

FUZZY CLASSIFICATION OF HETEROGENEOUS VEGETATION IN A COMPLEX ARID ECOSYSTEM

Ben McMahan, Keith T. Weber, and Joel Sauder
GIS Training and Research Center
Idaho State University
Pocatello, ID 83209-8130
(<http://giscenter.isu.edu>, e-mail: giscenter@isu.edu)

ABSTRACT

Traditional methods of remote sensing classification may not accurately portray the complexity of ecosystems where vegetation type and structure is diverse and variable. We used fuzzy classification to better characterize the complexity and heterogeneity of sagebrush-steppe vegetation on the Upper Snake River Plain in southeast Idaho. Unlike supervised classification where pixels are classified into discrete categories, fuzzy systems classify each pixel into multiple categories based on estimated membership in each class. Field data (n = 370) collected in summer 2002 were used as training sites for supervised classification and fuzzy classification. We compared the results of supervised and fuzzy classification to determine which produced a more accurate depiction of land cover. Our results show fuzzy classification produces more accurate predictions of sagebrush and grass cover compared to supervised classification. Additionally, there is an apparent relationship between fuzzy membership and percent cover of sagebrush, which provides the user vegetation structure information as well.

Keywords: Remote sensing, sagebrush, grasslands.

INTRODUCTION

Vegetation modeling can be problematic in ecosystems where vegetation types are heterogeneous or are poorly represented by large pixels. Fuzzy classification offers an alternative to crisp logic (supervised classification) by evaluating pixels based on their membership into each category. Fuzzy membership is based on fuzzy set theory, which assumes that membership to a given category will range from complete membership (100%) to non-membership (0%), and that pixels may be classified as partial members into two or more categories (Gopal and Woodcock, 1994).

The Upper Snake River Plain (SRP) in southeastern Idaho is a sagebrush-steppe ecosystem, where sagebrush (*Artemisia* spp.), forb, and grass species comprise a majority of the vegetation. Bare soil, lava flows of various ages, and cryptobiotic crust are also present. The complexity of SRP vegetation types has been poorly classified by previous landscape-scale mapping attempts (e.g. GAP and NLCC classifications, 2000) with accuracies of <50% in sagebrush and grassland areas. These classifications also perform poorly in their representation of percent cover of vegetation, which is an important characteristic of the ecosystem (Connely, et al, 2000).

Fuzzy logic can be used to classify satellite imagery based by calculating membership to homogeneous sagebrush and grassland training sites. If a correlation between membership and percent cover can be established, fuzzy classification will function as a better technique for producing land cover maps which include information on vegetation structure. Fuzzy classification has also been shown to improve classification accuracy of land cover types when compared to maximum likelihood supervised classification (Bardossy and Samaniego, 2002).

METHODS

Field data ($n = 370$) were used to create training sites for maximum likelihood supervised classification and fuzzy classification of Landsat 7 ETM+ imagery of the SRP (acquired August 07, 2002). Seven base images and indices (PVI, NDVI, SAVI, tasseled cap greenness, tasseled cap moistness, Landsat 7 band 5, and PCA composite 2) were created with this imagery and each was used as a training image in the models. These were chosen because they best characterized the spectral variability within the training sites. We created two models for each classification system (supervised and fuzzy): one that used two homogenous classes of training sites (sagebrush and grass), and one that used three classes of training sites (sagebrush, grassland and mixed). Sagebrush sites had >15% cover of sage and <15% cover of grass. Grass sites had <15% sage cover and >15% grass cover. Mixes sites had >15% cover of both sage and grass. Spectral signatures were extracted using these training sites from the base imagery and classify the imagery using both supervised and fuzzy classification. Supervised classification produces an image where each pixel is classified as having 100% membership in a category (e.g. sagebrush or grassland). Fuzzy classification produces a series of images where the membership of each pixel is determined based on spectral similarity to training sites and user confidence of homogeneity of training sites (z-score).

We validated the supervised classifications using traditional error matrices. Validation of fuzzy classification is problematic because fuzzy results must be converted into crisp categories (Congalton and Green, 1999). We validated fuzzy classification using two methods: 1) flexible error matrix validation described in Congalton and Green (1999), and 2) a simple presence-absence threshold validation. To test for a significant relationship between fuzzy membership and percent cover of sagebrush and grass an ANOVA (Analysis Of Variance) was used.

