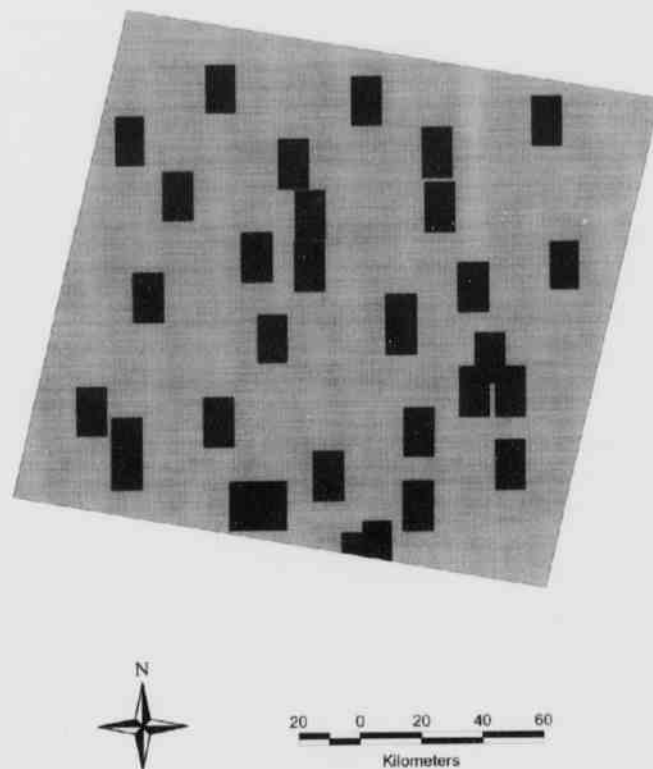

Fig. 3. Aerial photograph outlines (black) were used to delineate a subset of the study area (gray).

tant map for later stages of this research. Notwithstanding, in the authors' experience (Weih et al., 1993; Thomasson et al., 1994), the percentages in Table 1 are very agreeable for such a large and diverse area. In addition, spot checks of the 1988 classification by personnel experienced with the area indicate similar agreement. Therefore, it appears that the use of aerial photography as a verification medium for unsupervised classification can be a viable alternative to more expensive field checking. Future research in this area could investigate specific combinations of spectral bands for efficiency in classifying various vegetation types, development of spectral curves for vegetation communities and individual species, and the practical application of theoretical accuracy assessment concepts to production-oriented image classification endeavors.

Literature Cited

Congalton, R.G. 1988. A Comparison of Sampling Schemes Used in Generating Error Matrices for Assessing the Accuracy of Maps Generated from Remotely Sensed Data. *Photogrammetric Engineering and Remote Sensing.* 54(5):593-600.

Table 1. Percent of samples deemed correct from accuracy assessment.

	1988	1994
Water	30.9	100.0
Marsh	93.3	100.0
Pine Forest	90.7	89.6
Mixed Forest	54.4	67.7
Hardwood Forest	70.3	85.7
Urban/Roads	100.0	90.0

ERDAS, Inc. 1994. ERDAS Field Guide. 3rd ed. *ERDAS. Inc.. Atlanta. GA.* 628 pp.

Fenstermaker, L.K. ed. 1994. Remote Sensing Thematic Accuracy Assessment: A Compendium. *American Society for Photogrammatric Engineering and Remote Sensing.* 413 pp.

FIA Homepage. 1997 April 3. "Southern Region Forest Inventory and Analysis. " [Online] *USDA Forest Service.* URL: http://www.srsfia.usfs.msstate.edu/doc/lainfo.htm.

Keefer, Brent. 1994. Personal communication with the Manager of Technical Services for Rayonier, Inc. Southeastern Division.

Lillesand, T.M. and **R.W. Kiefer.** 1987. Remote sensing and image interpretation. 2nd ed. *Wiley. New York.* 721 pp.

Thomasson, J.A., C.W. Bennett, B.D. Jackson, and **M.P Mailander.** 1994. Differentiating Bottomland Tree Species with Multispectral Videography. *Photogrammetric Engineering and Remote Sensing.* 60(1):55-59.

Weih, R. R. Zakaluk, R. Pearson, T. Gress and **K. Gilmore.** 1993. Integrating Space Technology into Forest Management. Proceedings. *The World Space Congress.* Pergamon Press. 13(11):65-69.

Wilkie, D.S. and **J.T. Finn.** 1996. Remote Sensing for Natural Resources Monitoring: A Guide for First-Time Users. *Columbia Universitv Press, New York.* 295 pp.