# FORESTED LAND COVER CLASSIFICATION ON THE CUMBERLAND PLATEAU, JACKSON COUNTY, ALABAMA: A COMPARISON OF LANDSAT ETM+ AND SPOT5 IMAGES

**Yong Wang, Shanta Parajuli, Callie Schweitzer, Glendon Smalley, Dawn Lemke, Wubishet Tadesse, and Xiongwen Chen**[1]

**Abstract—**Forest cover classifications focus on the overall growth form (physiognomy) of the community, dominant vegetation, and species composition of the existing forest. Accurately classifying the forest cover type is important for forest inventory and silviculture. We compared classification accuracy based on Landsat Enhanced Thematic Mapper Plus (Landsat ETM+) and Satellite Pour l'Observation de la Terre (SPOT5) images for three land cover types (mixed oak forest, mixed hardwood forest, and agricultural) of the Cumberland Plateau, Jackson County, northern AL. The overall accuracy was 67 and 71 percent based on Landsat ETM+ and SPOT5 images, respectively. The most obvious commission error (misclassifying into wrong categories) was caused by mixed hardwood forest using SPOT-5 image and mixed oaks forest using Landsat ETM+ image, each was about 35 percent. The high omission error (omitting from correct categories) was associated with SPOT-5 data for the mixed hardwood and mixed oak forest. The low accuracy is typical for areas dominated by deciduous forest. Future research needs to explore the possibility of incorporating other GIS data such as variables derived from digital elevation model to improve the classification accuracy.

## INTRODUCTION

Forest cover classifications focus on the growth form (physiognomy) of the community, dominant vegetation, and species composition. The information is often used for forest inventory, sustainable management of forest resources, and conservation of biodiversity associated with the forest. Accurate classification of forest cover types can help forest resource managers to maker better decisions. Traditionally, forest cover classifications and mapping have been done by interpreting aerial photos or ground surveys. Forests are often complex; vary by topographic, edaphic, and climatic conditions; and are under constant change because of natural and human disturbances. Traditional forest classifications and mapping are time-consuming and cost-intensive. Over the last twenty years geographic information system (GIS) and remote sensing date have become important tools to generate digital maps and database of current forest types.

It has been demonstrated that visual and digital analysis based on Landsat Enhanced Thematic Mapper (Landsat ETM+) images (30 m resolution) could yield land cover maps useful for forest management purposes (Apan 1997, Sotomagor 2002). Although successful in many instances, forest cover classification based on Landsat ETM+ images still presents several difficulties particularly with complex topographic landscapes and among hardwood forest cover types. While the sun and viewing angles can be considered constant within an image, the topographic characteristics of the terrain may change the illumination geometry, affect spectral signatures of a cover type, and cause classification errors in the spectral classification (Holben and Justice 1980, Civo 1989). To address this problem, Madden (2003) used the elevation, slope, and aspect generated from Digital Elevation Model (DEM) to assist the vegetation classification based on Landsat ETM+ data, and achieved 75 percent classification accuracy; Fahsi and others (2000) used same technique and found classification accuracy increased by incorporating DEM data. Classification of deciduous hardwood forest cover types has shown difficulty with Landsat ETM+ data because the dominant deciduous tree species of different forest often have similar spectral signatures (Jensen 2002, Schriever and Congalton 1995). Czaplnewski and Patterson (2003) found that there was a geometric increase in the error rate as the number of forest strata in the classification system increased.

Recent development and availability of high resolution satellite image from Satellite Pour l'Observation de la Terre (SPOT) provides an opportunity to extract more ground information that was not extracted by LANDSAT. An important factor limiting classification accuracy at higher levels of detail is the spatial resolution of the sensor system used. According to Jensen (2004), typically, sensors such as LANDSAT could be successfully used for classification at Anderson Level I (forest vs. non-forest) and classification at Anderson Level II (evergreen forest, deciduous forest, and mixed forest) requires higher resolution sensors like SPOT5 multispectral (10 m resolution). The SPOT satellites' unique features, variable viewing geometry, stereo imaging, and frequent revisit capability provide a flexible platform for capturing imagery on request and opportunities to get more detailed information of the land cover at a specific time period. Using SPOT data, Williams (1992) found that the accuracy of classifying 16 non-forest, 6 forested, and 6 other land cover types of the Peter Lougheed Provincial Park of the Kananaskis Valley in southwest Alberta, Canada was improved compared to classifications based on LANDSAT data. However, May and others (1997) found that LANDSAT data was more effective than SPOT data in separating shrubs

