Comparing Classification Results of Multi-Seasonal TM against AVIRIS Imagery – Seasonality more Important than Number of Bands

SYLVIO MANNEL, Nevada, MO & MARIBETH PRICE, Rapid City, SD, USA

Keywords: AVIRIS, Landsat TM, multitemporal, forest, classification

Summary: We classified forest cover and tree density in the Black Hills, SD, in twenty spatially contiguous AVIRIS scenes. Results were compared to those derived from two-season Landsat TM imagery. A decision tree classifier was used to analyze the TM data as well as the over two hundred bands of the twenty AVIRIS scenes. The classification based on summer AVIRIS data was more accurate than the classification based on the comparable early fall TM data. However, classification of spring and especially, two-season TM data resulted in higher accuracies than the classification based on summer hyperspectral data. These results indicate that seasonality is more important than the number of spectral bands.

1 Introduction

Hyperspectral data offer the opportunity to explore the differences of land cover types without being restricted to a few wavelengths. AVIRIS (airborne visible/infrared imaging spectrometer) was the first hyperspectral sensor that measured over 200 bands between the 400 nm and 2500 nm spectrum with individual band widths of ~10 nm. AVIRIS employs four spectrometers in the following ranges: 400–710 nm, 670–1290 nm, 1250–1870 nm and 1830–2450 nm (GREEN et al. 1990). The flight altitude is about 20 km with a rate of 7300 spectra per second (GREEN et al. 1990). The covered area of a high altitude flight ranges from 11 km × 9 km (WILLIAMS & HUNT Jr. 2002) to 12.3 km × 10.2 km (RIANO et al. 2002). The resulting area of about 120 km² is much smaller than for example a Landsat TM scene, which usually covers about 20,000 km².

Vegetation studies can take advantage of the continuously available reflectance; for example, NIEMANN et al. (2002) note that narrow spectral bands are necessary to detect some forest related parameters whose spectral range may be small. Hyperspectral data have been used to map land cover types such as woody vegetation (WYLIE et al. 2000, USTIN & XIAO 2001), leafy spurge (WILLIAMS & HUNT Jr. 2002), shrub recovery after fire (RIANO et al. 2002), vegetation in semi arid ecosystems (ASNER & HEIDEBRECHT 2002, OKIN et al. 2001) or lake water quality (HOOGENBOOM et al. 1998, THIEMANN & KAUFMANN 2002).

The nearly continuous spectrum also has its costs, in an economical, computational, and spatial sense. A typical AVIRIS scene holds...
about 400 MB. Most classifications, for example, maximum likelihood, require all flight lines fused into one image, severely restricting their utility for hyperspectral classification of large areas with more than 10 AVIRIS scenes. However, a decision tree classification, allows the user to work with each AVIRIS scene separately. To our knowledge this might be one of the first studies to classify a medium scale forest of about 2,000 km², utilizing twenty AVIRIS contiguous scenes. In preliminary tests decision trees also showed to be comparable or even slightly better than maximum likelihood classifications of Landsat TM data (Mannel et al. 2002).

Decision trees use a binary recursive partitioning algorithm to divide the data into smaller subsets with increasing statistical homogeneity (Swain & Hauska 1977, Breiman et al. 1984, Clark & Pregibon 1993). These divisions can be represented as branches and nodes, where nodes are connected to a set of possible answers that will lead to a classification. This process is often referred to as data mining (Read 2000).

Projects that compared AVIRIS data to other data sources are not always in agreement. Ustin & Xiao (2001) found AVIRIS about 20% more accurate than SPOT data in classifying forest regions. Lee & Cohen (2002) had better success mapping leaf area index (LAI) with AVIRIS than with Landsat ETM+ data. Lefsky et al. (2001), on the other hand, found that multi-seasonal TM data performed better than AVIRIS for quantifying forest biomass and basal area, although their study used 1994 AVIRIS data in which the signal-to-noise ratio was lower than today. They used six temporally different TM scenes but also showed improved results with only two TM scenes. Different seasons have been found to be important for vegetation classifications (Schriever & Congalton 1995, Wolter et al. 1995, Lefsky et al. 2001).

