Supervised and Unsupervised Spectral Angle Classifiers

Youngsinn Sohn and N. Sanjay Rebello

Abstract

We report that the cosine of the angle θ (spectral angle) can be utilized as a metric for measuring distances in feature space for multispectral image classification and clustering. Due to the invariant nature of the cosine of the angle θ to the linearly scaled variations, when two spectra are exactly *linearly scaled variations of one another by distance* r, *the cosine of the angle θ becomes zero while spectral distance is scaled by* r. *The fact that the cosine of the angle θ becomes zero when two spectra are* exactly *linearly scaled variations of one another implies that if we only have spectral patterns that are* exactly *linearly scaled variations of one another, then we will not be able to define* distances *between pairs of signatures for classification and clustering. For this reason, the cosine of the angle has never been considered before as a metric for multispectral image classification. According to our study, however, the fact that spectra of the same type of surface objects are* approximately *linearly scaled variations of one another due to the atmospheric and topographic effects allows the spectral angle to be used as a metric for measuring "angular distances" in feature space. Our test results indicate that the new spectral angle classifier is robust and provides better results than do the existing major image classifiers. The spectral angle classifiers do not require the data to be normally distributed, and they are insensitive to data variance and the size of the training data set. A major difference between the spectral angle classifier and conventional classifiers (ISODATA, minimum distance, maximum likelihood, decision trees, neural nets, etc.) is that the spectral angle classifier rests on the spectral shape pattern, i.e., the "identity" of the spectral pattern, while conventional classifiers rest on the statistical distribution pattern. Even though it is true for all the classifiers, especially when the spectral angle classifiers are used, the analyst's ability to relate field information to spectral characteristics and spectral shape patterns of different land-cover/land-use types is an important factor for acquiring accurate and adequate mapping results. We believe that the spectral angle classifier can potentially be one of the most accurate classifiers and a valuable tool for land-cover/land-use mapping using remotely sensed multispectral image data.*

Introduction

The fundamental premise of the remote sensing of land cover/land use is that every surface object has its own unique pattern of reflected, emitted, and absorbed radiation across the spectral bands (Parker and Wolff, 1965, cited from Campbell, 1996), and the same types of surface objects show similar spectral response patterns. For multispectral image data, spectral patterns of surface objects are presented as a list of N (number of bands) real-number components. These patterns can then be viewed as vectors in N-dimensional space and each pattern corresponds to a point in feature space. In such metric space, "similarity" is measured as the distance between two points. Two spectral patterns (signatures) that represent like objects are expected to be very close to each other in feature space (Pao, 1989). Based on this simple assumption, patterns are classified in accordance with the class membership of the nearest prototype or cluster center using the distance concept. For instance, the ISODATA and minimum-distance classifiers use the straight Euclidean distance as a metric for calculating distance. The maximum-likelihood, Mahalanobis distance, and fuzzy classifiers use transformed or weighted distances. The decision trees and artificial neural network classifiers classify patterns in accordance with whether they are on one side or another of a hypersurface or of a set of hyperplanes. For these classifiers, the Euclidean distance is the conceptual foundation, and the "similarity" is related to "distance" in feature space. It is interesting to note that, in remote sensing of natural resources and land-cover/land-use mapping, while our primary concern is the "identity" of the spectral shape pattern, we have adopted pattern recognition algorithms that are insensitive to the pattern of spectral shapes. As discussed in Sohn *et al.* (1999), the Euclidean distance, including transformed and weighted forms such as Mahalanobis and likelihood distances, is inherently insensitive to the shapes of the spectral pattern, and the cosine of the angle θ (spectral angle) provides a better definition of "similarity" due to its invariant nature to the linearly scaled variations.

Currently, the cosine of the angle θ has been utilized only for the testing of spectral similarity between the endmembers and image pixels (Spectral Angle Mapper of ENVI) or for spectral pattern matching between spectral clusters and reference classes (Sohn *et al.*, 1999). Through our experiments, we found the fact that spectral response patterns of the same type of surface objects are *approximately* linearly scaled variations of one another due to the atmospheric and topographic variations makes it possible to use the spectral angle as a metric for measuring "angular distances" for classification and clustering of multispectral image data.

In this paper we implement new spectral angle classifiers: the *supervised spectral angle classifier* (SSAC) and *unsupervised spectral angle classifier* (USAC). The spectral angle classifiers classify the image pixels based on the minimum "angular

Y. Sohn is with the Department of Geography & Environmental Systems, 211 Social Sciences, University of Maryland Baltimore County (UMBC), 1000 Hilltop Circle, Baltimore, MD 21250 (ysohn@umbc.edu).

N.S. Rebello was with the Department of Physics, 132 Peirce Science, Clarion University of Pennsylvania, Clarion, PA 16214; he is currently with the Physics Department, Kansas State University, Manhattan, KS 66506-2601.

Photogrammetric Engineering & Remote Sensing
Vol. 68, No. 12, December 2002, pp. 1271–1280.

