Journal of the Arkansas Academy of Science

Volume 62 Article 19

2008

Land-Use/Land-Cover Characterization Using an Object-Based Classifier for the Buffalo River Sub-Basin in North-Central Arkansas

Robert C. Weih Jr.
University of Arkansas at Monticello, weih@uamont.edu

D. White Jr.
University of Arkansas at Monticello

Land-Use/Land-Cover Characterization Using an Object-Based Classifier for the Buffalo River Sub-Basin in North-Central Arkansas

R. Weih, Jr.[1] and D. White, Jr.

[1] Arkansas Agricultural Experiment Station, Arkansas Forest Resources Center (AFRC), University of Arkansas at Monticello (UAM), School of Forest Resources, Spatial Analysis Laboratory (SAL), 110 University Court, Monticello, Arkansas 71656

[1]Correspondence: weih@uamont.edu

Abstract

Sensors for remote sensing have improved enormously over the past few years and now deliver high resolution multispectral data on an operational basis. Most Land-use/Land-cover (LULC) classifications of high spatial resolution imagery, however, still rely on basic image processing concepts (i.e., image classification using single pixel-based classifiers) developed in the 1970s. This study developed the methodology using an object-based classifier to characterize the LULC for the Buffalo River sub-basin and surrounding areas with a 0.81-hectare (2-acre) minimum mapping unit (MMU). Base imagery for the 11-county classification was orthorectified color-infrared aerial photographs taken from 2000 to 2002 with a one-meter spatial resolution. The object-based classification was conducted using Feature Analyst®, Imagine®, and ArcGIS® software. Feature Analyst® employs hierarchical machine-learning techniques to extract the feature class information from the imagery using both spectral and inherent spatial relationships of objects. The methodology developed for the 7-class classification involved both automated and manual interpretation of objects. The overall accuracy of this LULC classification method, which identified more than 146,000 features, was 87.8% for the Buffalo River sub-basin and surrounding areas.

Introduction

Land-use/Land-cover is a distinct concept applied to the classification of the earth's land surface (Estes et al. 1982). Land-cover is defined as visible features on the landscape and land-use is defined as human activity on the landscape. For our classification of the Buffalo River sub-basin, we did not distinguish between land-use and land-cover because of the difficulty of identifying land-use of the landscape.

Numerous uses exist for digital LULC classification maps. For example, LULC classification maps provide insight into a region's soils and geology (Ustin et al. 1999, Gupta 2003). Land-use/Land-cover classification maps are used extensively in conservation planning (Turner et al. 2003, Kerr 2003), informing land development decisions in metropolitan areas (Ridd 1995, Weber and Puissant 2003), planning and implementing large-scale inventories of natural resources (Anderson 1982, Volgelmann et al. 1998), and monitoring change in ecosystem/landscape condition over time (Frohn 1998, Lambin 1996, Weng 2002). Land-use/Land-Cover data, particularly when used in conjunction with other data such as terrain maps available from Digital Elevation Models (DEMs), can be useful in identifying areas more or less suited to specific land management practices and thereby aid in the assessment of appropriate practices for use in a specific area to attain certain goals (Bonner et al. 1982).

Traditional methods of mapping vegetation for use in natural resource management/research and conservation planning consist of field surveying and manual mapping using aerial photography or medium to coarse resolution satellite imagery. These techniques, however, do not typically provide the level of resolution and spatial scales required by many natural resource applications. Many wildlife management and research applications, including resource selection modeling, require fine resolution data (<10 m) at large spatial scales (>10,000 ha). Until recently, such data were unavailable or impractical to obtain using field-based techniques and medium to coarse resolution satellite imagery.

Remotely sensed imagery, i.e., satellite and aerial photography, has become a cost efficient, accurate, and precise tool for developing LULC classifications (McRoberts and Tomppo 2007). This study summarizes a novel approach, using an object-based classifier instead of a pixel based classifier, to develop a highly delineated LULC classification map of the Buffalo River sub-basin in North-central Arkansas.