[1]Professor and Graduate Student, Center for Forestry, Ecology, and Wildlife, Alabama A&M University, Normal, AL; Research Forester and Emeritus Scientist, Southern Research Station, USDA Forest Service, Normal, AL; Assistant Professor, Center for Forestry, Ecology, and Wildlife, Alabama A&M University, Normal, AL; Associate Professor, Center for Hydrology, Soil Climatology and Remote Sensing, Alabama A&M University, Normal, AL; and Research Associate, Center for Forestry, Ecology, and Wildlife, Alabama A&M University, Normal, AL, respectively

*Citation for proceedings*: Stanturf, John A., ed. 2010. Proceedings of the 14th biennial southern silvicultural research conference. Gen. Tech. Rep. SRS-121. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 614 p.

from meadows, but neither LANDSAT nor SPOT data were effective for separating meadow types.

The land cover of Cumberland Plateau region of northern AL is dominated by oak-hickory hardwood forest with mixed oak forest occurring above the escarpment and mixed hardwood forest occurring mainly below the escarpment (Smalley 2003). The landform is complex and varies both in elevation and aspect. These features of the study area suggest that reliable classification of land cover for this area could be difficult to achieve using remotely sensed data and no such study exists for this area. In this study, we attempted to classify the land cover of two locations at Jackson County, AL. Our specific objectives were to (1) classify the land cover using remotely sensed data and (2) compare the accuracy of land cover classifications based on LANDSAT TM and SPOT5 images.

## METHODS

### Study Area

This research focused on Jackson County of northern AL. We selected two sites in the northern Jackson County: the Hytop (34°56'30"N, 86°04'00"W) and Estill Fork (34°58'30"N, 86°12'30"W) tracts (fig. 1) both within the strongly dissected southern sub-region of the Mid-Cumberland Plateau Ridge (Smalley 1982). The region has temperate climate characterized by long, moderately hot summers, and short, mild winters due to the region's proximity to the Gulf of Mexico. The mean temperature for the region is about 13 °C. Precipitation is heavy throughout the year with some periods of prolonged droughts (Smalley 1982).

### Remote Sensing Data Pre-processing and Classification

We acquired Landsat ETM$^+$ images of October 20, 2003 and SPOT-5 multispectral image of October 18, 2005. The images were first geo-referenced by identifying ground control points on each image and on the topographic map used as a reference map. The images were further georeferenced using digital orthophotographs with six reference points such as roads, crossroads, and waterways that were identified on both sources. The final images had rooted mean square error (RMS) < 50 percent of the pixel size. The image was then referenced to Universal Transverse Mercator (UTM) projection (Zone 16), NAD 83 coordinate system. The supervised maximum likelihood classification algorithm was used to separate the land cover to three major land covers: mixed oak forest, mixed hardwood forest, and agriculture (including pastures) based on signatures from *in situ* ground cover data and aerial photographs. Earth Resource Data Analysis (ERDAS) Imagine 8.7 software was used for image data pre-processing and classifications.

### Accuracy Assessment

The accuracy assessment was accomplished by comparing the classification results based on the images with the land cover type collected in the field. First, ninety-eight random points were first generated on the classified image and their geographic locations (longitude and latitude) were recorded. These random points were then located in the field using global position system (GPS) and the land types were identified. Deciduous forests were classified as mixed oak forest when oaks contributed $\geq$ 80 percent total basal area and mixed hardwood forest otherwise (modified from Smalley 1982). The land cover types of these random points from the field were compared with their classification type from image analysis, and an error matrix was then generated to assess the accuracy level (Rosenfield and Fitzpatrick 2001). Overall accuracy, producer accuracy, and user accuracy were calculated (Jensen 2004). Overall accuracy is the probability of correct classification of the image with respect to the reference data. The probability that a sample from the classified image actually represents that class in the reference (field) data is the producer accuracy. The probability that a reference sample is correctly classified by imagery analysis is the user accuracy. We also used the Kappa coefficient to assess the agreement between the classifications generated based on images and from field survey. The Kappa coefficient (Bishop and others 1975) is based on

$$K = [N\sum x_{ij} - \sum(x_{i+} \times x_{+i})]/[N^2 - \sum(x_{i+} \times x_{+i})] \qquad (1)$$

where K is the Kappa coefficient, N is the total number of pixels in the error matrix, $X_{ij}$ is the number of observations in row i and column j, and $X_{i+}$ and $X_{+i}$ are the marginal total of the error matrix table for row i and column j, respectively.

