FOREST LAND CLASSIFICATION USING ISOCLUSTERING PROCEDURE: AN EXPLORATORY ANALYSIS

Yong Wang, Shanta Parajuli, Dawn Lemke, Xiongwen Chen
Center for Forestry and Ecology
Alabama A&M University, Normal, AL 35762

Wubishet Tadesse
Center for Hydrology, Soil Climatology and Remote Sensing
Alabama A&M University, Normal, AL 35762

Callie Schweitzer and Glendon Smalley
Southern Research Station, USDA Forest Service
Alabama A&M University, Normal, AL 35762

ABSTRACT

Forest land classification is widely used by forest managers to subdivide forest land to logical segments (landtypes), each having similar productivities for the particular timber species. Forest land classification is often based on variables such as physiography, climate, geology, topography, soil, and vegetation. In the past, forest land classification and mapping have often relied on the expert experience and manual processes. We tested the possibility of classifying forest landtype associations and landtypes in the Mid-Cumberland Plateau of Jackson County of northern Alabama based on georeferenced data layers using geographic information system (GIS) and remote sensing approaches. We tested the consistency of the computer generated forest land classification maps with the base map developed by an expert of forest land classification. We first separated the forest land into two landtype associations (A and C) that occurred at the study area. We then used an isoclustering method to classify land types within each association with five variables: elevation, slope, aspect, soil texture, and soil types. We used 7 classes for Association A and 11 classes for Association C. The resultant map of land types was fragmented with many small polygons. We merged the polygons less than 2 ha with their adjacent dominant land classes. Although the degree of fragmentation decreased and the land classes showed some similarities with the classification from the expert, the agreement was low. We believe that this low agreement is because the isoclustering technique is an unsupervised classification technique which depends on natural breaks among the input variables and disregards information of relative importance of each variable with respect to forest management. We recommend using rule-based or other knowledge-based approaches for forest land classifications.

KEYWORDS. Site quality, productivity, GIS, Mid-Cumberland Plateau, remote sensing.

In Prisley, S., P. Bettinger, I-K. Hung, and J. Kushla, eds. 2006. Proceedings of the 5th Southern Forestry and Natural Resources GIS Conference, June 12-14, 2006, Asheville, NC. Warnell School of Forestry and Natural Resources, University of Georgia, Athens, GA.

INTRODUCTION

Managers of U.S. southern 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 and efficiencies while protecting the environment and sustainability of the forests. The existing knowledge about forest resources and their potential is inadequate in many aspects for the increasingly complex management challenges that we now face. Forest site quality is used to describe the inherent productive capacity of the land for certain tree species, often used as an index to the amount of timber that the forest can produce. Ever increasing demands for sustainable forest practices require better measurements of current productivity, future potential, and physiographic variables that affect forest productivity. Forest productivity is also an important component of forest ecosystem services, which interacts with other components such as soil, water, wildlife, and vegetation components in the system.

Forest site refers to the area in which trees of a stand grow; the environment of a site determines the type and quality of the vegetation the area can carry (Avery and Burkhart, 2002). Forest site (land) classification subdivides forest land into logical segments, each with similar productivity. Classification helps forest managers to make management decisions by identifying the potential productivity of forest stands, both present and future, and providing a frame of reference for land management diagnosis and prescription (Avery and Burkhart, 2002).

It is possible to measure site productivity based on factors such as soil nutrients and moisture, temperature regimes, available light, topography, and so on. Although attempts at direct measurement of site productivity have been made, such an approach may not be of immediate value to the practicing forester; consequently, indirect estimates of forest site are frequently used (Avery and Burkhart, 2002). Methods developed to measure forest site quality include volume measurements, plant indicators, and the use of height growth (Groninger et al., 2000). A most commonly used technique for quantify site quality in the United States is the species specific site index. Site index is estimated by determining the average total height and age of dominant and codominant trees in well-stocked and even-aged stands. The relationship between height and age is used for interpolating site index (height at a specified age). There are several limitations of site index including: stand age is often difficult to determine; not well-suited for uneven-aged stands or area of mixed-species composition; effects of stand density are not considered; and other variables such as disturbance and climate variation are often not considered. Smalley (1982, 1986, 1991, 2003) classified and evaluated forest sites on the Southern Cumberland Plateau for the management of commercially valuable tree species. He developed a land classification system that enables forest managers to subdivide forest land into logical segments (land types) with approximately equal productivity, allow them to rate productivity, and manage the forests to the limitations and hazards that the land types impose on forest management activities. His assessment was based on ecology, soils, site features, and yields and often extrapolated from adjacent regions. McNab et al. (2003) explored the possibility of classifying forest productivity based on species composition in western North Carolina and found that using indicator species for site classification has promise in hardwood stands as it avoids problems associated with conventional methods based on site index.