Materials and Methods

Our study area was located in the Ozark Plateau province (Boston Mountains; Bailey 1995) and included the entire Buffalo River sub-basin and surrounding area (Figure 1). The study area consisted of 788,474 ha and included most of Newton and Searcy counties as well as portions of Baxter, Boone, Carroll, Johnston, Madison, Marion, Pope, Stone, and Van Buren counties. Of the total area, 38,447 ha (4.9%) were managed by the National Park Service under the National Scenic Rivers Act of 1972.

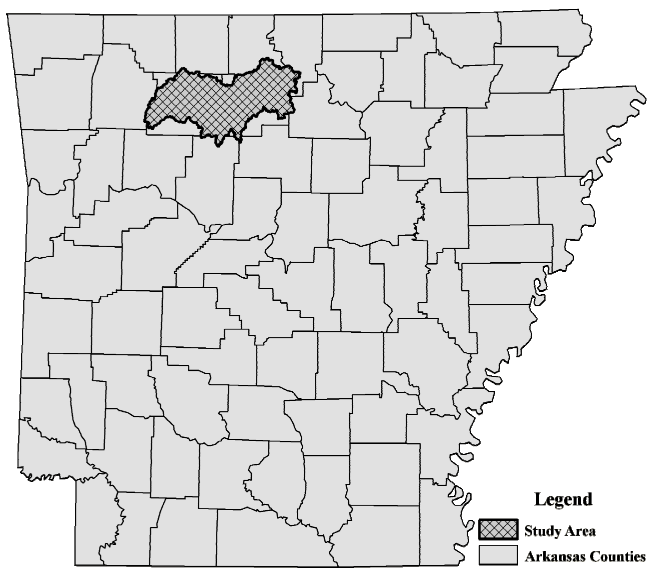

Figure 1. Location of the Buffalo River sub-basin and study area in north-central Arkansas.

The Boston Mountains are erosional remnants of a plateau that were dissected into rough terrain characterized by steep-slopes with flat ridge tops. Elevations range from 240 to 610 m. Our study area was predominately forested, consisting of oak (*Quercus spp.*), hickory (*Carya spp.),* and other hardwoods. Pine (*Pinus spp.)* and cedar (*Juniperus virginiana*) also occurred on selected sites (Bailey 1995). The area is important both ecologically and economically as it contains the states only elk herd. We developed our LULC classification as part of our research into the space use ecology of male elk (White et al. 2005).

The imagery used for the LULC classification was color Infrared (CIR) imagery. The CIR digital orthophoto quadrangle (DOQ) images used in the classification were acquired between May 2000 and January 2002, with most of the images acquired in late January and February 2001. The DOQ images had a pixel resolution of one meter. These images, acquired by the state of Arkansas, were obtained from the Natural State Digital Database (NSDD) (http://sal.uamont.edu) maintained by the Spatial Analysis Laboratory (SAL) at the University of Arkansas at Monticello (UAM).

Pixel-based image classification includes supervised and unsupervised methods (Enderle and Weih 2005). Supervised methods involve classification of pixels of unknown identity by means of a classification algorithm using spectral characteristics of pixels of known identity. Unsupervised methods involve the separation of image pixels into natural groupings based upon similar spectral characteristics by means of a classification algorithm and assignment of groupings into classes.

Marceau et al. (1990) and Hsieh et al. (2001) found that increasing spatial resolution does not necessarily increase classification accuracies because single pixels fail to capture the entire spectral signature of the object being classified. To circumvent this problem, we analyzed not only the individual pixel being classified but also neighboring pixels, resulting in the analysis of both the spectral and spatial structure of objects. Figure 2 illustrates the essential difference between pixel-based and object-based classifiers.

While the idea of using object-based classification to replace pixel-based methods has existed since the early 1970's, the first practical object-based classification model was not developed until 1984 when the Machineseg program was developed. Machineseg was an image-analysis technique that used object shapes, sizes, and spectral signatures obtained from aerial photographs (Flanders et al. 2003). Then in the late 1980's, a "road finder" program was developed that used a segmentation process to identify linear features such as roads, rivers, and field boundaries (Flanders et al. 2003). These early object-based classification models had difficulty combining information from multi-level analyses, validating classifications, reconciling conflicting results, attaining reasonable processing time, and automating analyses (Flanders et al. 2003). Pixel-based methods, which did not suffer from these problems, provided reasonably accurate classifications, and therefore maintained their position as the industry standard (Flanders et al. 2003).