In this study we compare twenty scenes of single-date AVIRIS data and two-season TM imagery to classify a forest in the Black Hills, South Dakota, USA. These two sensors were chosen to allow us to compare expensive hyperspectral data to a low-cost multispectral alternative. We received AVIRIS data via a grant by NASA and then selected freely available Landsat TM imagery over the same area. A decision trees classifier was used because of slightly better results in preliminary tests (Mannel et al. 2002) and because it offers the best alternative for processing large areas composed of many AVIRIS scenes (by being able to work with separate scenes without the need to fuse them).

Fig. 1: Study area.
2 Methods

2.1 Study Area

The study site lies in the northern Black Hills, SD. The Black Hills form an oval uplift approximately 200 km × 100 km in size with elevations reaching 2200 m in contrast to the surrounding plains at an elevation of about 900 m. The forested hills rise above the sea of prairie grass and form an orogenic and ecological island in the plains. Ponderosa pine (Pinus ponderosa) covers approximately 84% of the Black Hills (Bennett 1984), supplemented by other trees such as white spruce (Picea glauca), aspen (Populus tremuloides), paper birch (Betula papyrifera) and oak (Quercus macrocarpa). Ponderosa pine and white spruce each form forest stands typically dominated by one species, and they form two of the primary mapping classes in this study. Birch is hard to distinguish from aspen even from a few yards away. The spectral similarities between aspen and birch were confirmed by Singhroy et al. (2000) who measured both species with a field spectrometer. In addition the two trees usually occur as a species association in medium to dense stands. For these reasons, birch and aspen were mapped as a single class. Bur oak stands are isolated and too small to justify a separate land cover based on remote sensing. Open meadows are common in the Black Hills, and are dominated by mixed shortgrass prairie grasses such as western wheatgrass (Pascopyrum smithii), blue grama (Bouteloua gracilis), little bluestem (Schizachyrium scoparium), and others, with a variety of forbs (Larson & Johnson 1999). Non-vegetated classes include water and bare rock or soil.

2.2 Field Data and Cover Classes

Field data plots were sampled in the summer of 2000. Points were randomly distributed and stratified by cover class and forest density based on the Rocky Mountain Resource Information System (RIS) compiled by the Black Hills National Forest. Density refers to the percentage of ground covered by tree crowns when viewed from directly overhead. We established and measured 135 15-m radius plots. Overstory canopy cover was measured with a sighting tube (Ganey & Block 1994, Cook et al. 1995, Mannel et al. 2006). The sighting tube made it possible to record species while taking canopy measurements. Based on the forest service protocols, we distinguished the following densities: open, medium (40%–60% canopy cover), and dense (RIS, unpublished data 2000). Cover type was designated based on the dominant species that covered more than 70% of the plot. If no species covered more than 70%, the plot was labeled “mixed”.

Tree species differ in their abundance and densities. Ponderosa pine is abundant in the Black Hills but rarely forms very dense stands. We only encountered two field plots with a canopy cover of more than 70%. The spectral reflectance characteristics are therefore close to the medium class (40%–60%). We, therefore, combined medium and dense pine to one class. Aspen and white spruce are much less abundant than pine. Both tree classes usually grow in medium or dense stands in the Black Hills. We combined all densities for these two species into one aspen and one white spruce class.

Trees in the Black Hills typically form distinct stands dominated by one species, and areas with mixtures of different species (other than aspen/birch) were rare and of small size. As a result, an insufficient number of reference data for mixed classes could be collected. Water was identified using Digital Orthophoto Quadrangles (DOQs). DOQs are essentially georeferenced aerial photographs. Our DOQs were black and white with a ground resolution of one metre.

Training data for the nonvegetated class was based on an open sand pit, which was also the calibration site for AVIRIS atmospheric correction. Additional reference data for bare areas were difficult to collect, since bare areas large enough to buffer possible georeferencing errors and mixed pixels are rare in the Black Hills. Our final land cover scheme included water, nonvegetated bare areas, meadow, aspen (including birch), spruce, open pine, and medium/dense pine.

