FOREST/NON-FOREST STRATIFICATION IN GEORGIA WITH LANDSAT THEMATIC MAPPER DATA

William H. Cooke

ABSTRACT.—Geographically accurate Forest Inventory and Analysis (FIA) data may be useful for training, classification, and accuracy assessment of Landsat Thematic Mapper (TM) data. Minimum expectation for maps derived from Landsat data is accurate discrimination of several land cover classes. Landsat TM costs have decreased dramatically, but acquiring cloud-free scenes at optimum seasons for vegetation discrimination is still problematic. FIA plot locations determined from hand-held GPS units can vary ± 5-20 m. Landsat pixels can also vary ± 25 m. These spatial inaccuracies restrict the use of pixels on feature edges and decrease the usefulness of plots that have split conditions. Current research at the USDA Forest Service's, Southern Research Station involves aggregating forest types in the lab based on field plot measurements of dominant, co-dominant, and intermediate trees. We believe this methodology is most appropriate for tying FIA field plot data to the satellite imagery. We are testing methodological approaches for image processing that can satisfy the dual goals of repeatability and timeliness.

INTRODUCTION

Typically, remote sensing efforts at the Southern Research Station (SRS) of the USDA Forest Service have focused on large area estimates of forested and non-forested lands. Proportions of forested and non-forested lands within pixels of Advanced Very High Resolution Radiometer Data (AVHRR) have been predicted using high resolution Landsat Thematic Mapper (TM) data in a multiple regression scenario (Zhu and Evans 1994). To date, Forest Inventory and Analysis (FIA) managers have asked remote sensing analysts What can you do for me? with respect to rapid large area analysis for simplistic land cover conditions. I believe that our partners and cooperators in the Southern Annual Forest Inventory System (SAFIS) want more than delineation of forested from non-forested lands and attendant acreage calculations. At a bare minimum, we should be able to discriminate among broad land cover classes including pine, hardwood, scrub, grass, cultivated, and inert. As a remote sensing analyst, my question to FIA is What can you do for me? Or, how can plots taken under an annual inventory system be used to train and validate remotely sensed data to produce useful maps? Remote sensing efforts that benefit FIA should extend well beyond Phase I estimates for stratification. Remote sensing also plays an important role as a tool for providing timely information on natural disasters and for getting information about forest conditions in inaccessible areas. These applications of remote sensing should not be overlooked for funding.

BACKGROUND

Achieving the goal of providing the land cover classes of interest requires classification of large amounts of TM data with FIA field plots serving as the basis for classification and verification of those classes. Acquisition of large amounts of TM data is much less costly now that Landsat 7 has been successfully deployed. Full-scene (185 km x 185 km) costs have decreased from $4,000 to $600 per scene. However, acquiring cloud-free scenes at optimum seasons for vegetation discrimination is still problematic. Also problematic are the radiometric differences between adjacent scenes. Figure 1 illustrates the radiometric differences that exist among four full TM scenes in the Piedmont of Georgia.

Investigation by Zhiliang Zhu (U.S. Geological Survey, EROS Data Center, pers. comm.) indicates that normalizing the radiometric components of adjacent TM scenes before

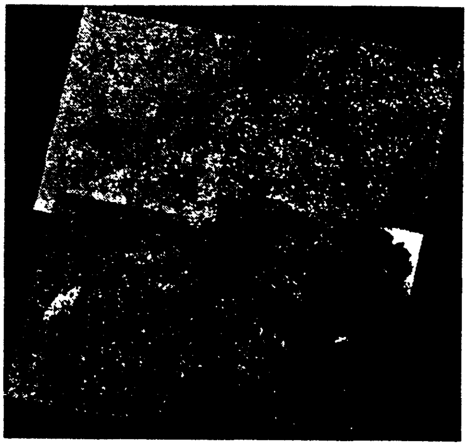

Figure 1.—*Radiometric differences for four winter scenes in Georgia.*

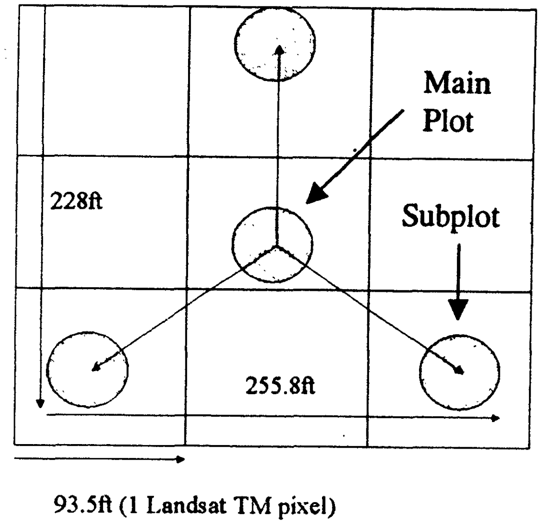

Figure 2.—*Comparison of the FIA plot design with Landsat TM data.*

classification can result in up to a 50 percent loss in reflectance characteristics important for automated classification efforts. At SRS, classification is done before scenes are mosaicked together. Although this approach maximizes classification accuracy on a scene-by-scene basis, it can result in a discontinuity of classification results between adjacent images. Higher per scene accuracy results from this methodology, but a more visually pleasing map product results from pre-classification scene normalization and mosaicking. Ultimately, there is a trade off between higher map accuracy versus more aesthetically pleasing map products. Managers and data users should be educated about this aesthetic problem if maximum information content is the desired outcome.