While a fully automated object-based classification process was highly desired, early efforts to develop such models failed due to limitations in hardware, software, image quality (poor resolution), and interpretation theories (Flanders et al. 2003). By the mid-1990's, however, these limitations were being resolved by the development of computers with large memory capacities, fast processing speeds, and the availability of images from high spatial resolution satellite sensors with increased spectral variability (Flanders et al. 2003). Advances in image-segmentation algorithms and intelligent machine - learning algorithms have led to "off-the-shelf" software packages such as Feature Analyst® and

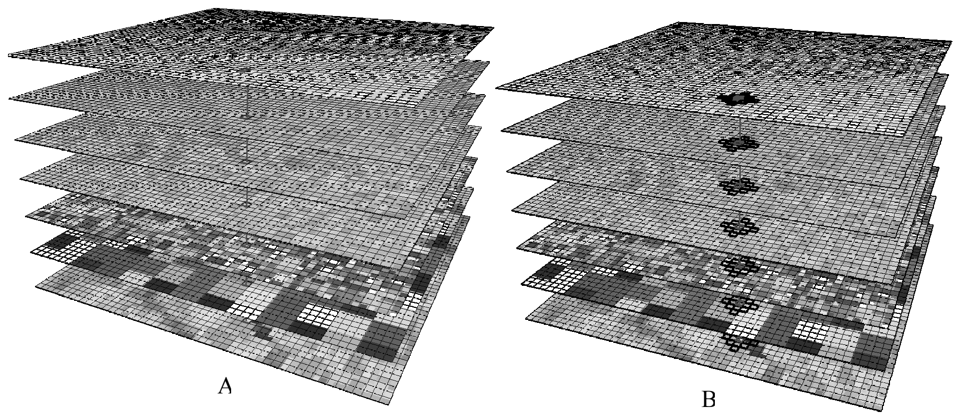

Figure 2. Pixel-based classifiers (A) classify objects within a single pixel using all layers. Object-based classifiers (B) classify objects within a defined region, including the focal pixel (the central-most pixel) using all layers.

eCognition capable of object-based classification methods that equal and often exceed the accuracy of pixel-based classification methods.

Feature Analyst®, which has been designed for use with ArcGIS®, GeoMedia®, SOCET SET®, and ERDAS Imagine® software, is a practical tool for use in LULC classification mapping (Visual Learning Systems 2004a). Feature Analyst® uses a machine-learning algorithm to achieve automated feature extraction (Visual Learning Systems 2004a). Once the software is given user-specified examples (training data sets), it utilizes software agent technology to “learn” to find similar landscape features and appoint a user-defined classification (Visual Learning Systems 2004a). If a series of images of the same area over time are correctly registered to each other, Feature Analyst® can extract changes that may have occurred in the features of the image by creating a change detection raster (Visual Learning Systems 2004b).

O’Brien (2003) at the National Imagery and Mapping Agency (NIMA) compiled a report on a series of tests that compared Feature Analyst® with manual methods currently employed for mapping operations. Feature Analyst® increased production over hand digitization, while at the same time achieved more accurate and consistent results (O’Brien 2003). Results from a questionnaire and discussions with participants of the test indicated a high level of enthusiasm for the Feature Analyst® system. Analysts agreed that the system was easy to learn and easy to use (O’Brien 2003).

The object-based LULC classification workflow used in our study involved 8 steps (Figure 3). The first step was to develop a training data set for the 7 LULC classifications of interest. The study area was divided into 17 approximately equal-area tiles to organize and facilitate the processing of such a large image data set. We developed more than 25 training polygons for each tile for each of our 7 LULC classes (hardwoods, agriculture, conifers, roads, rivers, water (other than river), and urban.