The field-sampled plots were mapped on AVIRIS, TM, and DOQ imagery and were visually inspected for potential mislocations.
prior to including them in the reference data set. Checking the spatial validity of remotely sensed data with respect to reference data location is necessary in boundary regions between classes. Removing reference data from boundary regions, such as roads or along meadow/forest boundaries, can increase the accuracy significantly (Foody 2002, Mannel et al. 2011).

In order to provide sufficient training and test data we used DOQs to collect additional points spatially close to field data (Mannel et al. 2006). Around all field-sampled plots we visually identified homogeneous areas using 1-m resolution DOQs, making the assumption that the vegetation characteristics within those regions were similar to those measured at the field site (Mannel et al. 2006, Mannel et al. 2011). Tab. 1 shows the final reference data distribution based on field data and data clusters sampled with DOQs. There are fewer clusters than field data for two reasons. First, some field data sites were spatially invalid and were omitted from consideration. Second, some of the original randomly sampled sites were close to each other, and were considered to belong to a single cluster.

It is obvious that the DOQ-collected reference data would be spatially autocorrelated within each cluster. Moran’s index of all the data including the clustered additional points was 0.67 (Mannel et al. 2011). However, autocorrelation between clusters was not present because the field-sampled plots at the core of the clusters were randomly stratified, and spatially close sites were aggregated to a single cluster. When assigning data points for either training or testing during analysis, the reference data were always divided by clusters such that a single cluster was assigned as either all training data or all testing data, but never both (Mannel et al. 2011). This technique was applied to negate the effects of spatial autocorrelation on the accuracy assessment.

2.3 Image Preprocessing

The high-altitude AVIRIS flight took place in the summer of 2000. The first flight on 26 June 2000 was repeated due to 5%–50% cloud cover within the flight rows. The second flight on 6 July 2000 covered six flight lines resulting in 30 scenes. The weather was clear with the exception of light haze and one cloud. We selected twenty scenes that covered an area containing most of our field data sites.

During the AVIRIS overflight we measured the reflectance of our calibration site, a bare sand pit, with a handheld spectrometer (ASD VNIR Dual, with a spectral range of ~370 nm–1100 nm). The atmospheric correction program ACORN (Analytical Imaging and Geophysics, LLC, 2001) was used in conjunction with the calibration site measurements to correct for atmospheric effects and to convert radiance to reflectance. The following AVIRIS bands were removed: 374 nm–394 nm (bands 1–3), 1354 nm–1424 nm (bands 107–114), 1812 nm–1951 nm (bands 153–168), and 2489 nm–2509 nm (bands 222–224) because of strong atmospheric effects. In these wavelengths atmospheric gases, such as water vapor increase the signal noise to unaccepta-

Tab. 1: Number of sites for each land cover class. The field measured plots were supplemented with additional points based on DOQs resulting in large point clusters.

<table>
<thead>
<tr>
<th>Land cover</th>
<th>Number of field data plots</th>
<th>Number of field- and DOQ-based reference data</th>
<th>Number of clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meadow</td>
<td>19</td>
<td>64</td>
<td>10</td>
</tr>
<tr>
<td>Bare</td>
<td>0</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Open pine</td>
<td>26</td>
<td>104</td>
<td>14</td>
</tr>
<tr>
<td>Dense pine</td>
<td>26</td>
<td>212</td>
<td>19</td>
</tr>
<tr>
<td>Spruce</td>
<td>21</td>
<td>126</td>
<td>12</td>
</tr>
<tr>
<td>Aspen</td>
<td>17</td>
<td>84</td>
<td>9</td>
</tr>
<tr>
<td>Water</td>
<td>0</td>
<td>106</td>
<td>2</td>
</tr>
<tr>
<td>Mixed</td>
<td>25</td>
<td>Not utilized</td>
<td>–</td>
</tr>
</tbody>
</table>
Fig. 2: AVIRIS spectra and standard deviation for land cover classes. The following AVIRIS bands were removed: 374 nm–394 nm (bands 1–3), 1354 nm–1424 nm (bands 107–114), 1812 nm–1951 nm (bands 153–168), and 2489 nm–2509 nm (bands 222–224) because of strong atmospheric effects.
 algorithms are regarded as proprietary (Quinlan 2002, personal communication). See5 assigns the classes by weighting the quantity of the input data, i.e. more training data of pine means there is more pine in the study area. See5 allows for the building of multiple decision trees to improve accuracy by utilizing several classifiers that predict a class. These predictions are counted and determine the final class. One type of multiple trees is “boosting”, which identifies the difficulties and mistakes made by the previous iteration and concentrates on them in the next iteration (Friedl et al. 1999, Chan et al. 2001, Wu et al. 2006). Boosting lasts until the predefined number of iterations is reached. We utilized boosting, because our data (Mannel 2003) and other studies showed an accuracy improvement by about 5% (Friedl et al. 1999, Chan et al. 2001).