0099-1112/02/6812–1271$3.00/0

distance" rule. The SSAC performs supervised classification using provided reference (training) signatures. The USAC performs unsupervised classification based on the minimum spectral angular distance rule using the Iterative Self-Organizing Data Analysis Technique (ISODATA). The angular distances are measured in degrees instead of radians. This is because degrees that range from zero to 90 give a better sense of spectral similarities or dissimilarities between the vectors in feature space compared to the radians that range from zero to $\pi/2$. When a pair of patterns shows the exact same shape or pattern or when they are *exactly* linearly scaled variations of one another, the angle would be zero degrees. For the pattern pair that shows completely different spectral patterns (shape), the maximum possible measurement is 90 degrees. The new classifiers were tested using Landsat 5 Thematic Mapper (TM) data. The spectral angle classifiers are written in C by the authors and compiled and run in PC DOS environment.

Spectral Angle and Spectral Distance

In two-dimensional feature space defined by bands *x* and *y*, two spectral signatures that represent two different surface objects can be represented as vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ (Figure 1). Then the spectral distance (Euclidean distance) is the length of the line segment *d* connecting the end points of the two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$. The spectral angle θ is the angle between the two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$: i.e.,

$$\theta_{\mathbf{v}_1,\mathbf{v}_2} = \cos^{-1}\frac{\mathbf{v}_1^{\mathrm{T}}\mathbf{v}_2}{\|\mathbf{v}_1\|\|\mathbf{v}_2\|}.$$

If we linearly scale the length of vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ by distance *r*, the spectral distance will be scaled by *r*. On the other hand the cosine of the angle θ between the two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ remains the same. Because of this invariant nature of the cosine of the angle θ to the linearly scaled variations, it becomes sensitive to the shape of the spectral patterns. Sohn *et al.* (1999) tested the sensitivity of the cosine of the angle θ to the spectral pattern using known field spectra of soil and sagebrush. Refer to Sohn *et al.* (1999) for further discussions on the sensitivity of the cosine of the angle θ to the spectral pattern.

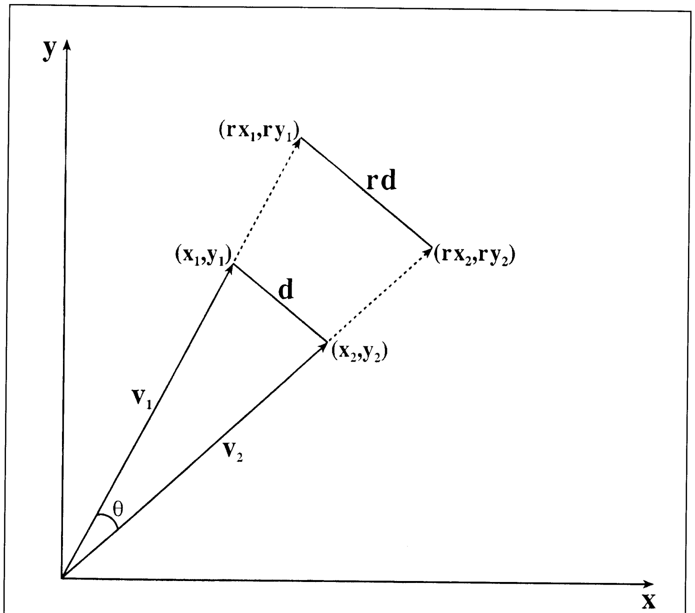

Figure 1. Spectral angle and spectral distance. Repoduced from Sohn *et al.* (1999).

When a pair of patterns is *exactly* linearly scaled variations of one another, the angle θ will be zero degrees; this implies that, if we have spectral patterns that are *exactly* linearly scaled variations of one another, then we will not be able to define distances between pairs of signatures in feature space for classification and clustering. For this reason, the cosine of the angle θ has never been considered as a metric for measuring distances in feature space for multispectral image classifications and clustering.

In reality, however, the spectra of the same type of surface objects are *approximately* linearly scaled variations of one another due to the atmospheric and topographic variations. So the actual vectors in feature space will fall slightly above or below the linearly scaled vectors. But the changes in the cosine of the angle θ caused by these variations remain very small (Sohn *et al.*, 1999). We found that this characteristic of multispectral satellite image data—that the spectra of the same type of surface objects are *approximately* linearly scaled variations of one another due to the atmospheric and topographic effects—allows the spectral angle to be used as a metric for measuring "angular distances" between patterns in feature space for classification and clustering of multispectral satellite image data.

The spectral angle classifiers we present in this study rest on the spectral "angular distances," while the conventional classifiers—ISODATA, minimum distance, maximum likelihood, Mahalanobis distance, artificial neural network, decision trees, and fuzzy—rest on the spectral distance concept. When the spectral distance concept is used, pixels that are close together in feature space will be classified together into the same class based on the statistical distribution pattern. The maximum-likelihood classifier, which is known as one of the most accurate classifiers, can be quite diagnostic in distinguishing surface objects if applied to image data that are normally distributed and include distinctive land-cover/land-use features with well-defined variances in each spectral band. Therefore, when the maximum-likelihood classifier is tested using simulated data or laboratory/field spectra collected in controlled environments, the classification results will always be very accurate. The laboratory data tend to have well-defined variances in each spectral band and do not include noise caused by atmospheric and topographic effects. In reality, multispectral satellite image data rarely show normal distributions and include noise caused by atmospheric and topographic effect; so often the data do not include well-defined variances in each band for each land-cover/land-use class. As a result, when the maximum-likelihood classifier is applied to multispectral satellite image data that do not show normal distribution and do not have well-defined variances in each band due to the atmospheric and topographic effects, we often get unreliable classification results.