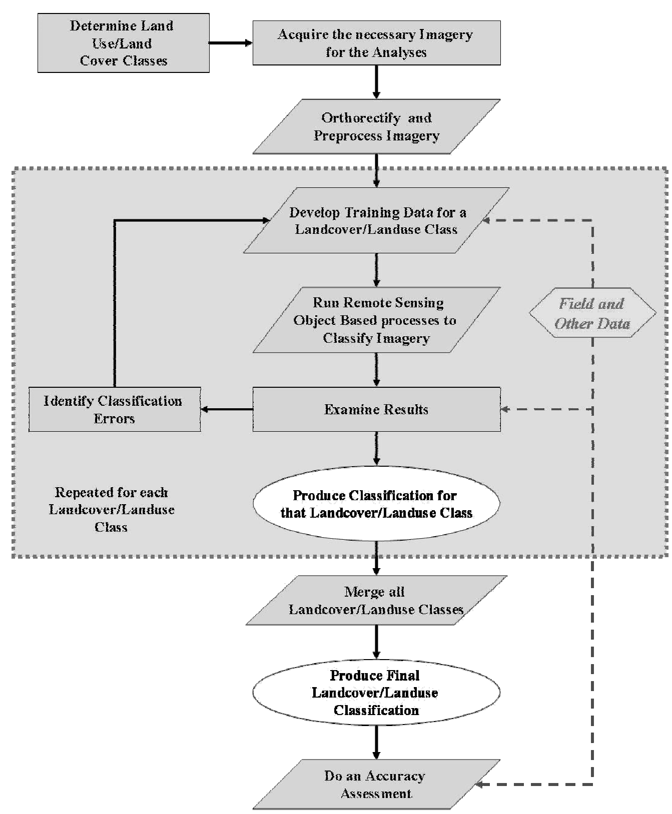

Figure 3. Workflow for developing the Land-use/Land-cover Map Layer using an object based classifier.

Step 2 was to determine the spatial context of neighbors for each LULC class being extracted.

Figure 4 shows an example of the geometric pattern of pixels used to define neighboring pixels that were used to classify a focal (or central) pixel. The geometric pattern of pixels was different for each of the classes. The characteristics of a LULC class can be better represented by an organized group of pixels (spatial feature representation) than single pixels as used in traditional pixel based classifiers for high spatial resolution images. In this step we ran the object-based classifier (Feature Analyst®) and visually examined the results.

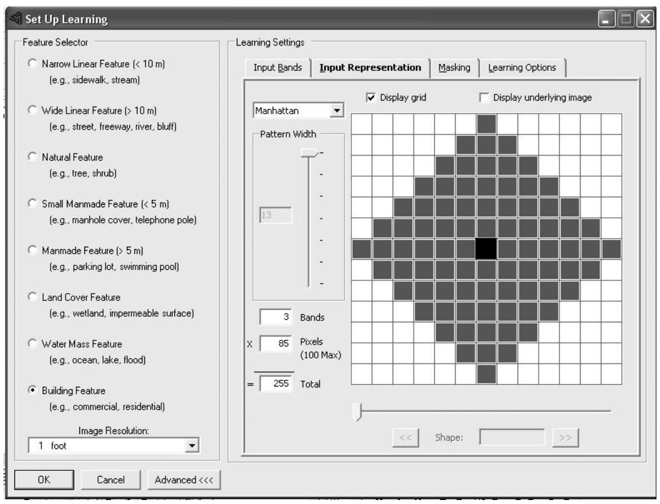

Figure 4. Spatial pattern of an object-based classifier used to classify a single pixel (black square).

Third, we examined the results and identified the features correctly and incorrectly classified. This is a function of Feature Analyst® to assist it in the learning process to classify a feature class. The MMU for the classification was 0.81 ha (2-acres) except for water (other than river), which was 0.04 ha (0.1 acre). The MMU determines the minimum size any feature must be to be considered a separate feature.

Fourth, we repeated steps 1 and 2 with correct and incorrect classified features. In the fifth step, the process was started over for the next LULC class (Figure 3).

For sixth step of the process, after all classes were extracted from the images on each tile, they were merged based on a model that prioritizes the LULC classes. This was done for each of the 17 image tiles in ERDAS Imagine. In the seventh step, we merged all tiles and produced a single LULC classification map for the study area (Figure 5).