One point of caution in using decision trees is “overfitting” the decision tree to the reference data visible in large “trees” (Kotsiantis 2007). A smaller “pruned” tree, with fewer branches is usually more robust than a larger tree (Quinlan 1996, Friedl et al. 1999). Therefore, we used the smallest tree with the highest accuracy for our final results.

2.5 Accuracy Assessment

We divided the reference data into different test and training data based on spatially unrelated clusters, rather than on individual points. An entire cluster was either training or test data (Mannel et al. 2006). In addition, we sought to reduce bias in our accuracy assessment by performing a 4-fold cross-validation (Mannel et al. 2011). For that we manually created four trial sets each with different test and training data that were chosen by randomly selecting 1/3 of the reference clusters as test data and the remaining 2/3 clusters as training data. Producer and user accuracies were averaged to calculate overall accuracy.

3 Results

The two pine densities were hardest to distinguish. Confusion mainly existed between the pine densities and between medium pine and
white spruce (Fig. 2). On the other hand bare soil and water were easiest to distinguish, because of their unique spectra and homogeneity of the reference pixels. This was followed by grass with almost 100 % accuracy when a spring image was included (Tab. 2). Misclassifications of meadow occurred with aspen/birch especially in the summer (AVIRIS) and early fall (fall TM).

Classification success depended on 1) the type of landcover, 2) sensor type and 3) season:

1) Landcover characteristics affecting classification success are: a) spectral properties and b) the homogeneity of the respective

20 m (AVIRIS) or 30 m (TM) mixed pixels. On the one hand, birch and aspen are virtually indistinguishable, because they have very similar spectral properties (Singhroy et al. 2000) and occur in mixed stands. Distinguishing different densities of the same forest composition is also tricky since the border between open and dense is fluent and artificial. On the other hand, we expected and achieved 100 % accuracies for homogeneous and spectrally unique bodies, such as water and bare soil. For meadow we also expected and partly achieved very high accuracies because of their relatively homogeneity (in comparison to forested areas) and

<table>
<thead>
<tr>
<th>Land cover</th>
<th>Spring TM (%)</th>
<th>Fall TM (%)</th>
<th>Two-season TM (%)</th>
<th>Summer AVIRIS (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meadow</td>
<td>99</td>
<td>86</td>
<td>99</td>
<td>88</td>
</tr>
<tr>
<td>Open pine</td>
<td>68</td>
<td>57</td>
<td>79</td>
<td>71</td>
</tr>
<tr>
<td>Dense pine</td>
<td>87</td>
<td>83</td>
<td>88</td>
<td>79</td>
</tr>
<tr>
<td>Spruce</td>
<td>86</td>
<td>68</td>
<td>90</td>
<td>83</td>
</tr>
<tr>
<td>Aspen</td>
<td>86</td>
<td>83</td>
<td>89</td>
<td>81</td>
</tr>
<tr>
<td>Average</td>
<td>85</td>
<td>75</td>
<td>89</td>
<td>81</td>
</tr>
</tbody>
</table>

Fig. 3: a) Classification of combined spring-fall Landsat TM data (left) and b) AVIRIS classification (right). The missing scene contained errors that the data provider was unable to fix. No reference data was used from the area of the missing AVIRIS scene.
their spectral properties. The higher wavelengths (>1.5 nm) provided grounds for distinguishing meadow from aspen/birch (Fig. 2).