When a spectral angle classifier based on "angular distances" is used, image pixels that have similar shape patterns will be classified together into the same cluster or information class. For example, consider the three spectra in Figure 2. All three spectra are from the TM image data of the study area. The spectrum A is the mean spectrum of the water training class; B is a known pixel from a pond in a golf course; and C is the mean spectrum of one of the urban/developed training classes, which represents a mixture of asphalt and concrete along the main street in downtown Clarion. Note that band 6 is the thermal band. Even with visual examination, spectra A and B can be identified as water due to their unique spectral shape pattern: generally low reflectance across the spectral bands with almost complete absorption in the water absorption bands 5 and 7. Only water will show this type of spectral pattern. On the other hand, spectrum C shows significantly increased reflectance in the near-infrared (band 4) and water absorption bands 5 and 7. Band 6 indicates the relative surface temperature. Note that

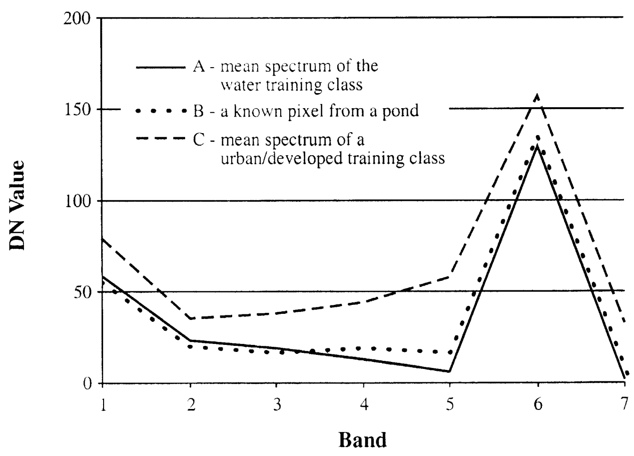

Figure 2. Spectra of a known pixel, and mean spectra of water and urban/developed training classes. The maximum-likelihood classifier classified the known pixel B, which is water, as urban/developed. The spectral angle classifier classified the pixel B as water.

both water spectra A and B have lower DN values in band 6 compared to asphalt and concrete in the downtown area. According to the classification results, pixel B is classified as urban/developed by the maximum-likelihood classifier. The relatively large variances (variances and covariances) of the urban/developed training classes and in image data are a major cause of this misclassification. Pixel B is classified as water by the supervised spectral angle classifier. This example clearly shows the behaviors of two different classifiers: the supervised spectral angle classifier and the maximum-likelihood classifier. Refer to Sohn *et al.* (1999) for further discussion.

New Multispectral Image Classifiers

The new spectral angle classifiers are implemented based on the following assumptions:

- A fundamental premise of the remote sensing of land-cover/land-use is that every surface object has its own unique pattern of reflected, emitted, and absorbed radiation across the spectral bands, and that the same type of surface objects show similar spectral response patterns (Campbell, 1996);
- Spectral similarity between two spectra can be measured in terms of the shape of spectral pattern using "angular distances" (Sohn *et al.*, 1999); and
- The fact that spectra of the same type of surface objects are *approximately* linearly scaled variations of one another due to atmospheric and topographic effects allows the spectral angle to be used as a metric for measuring "angular distances" in feature space for classification and clustering.

Supervised Spectral Angle Classifier (SSAC)

Given a set of reference signatures collected from an image with m bands, the SSAC classification consists of the following steps:

Step 1: Find the spectral angle $\theta_{i,r}$ between a pixel i in the image and every reference class r:

$$\theta_{i,r} = \cos^{-1}\left[\frac{\sum_{k=1}^{m} x_{i,k}\,\mu_{r,k}}{\sqrt{\sum_{k=1}^{m} x_{i,k}^2 \sum_{k=1}^{m} \mu_{r,k}^2}}\right],$$

where $\theta_{i,r}$ is the spectral angle between pixel i and a reference spectrum r, m is the number of bands, $x_{i,k}$ is the pixel value in band k, and $\mu_{r,k}$ is the mean pixel value of reference class r in band k.

Step 2: Assign each pixel to the reference class r that has the smallest spectral angular distance between pixel i and reference class r. For each pixel $i = 1$ to n, find the reference class r such that $\theta_{i,r}$ is the minimum for all r.

Unsupervised Spectral Angle Classifier (USAC)

The USAC performs an unsupervised classification based on the minimum spectral angular distance rule using the Iterative Self-Organizing Data Analysis Technique (ISODATA). On the first iteration, using the user specified number of clusters (n), the classifier randomly chooses n pixels in the image to be used as arbitrary means of n clusters, and calculates the cosine of the angles between each pixel in the image and the randomly chosen n cluster means. Based on the minimum angular distance rule, each pixel is assigned to one of the n clusters. After each iteration, a new mean of each cluster is calculated. Then the cosines of the angles between the new cluster means and each pixel are calculated and each pixel is reassigned to one of the clusters based on the minimum spectral angular distance rule. The iteration continues until the number of pixels remaining in the same clusters reaches the user specified convergence threshold. The following are the steps of the USAC classification:

Step 1: Select n random pixels in the image that will be used as arbitrary means of n clusters. For pixel $i = 1$ to n, $\mu_{c,k} = x_{i,k}$ where $x_{i,k}$ is one of the n pixels that was randomly selected to be an arbitrary cluster mean.