In the eighth and last step, we conducted an accuracy assessment of our map by randomly selecting 795 reference data points in the study area (Congalton and Green 1999). The selection of a proper and efficient sample design to collect valid reference data is an important component of any accuracy assessment because the design will determine both the cost and the statistical rigor of the assessment (Congalton and Green 1999). Congalton and Green (1999) list five common sampling schemes for acquiring reference data: simple random sampling, systemic sampling, stratified random sampling, cluster sampling, and stratified systemic unaligned sampling. They recommend stratified random sampling, where a minimum number of samples are selected from each stratum (i.e., map category) (Congalton and Green 1999). This study used this sampling technique in an attempt to collect representative samples from all the LULC classes in the study area. Each reference point was then identified as one of our 7 land-cover classes by an individual not associated with the construction of the classification.

What constitutes an acceptable level of classification accuracy is debatable. Foody (2002) recommended an 85% target for user's, producer's, and overall accuracies derived from the error matrix. While this level may exist as a *de facto* standard, accuracy assessments of Geographic Information System (GIS)-produced maps often fail to meet this criterion (Anderson et al. 1976, Foody 2002). This is probably due to the fact that for each component of accuracy there is a set of accuracy measures that may be calculated to express it (Foody 2002). In reality, it is probably impossible to specify a single, all-purpose measure of classification accuracy, because it depends on the application and the level of comfort the practitioner has with the classification.

When evaluating an image classification, there are two forms of accuracy that can be considered. Non-site-specific accuracy (NSSA) considers the overall agreement between the classified image and the reference data without examination of the agreement between them at specific locations. For example, NSSA involves the examination of the percent Mature Pine Forest in the classified image and the comparison of it to the percent Mature Pine Forest in the reference data. Relying solely on non-site-specific accuracy to evaluate a classification can hide errors resulting from disagreement in the placement of classes between the classified image and the reference data.

The second form of accuracy is site-specific accuracy (SSA), which examines the agreement between classes at specific locations on the classified image and in the reference data. This examination is done by means of an error matrix (also known as a confusion matrix or contingency table) to compare, for specific locations, what LULC class is the reference