The overall objective of this study is to classify the forest land using geographic information system (GIS), remotely sensed data, and statistical modeling approaches (Parajuli, 2006). Differing from forest cover type classifications which focus on the existing or past vegetation based on overall growth form (physiognomy) of the community and its dominant vegetation and species composition, forest land classification is often based on physiography, climate, geology, topography, soil, and vegetation. Forest land classification enables forest managers to subdivide forest land into logical segments (landtypes), each having similar productivity; these classifications help to recognize potentials, limitations, and hazards that the landtypes impose on forest management activities (Smalley, 1982). These classifications provide a foundation for modeling forest productivity in the Mid-Cumberland region of northern Alabama and for creating an information system that could be used to predict forest productivity and model the spatial and temporal changes of forests in the study area. This system also has the potential to be linked to other biodiversity databases such as wildlife and plant communities to provide a conservation tool for forest and natural resource managers. This is multi-phase research with its initial goals of developing models and techniques to classify forest cover type and forest land. We report here the forest land classification results based on an isoclustering approach.

METHODS

Study Area

The Southern Cumberland Plateau includes all of the Cumberland Plateau east and south of the Tennessee River. Covering all or portions of 18 counties in Alabama, three in Georgia, and two in Tennessee, it occupies about 24,030 km^2 and extends south to north from about 33° 15' N to 35° 7' N and east to west from about 85° 20' W to 87° 50' W. It extends southwest from Chattanooga, Tennessee, nearly to Tuscaloosa, Alabama, a distance of about 280 km. Included in the region are such well-known land-forms as the Warrior Coal Basin and Lookout, Sand and Little Mountains. Our research focused on Jackson County of northern Alabama. We selected two sites in northern Jackson County (Figure 1). The Hytop (34^o 56’ 30” N, 86^o 04’ 00” W) and Estill Fork (34^o 58’ 30” N, 86^o 12’ 30” W) tracts are both in the strongly dissected southern sub-region of the Mid-Cumberland Plateau Ridge. They fall under the Northern Cumberland Plateau section (221H) under the National Hierarchical Framework of Ecological Units (NHFEU) (Bailey et al., 1994; McNab and Avers, 1994). 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^{o}C. Precipitation is heavy throughout the year with some periods of prolonged droughts (Smalley, 1982).

GIS Data Layers

Smalley (2003) manually classified the forested land at the two study sites. The hardcopy map was scanned, georeferenced, and digitized to create a GIS digital map at the 1:24,000 scale. Biogeophysical properties such as soil types, soil physical properties, slope, and elevation collected from secondary sources were used as input in ArcGIS 9.0 (Environmental Systems Research Institute, 2004). The ancillary GIS data compiled include soil type, soil texture, available water supply, runoff and drainage extracted from Soil Survey Geographic Database (USDA Natural Resources Conservation Service, 2003) and slope, elevation, aspect, curvature, and surface profile extracted from United States Geological Survey (USGS, 2006) 10 m Digital Elevational Models (DEM). The SSURGO database is the digital soil association map developed by USDA Natural

Resources Conservation Service (NRCS). It is designed primarily for use by farmers, landowners, and county natural resource planners and managers.