Step 2: Find the spectral angle $\theta_{i,c}$ between every pixel i in the image and every cluster mean μ_c using the equation

$$\theta_{i,c} = \cos^{-1}\left[\frac{\sum_{k=1}^{m} x_{i,k}\,\mu_{c,k}}{\sqrt{\sum_{k=1}^{m} x_{i,k}^2 \sum_{k=1}^{m} \mu_{c,k}^2}}\right]$$

Step 3: For each pixel $i = 1$ to n, find the cluster c such that $\theta_{i,c}$ is minimum for all c.

$$\text{If } \theta_{i,cmin} = Min\,(\theta_{i,c}), \text{ then } x_{i,k} \rightarrow x_{i,k,c}$$

Step 4: Recalculate the cluster means:

$$\mu_{c,k} = \frac{1}{N_c}\sum_{i=1}^{N_c} x_{i,k,c}$$

where $\mu_{c,k}$ is the cluster mean value for cluster c in band k, N_c is the number of pixels in cluster c, and $x_{i,k,c}$ is the pixel value in band k for pixel i in cluster c.

Step 5: Repeat Steps 2 and 3 until the iteration reaches the user specified convergence threshold. For each pixel, check if it is reassigned to the same cluster as before: if in the $(j-1)$th iteration $x_{i,k} \rightarrow x_{i,k,c}$, and in the jth iteration $x_{i,k} \rightarrow x_{i,k,c'}$ and $c = c'$, then the convergence level f increases by $1/N$.

Step 6: If $f > p$ (where p is the specified convergence threshold), then the classification is complete. If not, repeat Step 5.

Test of the Classifiers

Study Area and Image Data

The study area covers approximately 150 km², which corresponds to the 1:24,000-scale U.S. Geological Survey (USGS) topographic quadrangle map of Clarion, Pennsylvania (Plate 1).

Plate 1. Landsat 5 Thematic Mapper image of Clarion, Pennsylvania (RGB: 4 3 2).

The study area is located at the southern edge of the Allegheny Plateau of northwestern Pennsylvania. The Clarion River and Mill Creek corridors lie within the unglaciated Allegheny Plateau Physiographic Province. This region displays relatively flat to gently rolling plateaus dissected by deep, V-shaped stream valleys. Elevations on plateaus generally range from 460 to 490 meters above sea level, and stream valley floors range from 330 to 400 meters in elevation (Davis, 1887; Zarichanksky *et al.*, 1964). Original forest types before European settlement in the study area included white pine, hemlock-beech, and beech-maple (Hough and Forbes, 1943). Hemlock-dominated forests of the region were heavily exploited for lumber and bark for tanning during the late 1800s and early 1900s, greatly diminishing their abundance (Whitney, 1990; Abrams and Ruffner, 1995). Between 1890 and 1920, the old growth and partially cut forests were almost completely clear-cut (Marquis, 1975). The second-growth forests after the initial clear cut are now 80 to 100 years old. After harvest, many hemlock stands, particularly on upland sites, were converted to stands of early successional Northern Allegheny hardwoods dominated by red maple, oak, and black cherry (Whitney, 1990). Regular timber harvesting since the initial clear cut left the forests in the region with even-aged groups: 30 to 40 years, 50 to 60 years, and 70 to 80 years old. Silvicultural practices that perpetuate dominance of valuable hardwoods, especially black cherry, and intense browsing by white-tailed deer continue to limit the development of late successional hemlock dominated forests in much of the region (Hough, 1965; Whitney, 1990; Rooney and Dress, 1997). Today, in the study area, hemlock dominated forests are found only on the slopes in mesic coves and in stream valleys.

Extraction of fossil fuels such as coal, oil, and natural gas is an important industry in the Clarion River and Mill Creek corridors, including the study area. Bituminous coal production, the primary industry of the region, is centered in the southern portion of Clarion (Zarichanksky *et al.*, 1964). Deep mining coal production prevailed in the region from the 1870s until World War II (Davis, 1887; Miller, 1949). At present, strip mining is the predominant recovery method used in the region (Puglio, 1983). Areas denuded by strip mining are apparent south of Interstate Highway 80 in the study area (Plate 1).

Image data used for the tests is a subset of the Landsat 5 TM scene, Path 17 and Row 31, taken on 26 August 1996. The scene has shifted 30 percent southward from the center of Path 17, Row 31 to place the Clarion area at the center of the scene. The entire scene was georectified using seven 1:24,000-scale USGS topographic maps that were evenly distributed. Out of 27 identified control points, seven were used for registering the image scene to the Universal Transverse Mercator (UTM) coordinate system. For georectification, the nearest-neighbor resampling method was used to minimize the distortion in spectral integrity. The root-mean-square (RMS) error was 0.98 pixel. The subset that covers the Clarion area was created using the georeferenced image scene. The image was not converted to reflectance and not corrected for atmospheric effects because analyses do not involve inter-scene comparisons.