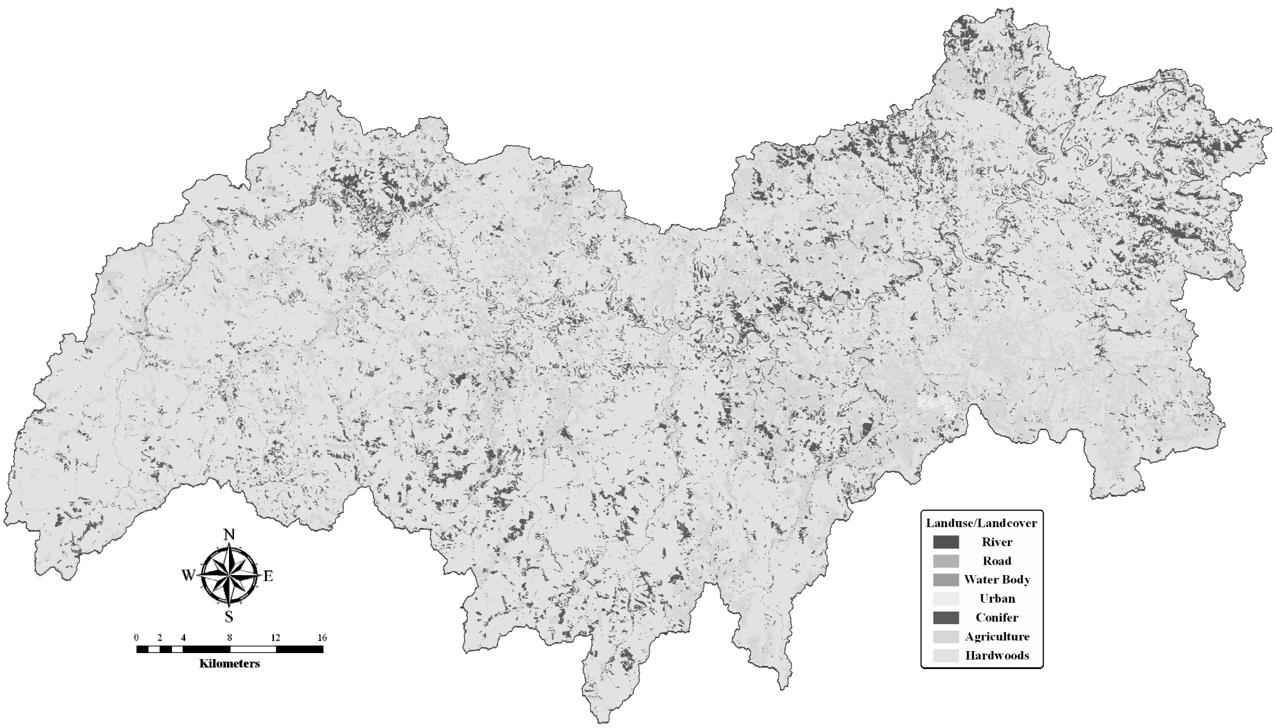

Figure 5. Buffalo River sub-basin Land-use/Land-cover classification.

data versus how that area was classified. The error matrix helps to identify instances of classification error for specific classes

There are two types of classification errors: errors of omission and errors of commission. Errors of omission are instances in which site has been excluded from a class to which it actually belongs. Errors of commission are instances in which a site is included in an incorrect class. Campbell (2007) noted that these errors tend to balance each other, as an error of omission for one class will also be tabulated in the error matrix as an error of commission in another class. Given the characteristics of these errors, it is best to examine them on a class-by-class basis before assuming the errors in one class reflect the errors found in all classes.

For SSA assessment using the error matrix, there are three primary measures of classification accuracy: overall classification accuracy, producer's accuracy, and user's accuracy. Overall classification accuracy is a measure of how much area was correctly classified for the entire area classified. From the error matrix, overall classification accuracy is the sum of the diagonals divided by the total.

Producer's accuracy is calculated for each class and provides an indication of how well a particular class has been classified by the producer of that classification. This accuracy is most often used by the producer as a means to assess how well the classifier performed. From the error matrix, the producer's accuracy for each class is the result of dividing the correctly classified pixels by the number of reference data pixels in that class.

User's accuracy is also calculated for each class and provides an indication of how often the areas assigned to a given class on the image classification actually belong to that class on the landscape. This accuracy is of greater importance to the users of the classification because this indicates how true the classified image is to the actual situation on the ground. From the error matrix, the user's accuracy for each class is the result of dividing correctly classified pixels in a given class by the total number of pixels in that class for the classified image. We report all three primary measures of classification accuracy in this study.

The area of each class in the study area was clipped using the watershed boundary of the Buffalo River in a GIS. The Buffalo River sub-basin boundary was calculated using a 5-m DEM in a GIS.

Results and Discussion

Number of features classified, area, and percentage of study area for each classification in the study area is summarized in Table 1. Almost 82% of the study area (282,967.10 ha) was forested. The two most common land-cover types in the study area were hardwoods (73.43%) and agriculture (16.49%). More than 6,000

Table 1. Number of features classified, area, and percentage of total study area by Land-use/Land-cover type.

Land Cover	Number of Features	Acres	Hectares	Percentage
Agriculture	21,243	141,068.0	57,072.9	16.49%
Roads	430	9,063.5	3,667.7	1.06%
Conifer	28,863	71,183.5	28,770.8	8.31%
Hardwoods	88,748	628,615.8	254,196.3	73.43%
Rivers	724	4,247.2	1,717.9	0.50%
Urban	149	873.7	353.2	0.10%
Water (non-river)	6,463	1,129.4	404.3	0.12%
Total	146,620	856,181.1	346,183.1	

water structures (mostly ponds) were identified and 1.06% of the study area was covered by roads.

Although classification accuracy varied by LULC type, the overall accuracy of our map was 87.8% (Table 2), which is the percentage of correct ground reference points for the LULC map. Producer and user accuracies varied from 46.6% to 100% and 79% to 100%, respectively (Table 2). Water (other than river) was accurately classified 100% of the time by both producers and users and linear features, such as rivers and roads, were correctly classified >97% of the time by both producers and users.