Isoclustering Classification

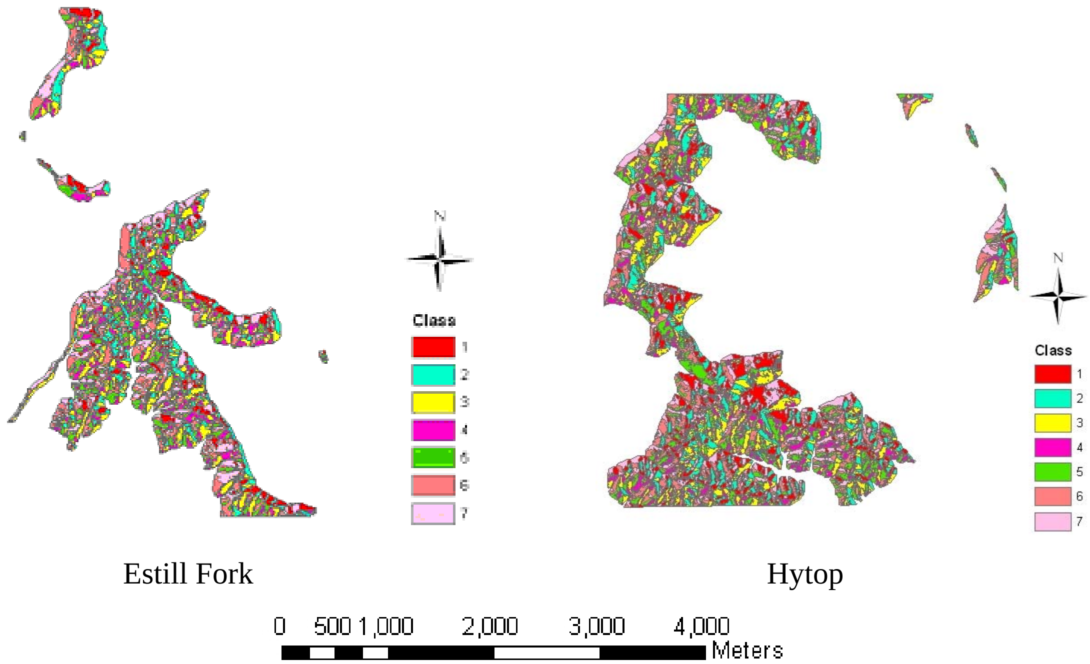
The isoclustering procedure uses an ISODATA (Iterative Self-Organizing Data Analysis Techniques) clustering algorithm for defining natural groupings of data points in attribute space. It is a commonly used algorithm in satellite image classification in which spectral signatures from multiple wavebands (equivalent to attributes) are used to determine classes. In our case, we used the five defined attributes for each pixel to determine class compositions. This technique was used to identify patches which were similar with regard to five variables: soil type, soil texture, slope, aspect, and elevation using the Spatial Analyst feature in ArcGIS. We selected these five variables after examining the classification criteria used by Smalley (2003). Isoclustering is a nonhierarchical clustering procedure similar to the K-mean clustering widely used for grouping subjects based on specific features. With isoclustering we specified 7 and 11 land classes for Association A and C, respectively. The polygons of the SSURGO-derived data were converted to raster grid cells and pixel size was set to 10 m, the resolution of a final clustered map. These five variables were used as input classification variables for cluster analysis. The clustering procedure divides all pixels into many discrete, separate clouds of cells as requested, separating them into cluster bins of pixels with similar combinations of values of the input variables. Once these cluster units (land unit signatures) were developed, a maximum likelihood discriminate analysis was used. The maximum likelihood classifier uses the cluster means and covariance matrices from the signature to determine to which land classes each pixel has the highest probability to belong (Jensen, 2004). The result from isoclustering was mapped and compared to the classification map manually created by Smalley (2003) for the same sites.

RESULTS AND DISCUSSION

The Smalley (1982, 1986, 1991, 2003) classification system is a hierarchical algorithm; it divides land to region, subregion, and land type association, and landtype, with landtype as the smallest land unit. In the Mid-Cumberland region, Smalley identified three subregions including true plateau, Walden ridge, and a strongly dissected portion. The study sites were located in the true plateau subregion with two landtype associations. Landtype Association A is the land with weakly dissected plateau surface, and Landtype Association C is the strongly dissected margins and sides of the plateau. Specifically, Association A is the forest land above the escarpment with 7 landtypes and 1,631 ha, and Association C is below the escarpment with 11 landtypes and 6,888 ha. Smalley (2003) suggested that the major feature separating the two associations was the elevation.