RESULTS

The two-class and three-class supervised classifications had an accuracy of 33.7% and 47.5%, respectively. The results of fuzzy classification using a z-score of 1.5 generated the most accurate results by minimizing type 1 and type 2 errors. The flexible error matrix validation showed the two-class fuzzy model had an accuracy of 39.5%, while the three-class model had an accuracy of 50.5%. The presence/absence threshold procedure showed the two-class fuzzy model had an accuracy of 65.6% predicting sagebrush cover and of 44.8% predicting grass cover. The three-class fuzzy model had an accuracy of 65.9% predicting sagebrush cover and 48.0% predicting grass cover. The ANOVA demonstrated a significant relationship between sagebrush membership and percent sagebrush cover ($F=8.114$, $P < 0.0001$). No significant relationship was detected between grass membership and percent grass cover.

DISCUSSION

Both classifications appear to perform at least as well as currently available vegetation maps (e.g. GAP or NLCC). The supervised classification did not have as much flexibility in classifying diversity in the region, and subsequently accuracy suffered. The improved accuracy of fuzzy classification demonstrates the utility of fuzzy logic systems for classification of heterogeneous vegetation in complex ecosystems. Additionally, since fuzzy classification is able to utilize remotely sensed data with coarse spatial resolution and still produce accurate results, it is more practical for landscape-scale analyses compared to other methods (e.g. hyperspectral or fine spatial resolution multispectral data).

A limitation of fuzzy classification using coarse-scale data is the increased heterogeneity of training classes. Higher spatial resolution multispectral data would allow for smaller, more homogeneous training sites that should yield improved results. The necessity to convert fuzzy results into crisp data for validation illustrates another limitation of fuzzy classifications. Validating relationships between membership and vegetation structure (such as those demonstrated in this study) are less susceptible to this limitation, as the process does not require prior conversion to crisp data. A final limitation of our application is the nature of training classes used in this study. Sagebrush and grass vegetation types present difficulties regarding how a sensor images these targets. Sagebrush has open spaces between branches as well as dead and dying limbs which do not reflect spectra the way “pure” sagebrush pixels would. Further, the elongated vertical growth-form of grasses means most of the plants biomass is perpendicular to the viewing angle of a satellite. Increased spatial resolution will help, but these issues are characteristic of the general problems associated with accurate representation of complex vegetation types and structures.

In conclusion, fuzzy classification provides a means by which vegetation heterogeneity and variability can be modeled if a relationship between fuzzy membership and percent cover can be reliably established. Even if such relationships cannot be established, fuzzy classification produces more accurate vegetation cover "maps" compared to other methods. Improved classification accuracy and the potential to model vegetation structure and density will prove useful to land managers. This is especially true in areas where existing classifications do not adequately portray the complexity of vegetation found in the region. Further research will attempt to establish a stronger correlation between fuzzy membership and vegetation structure using imagery with improved spatial resolution.

ACKNOWLEDGEMENTS

This study was made possible by a grant from the National Aeronautics and Space Administration Goddard Space Flight Center. ISU would also like to acknowledge the Idaho Delegation for their assistance in obtaining this grant.

LITERATURE CITED

Bardossy, A., and L. Samaniego. 2002. Fuzzy Rule-Based Classification of Remotely Sensed Imagery. *IEEE Transactions on Geoscience and Remote Sensing* 40 i2:p362(13).

Congalton, R. G., and K. Green. 1999. *Assessing the Accuracy of Remotely Sensed Data*. CRC Press, Boca Raton, FL. 137 pp.

Connelly, J.W., et al. 2000. Guidelines to Manage Sage Grouse Populations and Their Habitats. *Wildlife Society Bulletin* 28(4):967-985.

Gopal, S., and C. Woodcock, 1994. Theory and Methods for Accuracy Assessment of Thematic Maps Using Fuzzy Sets. *Photogrammetric Engineering and Remote Sensing* 60:181-188.

National GAP Analysis. 2000. GAP-Homepage. www.gap.uidaho.edu 1 Oct 2002.

National Land Cover Characterization. 2000. NLCC-Homepage. <http://landcover.usgs.gov/nationallandcover.html> 1Oct 2002.