2) Summer AVIRIS data led to 6 % better results than the comparable early-fall TM image.

3) Seasons clearly influence success of tree classifications. Classifying just the TM spring image showed an accuracy increase of about 10 % in comparison to the fall image (Tab. 2). Utilizing both spring and fall increased the TM classification accuracy by 4 % compared to just the spring season (Tab. 2). We found the two-seasonal TM classification to be about 8 % more accurate than the classification based on summer AVIRIS data (Tab. 2). All vegetation covers were better when using the two-seasonal TM approach, including meadow, aspen, spruce and the two pine densities (Tab. 2).

Fig. 3 shows the classified area based on two-seasonal TM classification and the corresponding AVIRIS flight lines.

4 Discussion and Conclusion

We found that the decision tree classification of summer-AVIRIS data provided 6 % better results than the comparable early-fall TM image. However, the summer AVIRIS classification fell short of accuracies reached via the spring TM scene. This result confirms the importance of seasonality for vegetation analysis. The AVIRIS data were taken in the middle of the summer when leaves were fully developed, which may not be the best season for a forest classification (SCHRIEVER & CONGALTON 1995). Accuracy can be further improved by combining several seasons. The two-seasonal TM data showed the best classification accuracy in all cover classes. Furthermore, multi-seasonal multispectral data is less demanding in terms of processing time and operator effort than hyperspectral AVIRIS data.

The advantages of hyperspectral remote sensing cannot be doubted, but for large-area land cover mapping applications, inexpensive multispectral data would appear to remain the sensor of choice. Multispectral data covers a larger area in a single scene, is less expensive to obtain, and requires less effort to process and analyze. Our study seems to indicate that seasonality is crucial and exceeds advantages of hyperspectral data.

Acknowledgements

The authors would like to acknowledge the efforts of Dong Hua for his collaborative input on this research, Mark Rumble and the USDA Rocky Mountain Forest Service Research Station for providing vegetation data and support with the reference data collection, Bruce Wylie and Chengquan Huang for assistance with the decision tree analysis, and Doug Baldwin and Krystal Price for collecting and processing field data. Furthermore, we would like to thank Teri Peterson for her help with the statistical analysis and Paul Cook for his review and suggestions. This research was supported by National Aeronautics and Space Administration (NASA) Food and Fiber Applications of Remote Sensing Program, Grant NAG13-99021, South Dakota School of Mines and Technology, National Science Foundation/Experimental Program to Stimulate Competitive Research (EPSCoR) Grant EPS-0091948, EPS-9720642 and the South Dakota Space Grant Consortium.

References


Bennett, D., 1984: Grazing Potential of Major Soils within the Black Hills of South Dakota. – Master of Science Thesis, South Dakota State University, Brookings, SD, USA.


Chan, J., Huang, C. & DeFries, R., 2001: Enhanced Algorithm Performance for Land Cover Classification from Remotely Sensed Data Using Bag-


LARSON, G. & JOHNSON, J., 1999: Plants of the Black Hills and Bear Lodge Mountains, South Dakota State University, Brookings, SD, USA.


MANDEL, S., 2003: Contribution of remote sensing to forest inventories in the Black Hills, SD – a focus on reference data acquisition, decision, tree classification and hyperspectral versus multispectral data. – South Dakota School of Mines and Technology, Ph.D. Dissertation.


QUINLAN, J., 2002: See5/C5.0. RuleQuest Research Pty Ltd, St Ives NSW 2075, Australia.


Addresses of the Authors:

Dr. SYLVIO MANNEL, Cottey College, Environmental Studies, USA-64772 Nevada, MO, Tel.: +1-417-667-6333, e-mail: smannel@cottey.edu

Dr. MARIBETH H. PRICE, Professor, Dept of Geology and Geological Engineering, South Dakota School of Mines and Technology, 501 E. St. Joseph St., Rapid City, SD 57701, USA, Tel. 1-605-394-2468, Fax: 1-605-394-6703, e-mail: maribeth.price@sdsmt.edu

Manuskript eingereicht: Dezember 2011
Angenommen: Mai 2012