Supervised Classification

To test the performance of the supervised spectral angle classifier (SSAC), a set of reference signatures was provided to the SSAC, maximum-likelihood, Mahalanobis distance, and minimum-distance (Euclidean distance) classifiers, and the classification results were compared. All seven TM bands were used for collecting training (reference) signatures and for classification. A total of 23 reference signatures were collected to represent 14 major land-use/land-cover categories in the study area: waterbody, urban/developed, shrub, managed grass (golf courses and pastures), dense grass (hay fields), sparse grass (strip-mined areas), dry/senesced grass, Allegheny hardwood I (30 to 40 years), Allegheny hardwood II (50 to 60 years), Allegheny hardwood III (70 to 80 years), hemlock-dominated forest, mixed conifer plantation (red pine mixed with Norway spruce), pine plantation (red pine), and mixed forest (white pine and hardwood). In the study area, tree canopy structure starts to develop from the age of 40 to 50 years and is fully developed around the age of 70 to 80 years. Allegheny hardwoods of 30 to 40 years exhibit even, closed, and relatively smooth canopies. Fully developed canopy structures are found in Allegheny hardwoods, hemlock-dominated, and mixed forests of 70 to 80 years old. Hemlock-dominated forests are found along the steep slopes of stream valleys. These physiological and distributional characteristics of different forest types made spectral identification of various forest types and collecting training signatures relatively easy. In the study area, planted grasses that grow in strip-mined areas usually show about 50 to 60 percent ground cover. Some planted grasses in the study area grow tall but become senesced starting in early August. Grasses in hayfields show almost a 100 percent ground cover but display a lack of vigor compared to the managed grasses in golf courses and pastures. Accordingly, grasses are grouped into four different categories: sparse grass, dense grass, managed grass, and dry/senesced grass. To collect reference signatures, the Clarion area road network layer registered to the UTM projection was overlaid on the 1996 study area TM image and the reference sites identified in the field were located on the image. Then the signatures were collected from the image.

Classification results of the supervised spectral angle classifier (SSAC, 2a) and maximum-likelihood classifier (2b) are shown in Plate 2. The accuracy assessment was done using the mapping results from the SSAC, maximum-likelihood, and minimum (Euclidean) distance classifiers. Because the Mahalanobis distance classifier produced a mapping result very similar to that of the maximum-likelihood classifier, it was not included for accuracy assessment. The results of the accuracy assessment are in Tables 1A (SSAC), 1B (maximum-likelihood), and 1C (minimum distance). For accuracy assessment, 200 stratified random sample points were generated and compared to the field data. Field checking was done mostly in the summer of 1999 and another field check was done in November of 2000. When we could not confirm the 1996 land-use/land-cover information or when recent disturbances such as recent timber harvesting, new development, rotation of crop, etc. were apparent, the sample sites were discarded. A total of 124 sample points were used for accuracy assessment. The overall measured accuracies of the SSAC, maximum-likelihood, and Euclidean distance classifiers were 95.2 percent, 70.2 percent, and 93.6 percent, respectively, and the kappa values were 0.947, 0.674, and 0.929, respectively.

In general, the supervised spectral angle classifier and minimum-distance classifier generated better classification results than did the Mahalanobis distance and maximum-likelihood classifiers did. Both Mahalanobis distance and maximum-likelihood classifiers largely over-classified the surface features that show a large inter-class spectral variance such as in urban/developed, dense grass, pine, mixed conifer plantation, and Allegheny hardwood III classes while under-classifying the surface features that show a small inter-class variance such as in shrub, Allegheny hardwood II, and mixed forest (white pine and hardwood) classes. The classification results of the maximum-likelihood and Mahalanobis classifiers are not surprising when we consider the fact that both maximum-likelihood and Mahalanobis distance classifiers use class variances in each spectral band for calculating distances for classification. Both Mahalanobis distance and maximum-likelihood classifiers use parametric rules that require normally distributed data and well defined variances for image data and each training class, while most image data do not show normal distribution, and most of the training classes have high variances of pixel values in each band. The maximum-likelihood and Mahalanobis distance rules can be quite diagnostic in distinguishing different features with image data that show normal distribution and have well defined variances in each spectral band for each surface object. However, when those assumptions are violated, their performances are less than desirable.