Table 2. Producer's and user's accuracy by Land-use/Land-cover type.

Land Cover	Producer's Accuracy	User's Accuracy
Rivers	97.2%	100.0%
Roads	97.0%	98.5%
Water (non-river)	100.0%	100.0%
Urban	48.5%	100.0%
Conifer	46.6%	80.4%
Agriculture	88.3%	79.0%
Hardwood	95.6%	86.6%
Overall Accuracy	87.8%	

As previously stated, producer's accuracy relates to the probability that a reference sample point will be correctly mapped and measures the errors of omission and producer's accuracy indicates the probability that a sample from LULC map actually matches the reference sample data and measures the error of commission. Users of the LULC map are interested in user's accuracy.

Producers misclassified conifers and urban features most frequently, whereas users misclassified agriculture most frequently. Conifers and urban features were correctly classified <46% of the time by producers but correctly classified >80% of the time by users. Although only 2.3% of 432 hardwood reference points were misclassified as conifer, 52.3% of 88 conifer reference points were misclassified as hardwood. This led to the low classification accuracy for conifers by producers (Table 2). It was easier to accurately identify hardwoods than it was conifers in our study area. This is probably explained by the criteria we set to designate an area as conifer (i.e., an area must be >50% conifer to be designated conifer). In our study area, conifers do not typically occur in large, dense stands but occur at relatively low basal areas mixed with hardwoods. Visually estimating percent coverage of a sparely distributed land-cover type is difficult and error prone.

The classifier classified 11.7% of 94 agriculture reference points as hardwood, which lowered the accuracy of this LULC class (Table 2). Typically, an agricultural field in the study area contained hardwoods, which complicated classification efforts for the same reason conifers were difficult to distinguish from hardwoods.

All urban areas in the classified image were classified correctly, 100% User's Accuracy (Table 2), but some urban area field data points were classified as agriculture and hardwood. This is reflected in the 48.5% producer's accuracy.

In summary, Feature Analyst® processes are similar to the way human interpreters identify objects, which involves: association, color, pattern, shadow, shape, size, and texture (Caylor 1998). A pixel based classifier might only look at color (spectral signature) and possibility texture or pattern in an advanced classification process workflow. The methodology developed for this study showed that an object-based classifier can produce accurate LULC classifications with high spatial resolutions.

Acknowledgments

This research would not have been possible without the efforts of James Hartshorn, Alex Harper, David General, Jake Freeman, Jonah Freeman, Dale Riggan, and Chris Kleinhofs. The Arkansas Game and Fish Commission and the Forest Resources Center at the University of Arkansas at Monticello supported this research.

Literature Cited

Anderson R, E Hardy, J Roach, and **R Witmer**. 1976. A land Use and Land Cover Classification for Use with Remote Sensing. USGS Professional Paper 964. Washington, D.C. 28 p.

Anderson R. 1982. Land resources map making from remote sensing products. Pp. 63-72, In Remote Sensing for resource management (C.J. Johannsen and J.L. Sanders, eds.). Soil Conservation Society of America, Ankeny, IA. 665 p.

Bailey R. 1995. Description of the Ecoregions of the United States. 2nd ed. USDA Forest Service, Publication No. 1391. Washington, DC. 108 p.

Bonner J, **W Rohde**, and **W Miller**. 1982. Mapping wildland resources with digital Landsat and terrain data. Pp. 73-80, In Remote Sensing for resource management (C. Johannsen and J. Sanders, eds.). Soil Conservation Society of America, Ankeny, IA. 665 p.

Campbell J. 2007. Introduction to Remote Sensing, 4th ed. Guilford Press, New York. 625 p.

Caylor J. 1988. How to use Aerial Photography for Natural Resource Applications. USDA Forest Service, National Forestry Application Program. 328p.

Congalton R and **K Green**. 1999. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. Boca Raton, Florida: Lewis Publishers. 137 p.

Enderle D and **R Weih**. 2005. Integrating supervised and unsupervised classification methods to develop a more accurate land cover classification. Journal of the Arkansas Academy of Science 59:65-73.