The initial isoclustering process separated Association A to 7 classes and Association C to 11 classes. The landscape map (Figure 1) based on these classifications was fragmented, and the boundaries of landtype polygons were difficult to define compared to Smalley's. The maps based on the results of the isoclustering were less well defined for Association A than for Association C. The fragmentation was caused by the nature of isoclustering procedure, which divides the land into different classes based on natural breaks of the input variables and does not merge polygons of small area. To reduce the fragmentation, we merged polygons that had an area less than 2 ha (5 acres) to the surrounding polygons based on their shared boundaries. The final land class maps were less fragmented. The size and the boundaries of the new land unit polygons were better defined and more meaningful for forest land management purpose (Figure 2).

Association A



Association C

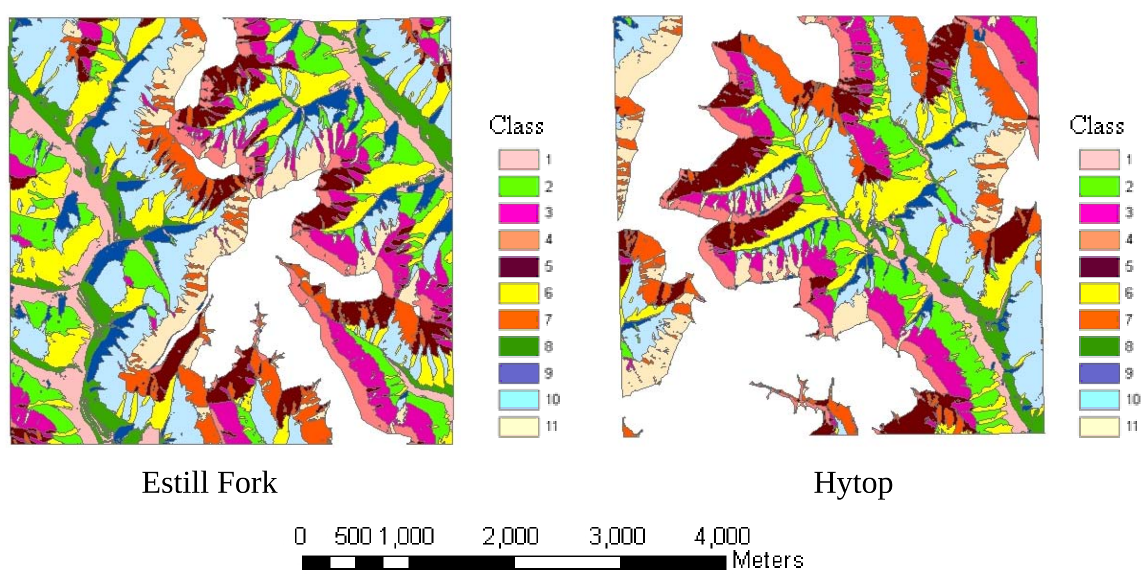


Figure 1. Land classes produced by isoclustering procedure for Landtype Association A (above) and C (below) before merging polygons ≤ 2 ha for Estill Fork and Hytop in Jackson County, AL.

Elevation, slope, and aspect were the three most important variables for separating the land classes in Association A. All land classes in this association had similar soil series and soil texture except for land class 3 which was characterized by Hartsell and Colbert series with fine sandy loam texture. This indicated that soil series and soil texture were two of the major factors differentiating class 3 from the others. Class 1 was characterized with higher slope and west to northwest aspect, indicating that slope and aspect were the important variables to separate this land class from the others. Similarly, class 2 had the lowest slope and elevation among the classes in Association A, indicating that the elevation and slope were the important factors characterizing class 2. Class 5 was dominated by forest land facing north. In Association C, class 4 was characterized by the highest elevation range with sandy loam soil among the classes. A low elevation and slope range were found for classes 1 and 8 indicating slope and elevation were most important for these two classes than for others. Between these two classes, class 1 was characterized with silty clay soil texture and faced mainly to north, while class 8 was characterized by Ceda and Barfield soil series with fine sandy loam soil and facing south. Some agreement between isoclustering classes and Smalley's landtypes were found in both Associations; however, the overall agreement was low based on the visual inspections.