## RESULTS AND DISCUSSION

### LANDSAT Classifications

The overall classification accuracy was 67 percent based on the Landsat ETM$^+$ images (table 1). The agricultural land had the lowest user accuracy (50 percent) and the highest producer accuracy (100 percent), which indicates that the agriculture lands could be misclassified as forests, but forests were never misclassified as agriculture land. Producer accuracy was similar between the mixed oak forest (72 percent) and mixed hardwood forest (70 percent). The highest user accuracy was for the mixed hardwood forest (77 percent) compared to 64 percent of mixed oak forest, suggesting that mixed oak forest was more likely to be misclassified to mixed hardwood forest. Kappa coefficient was 49 percent for land covers combined, was the highest for mixed hardwood forest (55 percent), and the lowest (44 percent) for the agriculture land cover. The mixed hardwood was the most abundant cover type (4 402 ha) in the study area followed by mixed oak forest (table 2).

### SPOT-5 Classification

The overall classification accuracy was improved to 71 percent (table 1) based on SPOT-5 images. However, the producer accuracy decreased for both forest covers, and did not change for agriculture land. User accuracy was increased for mixed oak forest (from 64 to 72 percent), decreased for agriculture land (from 100 to 82 percent) and mixed hardwood forest (from 77 to 65 percent). Overall Kappa coefficient was 53 percent. The Kappa coefficient was almost doubled for agriculture land, increased 4 percent for mixed oak forest, and reduced 16 percent for mixed hardwood forest

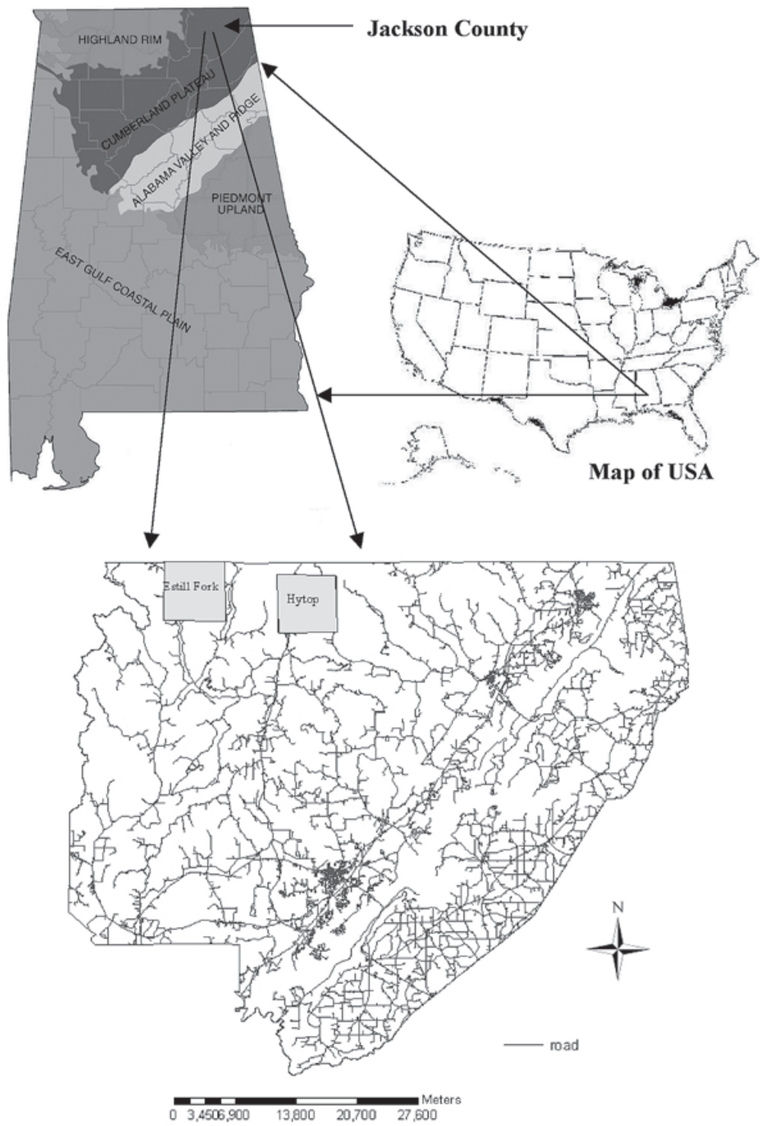


Figure 1—Hytop and Estill Fork study sites in the northern Jackson County, Alabama, USA.