Estes J, **D Stow**, and **J Jensen**. 1982. Monitoring land use and land cover changes. Pp. 100-110, In Remote Sensing for resource management (C. Johannsen and J. Sanders, eds.). Soil Conservation Society of America, Ankeny, IA. 665 pp.

Flanders D, **M Hall-Beye**, and **J Pereverzoff.** 2003. Preliminary evaluation of eCognition object-based software for cut block delineation and feature extraction. Canadian Journal of Remote Sensing 29(4):441-452.

Foody G. 2002. Status of land cover classification accuracy assessment. Remote Sensing of Environment 80:185-201.

Frohn R. 1998. Remote Sensing for Landscape Ecology: New Metric Indicators for Monitoring, Modeling, and Assessment of Ecosystems. CRC Press LLC. 99 p.

Gupta R. 2003. Remote Sensing in Geology. Springer-Verlang Berlin Heidelberg New York. 656 p.

Hsieh F, **C Lee**, and **N Chen.** 2001. Effect of spatial resolution on classification errors of pure and mixed pixels. IEEE Transactions on Geoscience and Remote Sensing. 39(12):2657-2663.

Kerr J and **M Ostrovsky**. 2003. From space to species: ecological applications for remote sensing. Trends in Ecology and Evolution. Volume18 6:299-305.

Lambin F. 1996. Change detection at multiple temporal scales: Seasonal and annual variations in landscape variables. Photogrammetric engineering and remote sensing. 62:931-938.

Marceau DJ, **PJ Howarth**, **JMM Dubois**, and **DJ Gratton**. 1990. Evaluation of the gray-level co-occurrence matrix-method for land cover classification using SPOT Imagery. IEEE Transactions on Geoscience and Remote Sensing. 28(4):513-519.

McRoberts E and **E Tomppo**. 2007. Remote sensing support for national forest inventories. Remote Sensing of Environment. 110(4):412-419.

O'Brien M. 2003. Feature extraction with the VLS Feature Analyst System. Report to the ASPRS 2003 Annual Conference Proceedings, Anchorage, AK. http://www.featureanalyst.com/featureanalyst/publications.htm Accessed 5 April 2008.

Ridd M. 1995. Exploring a V-I-S (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities. International Journal of Remote Sensing. 16:2165-2185.

Turner W, **Spector**, **N Gardiner**, **M Fladeland**, **E Sterling**, and **M Steininger**. 2003. Remote sensing for biodiversity science and conservation. Trends in Ecology and Evolution. 18:306-314.

Ustin L, **M Smith**, **S Jacquemoud**, **M Verstraete**, and **Y Govaerts**. 1999. Geobotany: vegetation mapping for earth sciences. Pp. 189-248, In Remote Sensing for the Earth Sciences: Manual of Remote Sensing, 3rd ed., Vol. 3 (A. Rencz, ed.). John Wiley & Sons, New York. 707 p.

Visual Learning Systems, **Inc**. 2004a. Using Feature Analyst TM for land cover and land use mapping. A Visual Learning Systems, Inc. White Paper, Missoula (MT). http://www.featureanalyst.com/featureanalyst/publications.htm Accessed 5 April 2008.

Visual Learning Systems, **Inc**. 2004b. Using Feature Analyst to perform object-specific change detection. A Visual Learning Systems, Inc. White Paper, Missoula (MT). http://www.featureanalyst.com/featureanalyst/publications.htm. Accessed 5 April 2008.

Vogelmann J, **T Sohlt**, and **S Howard**. 1998. Regional characterization of land cover using multiple sources of data. Photogrammetric engineering and remote sensing. 64:45-57.

.

Weber C and **A Puissant**. 2003. Urbanization pressure and modeling of urban growth: example of the Tunis Metropolitan Area. Remote Sensing of Environment. 86(3):15-352.

Weng Q. 2002. Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS and stochastic modeling. Journal of Environmental Management. 64:273-284.

White D, M Cartwright, and **R Weih** 2005. Answers about Elk: Helicopters, Radio Waves Deliver Data about Elk. Arkansas Wildlife. pp. 6-12