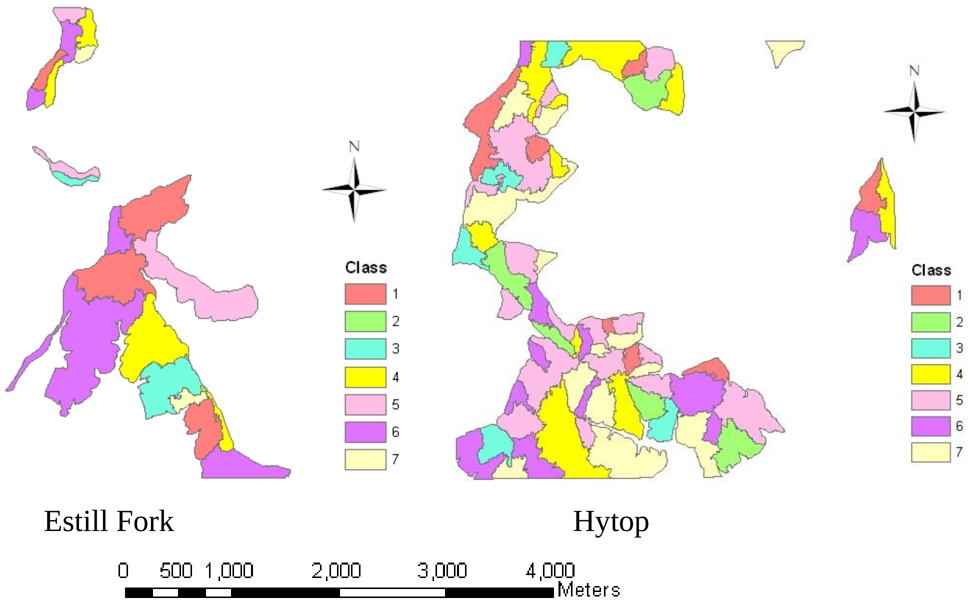
CONCLUSIONS

Isoclustering is an unsupervised classification procedure widely used for classifying land cover. The main difference between isoclustering process and the Smalley classification is the input of expert knowledge. While isoclustering technique separates subjects into different groups based on the natural breaks within the input variables, it ignores the relative importance of each input variable in the real world. In our particular case, the specific objective is to separate forest land into relatively uniform units so that forest productivity could be estimated and similar management practices may be applied for a specific unit. Physiographic variables such as the ones used in this study including elevation, slope, aspect, soil type, and soil texture are likely to differ in relative importance for forest productivity and also may change by timber species and geographic locations. For example, elevation and aspect are two important variables that could have deterministic effect on the species composition and growth rate for the forest stands at our study site. Because the isoclustering process is not explicit, it also creates the difficulties in pin-pointing features that separate different classes. Isoclustering benefits include its objective process and lack of human perception biases. Isoclustering can be applied to large areas to separate land subjectively and produce classification quickly. The results from isoclustering may provide information for rule-based or supervised classification systems.