**Table 1—Accuracy of the land cover classification based on Landsat $ETM^+$ and SPOT5 images of Cumberland region of Jackson County, AL**

| | Landsat $ETM^+$ | | | | SPOT5 | | | |
|---|---|---|---|---|---|---|---|---|
| Class | PA[1] | UA | Kappa coefficient[2] | OA | PA | UA | Kappa coefficient | OA |
| Mixed oak forest | 72 | 64 | 46 | 67 | 65 | 72 | 50 | 71 |
| Mixed hardwood | 70 | 77 | 55 | | 67 | 65 | 39 | |
| Agriculture and pasture | 100 | 50 | 44 | | 100 | 82 | 78 | |

[1] PA is the producer accuracy, UA is the user accuracy, and OA is the overall accuracy in percentages.
[2] Kappa coefficient measures the agreement between the classifications based on remotely sensed images and the reference points from the field, higher values indicate greater agreement.

**Table 2—Total area calculated for each land cover type using Landsat ETM+ and SPOT5 for the study sites of the Cumberland region of Jackson County, AL**

| | Landsat ETM+ | | SPOT5 | |
|---|---|---|---|---|
| Class | Area (ha) | Percent | Area (ha) | Percent |
| Mixed oak forest | 3790 | 44 | 3641 | 43 |
| Mixed hardwood | 4402 | 52 | 4673 | 55 |
| Agriculture and pasture | 327 | 4 | 205 | 2 |

compared to the classification from Landsat ETM+ data. The results suggest that with the higher resolution of SPOT-5 image, the agricultural lands and mixed oak forest were more likely to be accurately identified while the accuracy for mixed hardwood forest was lower compared to the classification based on Landsat ETM+. The area estimated based on SPOT5 image decreased from 327 ha to 205 ha (a 37.3 percent reduction) for agriculture land and increased from 4402 ha to 4673 ha for mixed hardwood forest (a 5.8 percent increase) (table 2).

## CONCLUSIONS

The use of SPOT5 images for classifying land cover of the Cumberland Plateau of Jackson County, AL, a landscape dominated by deciduous hardwood forest, improved classification accuracy compared to the classification based on Landsat ETM+. However, the accuracy of the classification based on Landsat ETM+ and SPOT5 data was relatively low (about 70 percent) and below the Anderson criterion (80 percent) for image application. This is typical for areas dominated by deciduous forest (Jensen 2002, Schriever and Congalton 1995). Hardwood forests are difficult to distinguish because of similar vegetation components and hence, the spectral similarity (Jensen 2002). The most obvious commission error (misclassifying to wrong categories) was caused by mixed hardwood forest using SPOT-5 image and mixed oaks forest using Landsat ETM+ image, each was about 35 percent. The high omission error (omitting from correct categories) was associated with SPOT-5 data for the mixed hardwood and mixed oak forest.

We classified forest type based the criteria of mixed oak forest (forests with $\geq$ 80 percent oaks) and mixed hardwood (forest with < 80 percent oaks). According to Smalley (1982), mixed oak forests contain primarily white oak (*Quercus alba* L.), scarlet oak *(Q. coccinea* Muench.*),* southern red oak (*Q. falcata* Michx.), black oak (*Q.velutina* Lamarck), chestnut oak (*Q. prinus* L.), and have associations with hickories (*Carya spp*.), black gum (*Nyssa sylvatica* Marsh.), red maple (*Acer rubrum* L.), shortleaf pine (*Pinus echinata* Mill.), and Virginia pine (*P. virginiana* Mill.); they occupy the drier sites on top of the Plateau including ridges above the base level of the plateau and on the upper warm escarpment slopes. In places, shortleaf and Virginia pines are prevalent on upper warm escarpment slopes perhaps reflecting a fire history. Mixed hardwoods (i.e., greater percentage of species other than oaks) are on the more moist sites on top of the plateau, in stream channels, on cool slopes above the base level of the plateau, and on the warm upper escarpment slopes (Smalley 1982).The lower escarpment slopes are sometimes an Eastern redcedar (*Juniperus virginiana* L.)-hardwood mixture. Under natural environmental conditions, there are gradations between all of these vegetation types, reflecting the variations in geophysical features such as elevation, slope, and relief. This could result in the errors of our classification with Landsat ETM+ and SPOT5 images.