Classification accuracy results of the supervised spectral angle classifier are presented in Table 1A with the overall accuracy of 95.2 percent. Notice from the table that mixed forest has the lowest classification accuracy due to the confusion between similar classes of mixed forest (white pine and Alle-

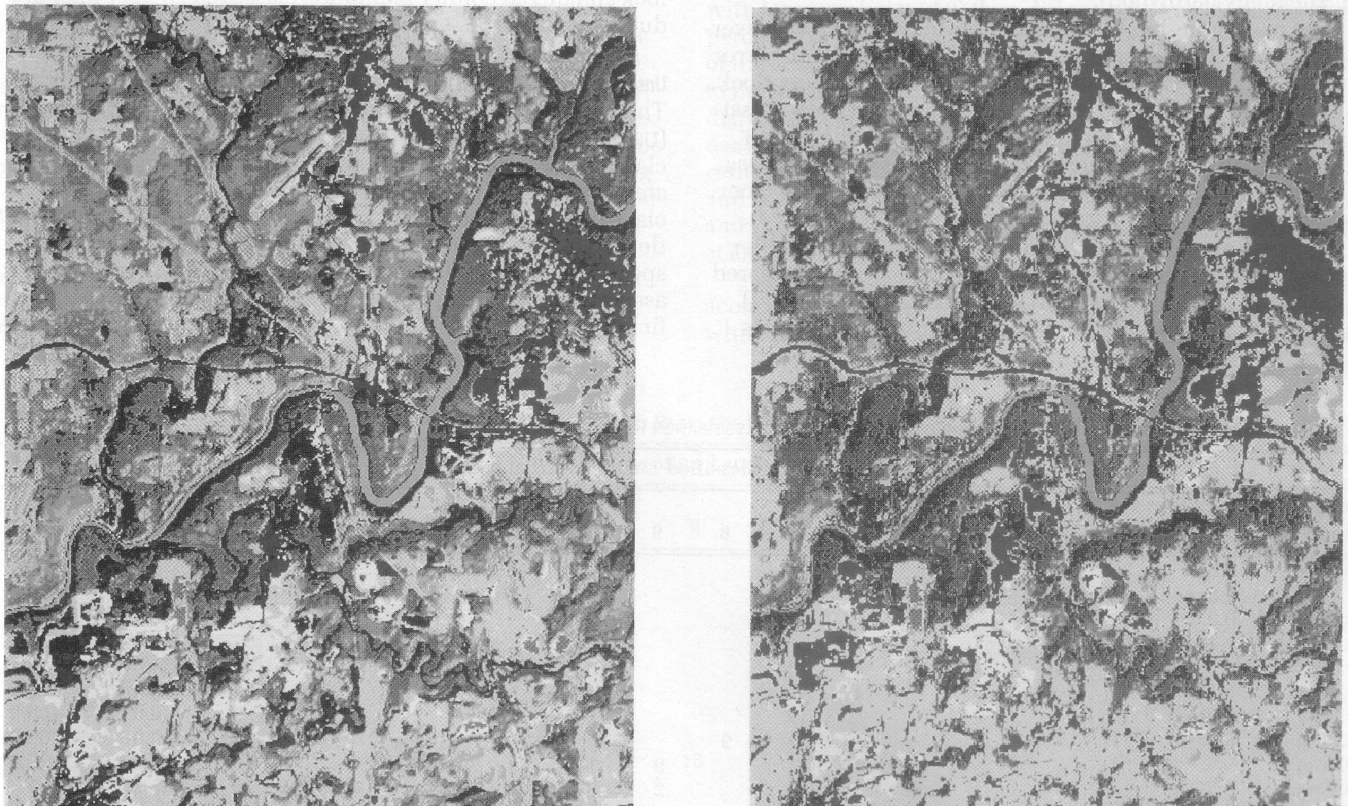

Plate 2. Supervised classification results. (a) The spectral angle classifier. (b) The maximum-likelihood classifier.

Table 1A. Error Matrix for the Classification Result of Supervised Angle Classifier (SSAC)

Actual Land Use	Classified Land-Cover/Land-Use 1	2	3	4	5	6	7	8	9	10	11	12	13	14	Row Total	Omission Error (%)	Accuracy (%)
1	9														9	0.0	100.0
2		9													9	0.0	100.0
3			13				1								14	7.1	92.9
4				10											10	0.0	100.0
5				1	14										15	6.7	93.3
6						5									5	0.0	100.0
7							7								7	0.0	100.0
8								9							9	0.0	100.0
9									13						13	0.0	100.0
10										9					9	0.0	100.0
11											4				4	0.0	100.0
12												4			4	0.0	100.0
13						1							6		7	14.3	85.7
14										3				6	9	33.3	65.7
Column Total	9	9	13	11	14	6	8	9	13	12	4	4	6	6	124	Overall Accuracy	
Commission Error (%)	0.0	0.0	0.0	9.0	0.0	16.7	12.5	0.0	0.0	25.0	0.0	0.0	0.0	0.0		118/124 = 95.1% kappa value = 0.947	

1. Water 2. Urban/Developed 3. Shrub 4. ManagedGrass 5. DenseGrass 6. SparseGrass 7. Allegheny Hardwood I (30–40years) 8. Allegheny Hardwood II (50–60years) 9. Allegheny Hardwood III (70–80years) 10. Hemlock 11. Mixed Conifer Plantation 12. Pine Plantation 13. Dry Senesced Grass 14. Mixed Forest (White Pine/Allegheny Hardwood).

gheny hardwood) and hemlock-dominated forest (hemlock and Allegheny hardwood).

The maximum-likelihood classifier shows the lowest overall accuracy of 70.2 percent as presented in Table 1B. The error matrix shows that the urban/developed, dense grass, hemlock, pine plantation, and mixed forest classes have very low classification accuracy. The mixed forest class shows the lowest accuracy of 33.3 percent. Notice that even most of the classes that show high user's accuracies exhibit high producer's (commission) errors that range from 12.5 percent to 76.9 percent.