METHODS

Scientists at SRS are currently examining the usefulness of FIA plot variables for training and for verifying Landsat TM imagery. SAFIS inventories employ Global Positioning System (GPS) receivers to acquire geographic coordinates for the center plot of the four-plot cluster design (Rockwell Avionics 1996). Figure 2 illustrates how this four-plot cluster compares to a nine-pixel window of Landsat TM imagery.

On the surface, the correspondence between the FIA plot design and Landsat TM data appears conveniently located within this nine-pixel window. But misregistration of the imagery and sources of error in the GPS measurement require a model more like the one shown in figure 3. Each subplot falls close to a pixel center, but GPS coordinates are accurate to ± 5 m to ± 20 m depending on averaging techniques, time of acquisition, number of satellites acquired, and overhead "line-of-sight" (Rockwell Avionics 1996). Each TM pixel's registration to its "real-world" location is assumed to be about ± 1 pixel (28.5 m) in relatively flat terrain, possibly greater in steeply dissected terrain. Misregistration errors can be additive in a worst-case scenario (20 m + 28.5 m = 48.5 m). The reality is that the main plot falls somewhere within a 25-pixel window. The full

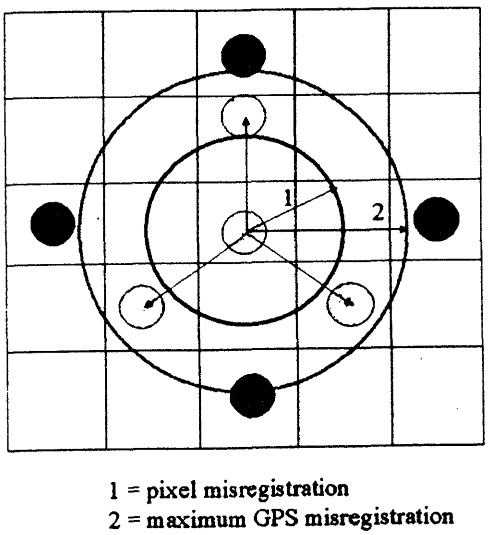

Figure 3.—*Pixel and GPS misregistration problems.*

cluster of four plots falls within a 7 x 7 pixel window or greater. This translates into roughly an 11-ac (4.4 ha) ground area. This inherent "slop" in location makes training of automated classifiers and accuracy assessment procedures more difficult. These spatial registration problems will likely restrict the use of pixels on feature edges and limit the potential usefulness of plots that have split land cover conditions. The possibility of deriving an edge class is being investigated by SRS scientists. Classification techniques being used for wall-to-wall TM efforts in Georgia are variations on methods used by Coppin and Bauer (1994) and Cooke (1991).

Classification Techniques

1. Stratify TM scenes by physiographic/ ecological condition.
2. Use National Wetlands Inventory data to mask wetlands.
3. Use Census data to mask high-density urban areas.
4. Allow low-density urban areas to be classified.
5. Use differential highway masks.
6. Use edge detection spatial filtering algorithms to locate and eliminate some edge pixels before classification.
7. Classify the data using these TM channels:
   a. Raw data channels 3, 4, and 5
   b. 1st Principal Component
   c. Brightness and Greenness components of the Kauth-Thomas transformation
   d. Ratio of channels 3 and 4 (NDVI).
8. Classify 75 classes using unsupervised techniques to reduce class variance.
9. Aggregate in classes (Pine, HW, Brush, Inert, at a minimum), then iteratively reclassify if necessary.
10. Aggregate classes by following methods developed by Linda Garnett for her Master's Thesis.
11. Use a 5 x 5 majority scan to filter out "salt and pepper" pixels.
12. Assess accuracy/refine classifications for areas > 25 pixels using FIA plots.
13. Assess supervised classifications for accuracy with FIA plots for cross validation.

RECOMMENDATIONS

Work is needed to determine which plot variables are appropriate for training and verifying TM classifications. Current efforts involve aggregating from individual tree data for dominant and co-dominant species. We believe that the satellite "view-from-above" makes these crown classes most likely to provide useful information for modeling land cover. Plot-level variables like forest type are subject to field-level interpretation. A forest type designation calculated in the lab from the dominant and co-dominant members of a stand is more likely to be representative of crown reflectance in TAM data and is easily reproducible in the lab.

LITERATURE CITED

Cooke, W.H. 1991. **Use of temporal Landsat data for discrimination of selected forest species composition group**. Mississippi State University. M.S. thesis.

Coppin, P.R.; Bauer, M.E. 1994. **Processing of multitemporal Landsat TM imagery to optimize extraction of forest cover change features**. IEEE Transactions on Geoscience and Remote Sensing. 32(4): 918-927.

Rockwell Avionics. 1996. **Precision Lightweight GPS Receiver PLGR+96 and PLGR FEDERAL Precise Positioning Service (PPS)**. Operations and Maintenance Manual.

Zhu, Z.; Evans, D.L. 1994. **U.S. forest types and predicted percent forest cover from AVHRR data**. Photogrammetric Engineering and Remote Sensing. 5: 525-531.

ABOUT THE AUTHOR

William Cooke is a Research Forester, Southern Research Station, USDA Forest Service, Forest Inventory and Analysis program, Starkville, MS.