Managers of Southern United States forests are under increasing pressure to balance the economic, social, and ecological aspects of the resource. Meeting contemporary demands for healthy forests as well as forest products depends on increasing productivity while protecting the environment and sustainability of the forests. The accurate inventory of different forest covers in a timely manner is critical. Remote sensing and GIS-based classification such as those from this study can provide quick and relatively inexpensive mapping and quantitative estimation of forest covers. Further study will explore the possibility of incorporating other GIS data such as those variables derived from digital elevation model for image analysis to improve the classification accuracy.

## ACKNOWLEDGMENTS

We thank the Southern Research Station of USDA Forest Service, National Science Foundation, and the School of Agriculture and Environmental Science of Alabama A&M University for providing funds for this project. Doug

Clendenon and Joe Cardinski of USDA Nature Resource Conservation Service provided technique assistance and some digital data used in this study. Daryl Lawson and Ryan Sisk assisted the ground truthing of the classification result. We also thank Drs. Luben Dimov and Zachary Felix for reviewing the draft of this manuscript.

## LITERATURE CITED

Apan, A. 1997. Land cover mapping for tropical forest rehabilitation planning using remotely sensed data. International Journal of Remote Sensing. 18: 1029-1049.

Bishop, Y.M.M.; Fienberg, S.E.; Holland P.W. 1975. Discrete Multivariate Analysis: Theory and Practice. The MIT Press, Cambridge, MA: 557 p.

Civo, D.L. 1989. Topographic normalization of Landsat thematic mapper digital imagery. Photogrammetric Engineering and Remote Sensing. 55: 1303-1309.

Czaplenwski, R.L.; Patterson, P.L. 2003. Classification accuracy for stratification with remotely sensed data. Forest Science. 49: 402-408.

Fahsi, A.; Tsegaye, T.; Tadesse, W. [and others]. 2000. Incorporation of digital elevation models with Landsat TM data to improve land cover classification accuracy. Forest Ecology and Management. 128: 57-64.

Holben, B.N.; Justice, C.O. 1980. An examination of spectral band rationing to reduce the topographic effect on remotely sensed data. NASA/Goddard Space Flight Center, Washington, DC. 28 p.

Jensen, J.R. 2002. Remote Sensing of the Environment: An Earth Resource Perspective. 2nd ed. Prentice Hall, New Jersey: 545 p.

Jensen, J.R. 2004. Sensing of the Environment: An Earth Resource Perspective. 3rd ed. Prentice Hall, New Jersey: 545 p.

Madden, M. 2003. Vegetation modeling, analysis and visualization in United States National Parks. Photogrammetric Engineering and Remote Sensing. 65: 171-177.

Rosenfield, G.H.; Fitzpatrick, L.K. 2001. A coefficient of agreement as a matter of thematic classification accuracy. Photogrammetric Engineering and Remote Sensing. 52: 223-227.

Schriever, J.R.; Congalton, R.G. 1995. Evaluating seasonal variability as an aid to cover type mapping from Landsat Thematic Mapper data in the northeast. Photogrammetric Engineering and Remote Sensing. 61: 321-327.

Smalley, G.W. 1982. Classification and evaluation of forest sites on the Mid-Cumberland Plateau. Gen. Tech. Rep. SO-38. U.S. Forest Service, Southern Forest Experiment Station, New Orleans, LA.

Smalley, G.W. 2003. Ecological land classification maps of United States Forest Service research study areas on Coastal Lumber Company properties. A contract report. U.S. Forest Service, Southern Research Station, Asheville, NC.

Sotomagor, A.I.T. 2002. A spatial analysis of different forest cover types using GIS and remote sensing techniques. Ph.D. dissertation. Forest Science Division, University of Netherlands. 270 p.

Williams, J.A. 1992. Vegetation classification using Landsat TM and SPOT-HRV imagery in mountainous terrain, Kananskis Country, South West Alberta. Forest Science. 44: 427-435.