The mapping result of the minimum-distance classifier shows much higher overall accuracy of 93.5 percent compared to that of the maximum-likelihood classifier (70.2 percent). According to the measured accuracy, even though it shows generally high user's accuracy for most of the classes, the hemlock and mixed conifer plantation classes show very high producer's errors of 33.3 percent and 50.0 percent respectively.

Unsupervised Classification

The performance of the unsupervised spectral angle classifier (USAC) was compared to that of the ERDAS Imagine ISODATA classifier, which is based on the Euclidean distance. We generated 100 and 200 spectral clusters using the USAC and ISODATA classifiers. Comparing, discriminating, and identifying the subtle differences in spectral patterns among a large number of spectral clusters (100 or more) through visual examination and assigning them into information classes will be extremely difficult. To avoid the inconsistency involved in visual interpreta-

Table 1B. Error Matrix for the Classification Result of Maximum-Likelihood Classifier

Actual Land Use	Classified Land-Cover/Land-Use 1	2	3	4	5	6	7	8	9	10	11	12	13	14	Row Total	Omission Error (%)	Accuracy (%)
1	3														3	0.0	100.0
2	5	9											1		15	40.0	60.0
3			3												3	0.0	100.0
4				7											7	0.0	100.0
5			10	4	14	1	1							1	31	54.8	45.2
6						4									4	0.0	100.0
7							7								7	0.0	100.0
8								9							9	0.0	100.0
9									9	1					10	10.0	90.0
10									2	9				3	14	35.7	64.3
11											2				2	0.0	100.0
12	1										2	4			7	42.9	57.1
13						1							5		6	16.7	83.3
14									2	2				2	6	66.7	33.3
Column Total	9	9	13	11	14	6	8	9	13	12	4	4	6	6	124	Overall Accuracy	
Commission Error (%)	66.7	0.0	76.9	36.4	0.0	33.3	12.5	0.0	30.8	25.0	50.0	0.0	16.7	66.7		87/124 = 70.2% kappa value = 0.674	

1. Water 2. Urban/Developed 3. Shrub 4. ManagedGrass 5. DenseGrass 6. SparseGrass 7. Allegheny Hardwood I (30–40years) 8. Allegheny Hardwood II (50–60years) 9. Allegheny Hardwood III (70–80years) 10. Hemlock 11. Mixed Conifer Plantation 12. Pine Plantation 13.Dry/ Senesced Grass 14. Mixed Forest (White Pine/Allegheny Hardwood).

TABLE 1C. ERROR MATRIX FOR THE CLASSIFICATION RESULT OF EUCLIDEAN-DISTANCE (MINIMUM-DISTANCE) CLASSIFIER

Actual Land Use	Classified Land-Cover/Land-Use 1	2	3	4	5	6	7	8	9	10	11	12	13	14	Row Total	Omission Error (%)	Accuracy (%)
1	9														9	0.0	100.0
2		9												1	10	10.0	90.0
3			13												13	0.0	100.0
4				11			1								12	8.3	91.7
5					14										14	0.0	100.0
6						6									6	0.0	100.0
7							7								7	0.0	100.0
8								9							9	0.0	100.0
9									13						13	0.0	100.0
10										8					8	0.0	100.0
11											2				2	0.0	100.0
12										4	2	4			10	60.0	40.0
13													6		6	0.0	100.0
14														5	6	0.0	100.0
lumn Total	9	9	13	11	14	6	8	9	13	12	4	4	6	6	124	Overall Accuracy	
mmission																116/124 = 93.5%	
ror (%)	0.0	0.0	0.0	0.0	0.0	0.0	12.5	0.0	0.0	33.3	50.0	0.0	0.0	16.7		kappa value = 0.929	

Water 2. Urban/Developed 3. Shrub 4. ManagedGrass 5. DenseGrass 6. SparseGrass 7. Allegheny Hardwood I (30–40years) 8. Allegheny rdwood II (50–60years) 9. Allegheny Hardwood III (70–80years) 10. Hemlock 11. Mixed Conifer Plantation 12. Pine Plantation 13. Dry/ nesced Grass 14. Mixed Forest (White Pine/Allegheny Hardwood).

ons, we adopted the spectral pattern matching method esented by Sohn *et al.* (1999).

To assign generated spectral clusters to information asses, first, we calculated the cosine of the angle θ between e spectral clusters and each of the provided reference signa- res. Then based on the minimum spectral angle rule, each pectral cluster was assigned to one of the information classes. or assigning spectral clusters to information classes, the same t of reference (training) signatures that were used for testing upervised classifiers was used. The same sample points used r testing the accuracies of the supervised classifiers were sed for accuracy assessment. The results of the accuracy ssessment are in Tables 2A (ISODATA, 100 clusters), 2B (USAC, 00 clusters), 2C (ISODATA, 200 clusters), and 2D (USAC, 200 usters).