ACKNOWLEDGEMENTS

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

Association A



Association C

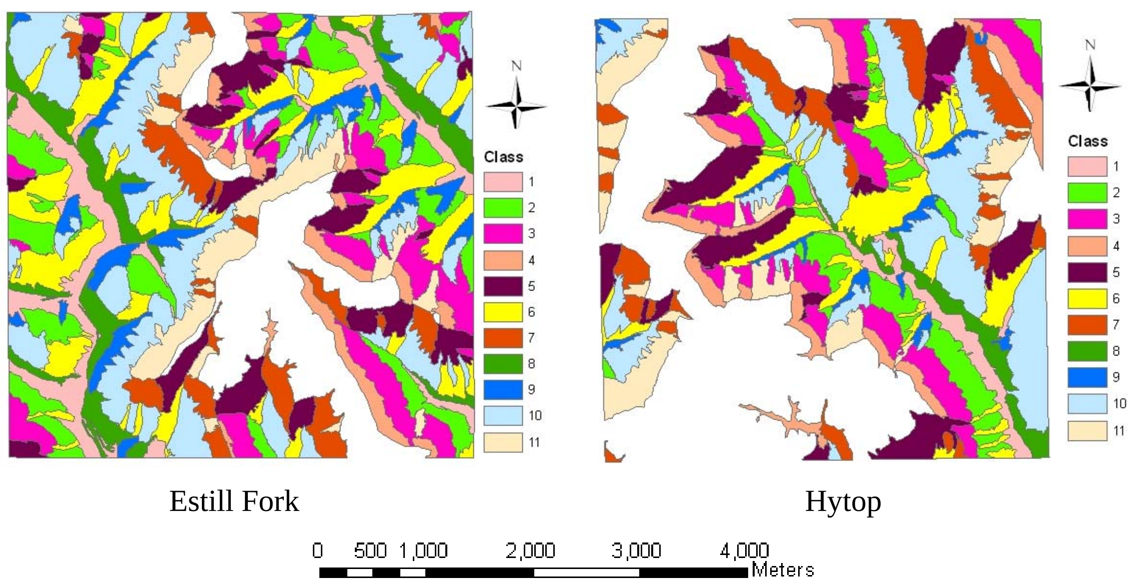


Figure 2. Land classes produced by isoclustering procedure for Landtype Association A and C after merging polygons ≤ 2 ha for Estill Fork and Hytop in Jackson County, AL.

References

Avery, T.E., and H.E. Burkhart. 2002. *Forest Measurements*. *5th Ed*. New York: McGraw-Hill.

Bailey, R.G., P.E. Avers, and T. King. 1994. Ecoregions and subregions of the United States (map). W. H. McNab and R. G. Bailey eds. Washington, DC: USDA Forest Service.

Environmental Systems Research Institute. 2004. *ArcGIS version 9.0*. Redlands, CA: Environmental Systems Research, Inc.

Groninger, J.W., W.M. Aust, M. Miwa, and J.A. Stanturf. 2000. Growth predictions for tree species planted on marginal soybean lands in the Lower Mississippi Valley. *Journal of Soil and Water Conservation*. 55: 91-95.

Jensen, J.R. 2004. *Remote Sensing of the Environment: An Earth Resource Perspective*. *3rd ed.* Upper Saddle River, NJ: Prentice Hall.

McNab, W.H., and P.E. Avers. 1994. Ecological subregions of the United States: section description. WO-WSA-5. Washington, D.C.: USDA Forest Service.

McNab, W.H., F.T. Lloyd, and D.L. Loftis. 2003. Preliminary evaluation of methods for classifying forest site productivity based on species composition in western North Carolina. In *Proceedings of the Southern Forest Mensturationists' Conference*, 10-18, P. F. Doruska and D. C. Bragg, eds. Chattanooga, TN: Arkansas Forest Resource Center.

Parajuli, S. 2006. Forest type mapping and forest land classification of northeastern Alabama using remote sensing and geographic information system. MS Thesis. Normal, AL: Alabama A & M University, Department of Plant and Soil Science.

Smalley, G.W. 1982. Classification and evaluation of forest sites on the Mid-Cumberland Plateau. General Technical Report SO-38. New Orleans, LA: USDA Forest Service, Southern Forest Experiment Station.

Smalley, G.W. 1986. Site classification and evaluation for the Interior uplands: forest sites of the Cumberland Plateau and Highland Rim/Pennyroyal. Technical Publication R8-TP9. Atlanta, GA; USDA Forest Service, Southern Forest Experiment Station and Southern Region.

Smalley, G.W. 1991. No more plots, go with what you know: developing a forest land classification system for the Interior Upland. In *Ecological Land Classification: applications to identify the productive potential of southern forests*, 48-58, D.L. Menge and D.T. Twe, eds. General. Technical Report SE-68. Asheville, NC: USDA Forest Service, Southeastern Forest Experiment Station.

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

USDA Natural Resources Conservation Service. 2003. *Soil Survey Geographic (SSURGO) Database*. Washington D.C.: USDA Natural Resources Conservation Service.

USGS. 2006. *Digital Elevation Model (DEM) database*. Washington D.C.: US Geological Survey.