With 100 generated clusters, the USAC classifier showed a slightly better mapping result (87.1 percent of overall accuracy with kappa value 0.858) than did the ISODATA classifier (86.3 percent of overall accuracy with kappa value 0.849). The results of the accuracy assessment are in Tables 2A and 2B. According to the measured accuracy, both ISODATA and USAC classifiers display a similar pattern of accuracy for each land-cover/land-use class. For both classifiers, with only 100 spectral clusters generated, hemlock shows the lowest accuracy (52.6 percent and 55.0 percent, respectively), and mixed conifer plantation and pine plantation classes were not separated from hemlock at all.

When 200 spectral clusters were generated, the overall accuracy of the USAC was increased from 87.1 to 93.5 percent, while the overall accuracy of the ISODATA classifier remained

TABLE 2A. ERROR MATRIX FOR THE CLASSIFICATION RESULT OF ISODATA CLASSIFIER (100 SPECTRAL CLUSTERS)

Actual Land Use	Classified Land-Cover/Land-Use 1	2	3	4	5	6	7	8	9	10	11	12	13	14	Row Total	Omission Error (%)	Accuracy (%)
1	7														7	0.0	100.0
2	2	9													11	18.2	81.8
3			11												11	0.0	100.0
4			2	11			1								14	21.4	78.6
5					14										14	0.0	100.0
6						6							1		7	14.3	85.7
7							7								7	0.0	100.0
8								9		2					11	18.2	81.8
9									13						13	0.0	100.0
10										10	4	4		1	19	47.4	52.6
11											0				0	—	—
12												0			0	—	—
13													5		5	0.0	100.0
14														5	5	0.0	100.0
olumn Total	9	9	13	11	14	6	8	9	13	12	4	4	6	6	124	Overall Accuracy	
ommission																107/124 = 86.3%	
rror (%)	22.2	0.0	15.4	0.0	0.0	0.0	12.5	0.0	0.0	20.0	100.0	100.0	0.0	16.7		kappa value = 0.849	

. Water 2. Urban/Developed 3. Shrub 4. ManagedGrass 5. DenseGrass 6. SparseGrass 7. Allegheny Hardwood I (30–40years) 8. Allegheny Hardwood II (50–60years) 9. Allegheny Hardwood III (70–80years) 10. Hemlock 11. Mixed Conifer Plantation 12. Pine Plantation 13. Dry/ senesced Grass 14. Mixed Forest (White Pine/Allegheny Hardwood).

Table 2B. Error Matrix for the Classification Result of Unsupervised Spectral Angle Classifier (100 Spectral Clusters)

Actual Land Use	Classified Land-Cover/Land-Use 1	2	3	4	5	6	7	8	9	10	11	12	13	14	Row Total	Omission Error (%)	Accuracy (%)
1	8														8	0.0	100.0
2	1	9											1		11	18.2	81.8
3			13	1											14	0.0	92.9
4				10			1								11	9.1	90.9
5					14										14	0.0	100.0
6						6							1		7	14.3	85.7
7							6								6	0.0	100.0
8							1	9		1					11	18.2	81.8
9									13						13	0.0	100.0
10										11	4	4		1	20	45.0	55.0
11											0				0	—	—
12												0			0	—	—
13													4		4	0.0	100.0
14														5	5	0.0	100.0
Column Total	9	9	13	11	14	6	8	9	13	12	4	4	6	6	124	Overall Accuracy 108/124 = 87.1%	
Commission Error (%)	11.1	0.0	0.0	10.0	0.0	0.0	25.0	0.0	0.0	8.3	100.0	100.0	33.3	16.7		kappa value = 0.858	

1. Water 2. Urban/Developed 3. Shrub 4. ManagedGrass 5. DenseGrass 6. SparseGrass 7. Allegheny Hardwood I (30-40years) 8. Allegheny Hardwood II (50–60years) 9. Allegheny Hardwood III (70-80years) 10. Hemlock 11. Mixed Conifer Plantation 12. Pine Plantation 13. Dry/Senesced Grass 14. Mixed Forest (White Pine/Allegheny Hardwood).

almost the same (86.3 percent vs. 87.1 percent). For the ISODATA classifier, hemlock, urban/developed, Allegheny hardwood I, and managed grass show relatively low accuracies of 57.9 percent, 75.0 percent, and 77.8 percent, respectively. Even with 200 spectral clusters, the ISODATA classifier did not separate pine plantation and mixed conifer plantation from hemlock. For the mapping result of the USAC classifier, almost all of the classes show 100 percent accuracy except for urban/developed, which shows 90 percent accuracy. When 200 spectral clusters were generated, the accuracy of pine plantation improved from zero percent to 57.0 percent. Pine plantation was mostly confused with mixed conifer plantation, which is a similar class.

Discussion

In this paper we have proposed new spectral angle classifiers based on the cosine of the angle θ as the metric for measuring spectral "angular distances" for classification and clustering of multispectral image data. According to our test results with Landsat 5 Thematic Mapper data, the new classifiers seem robust. A major difference between the spectral angle classifier and conventional classifiers (ISODATA, minimum distance, maximum likelihood, decision trees, neural nets, etc.) is that the spectral angle classifier rests on the spectral shape pattern, i.e., the "identity" of the spectral pattern, while conventional classifiers rest on the statistical distribution pattern. When

Table 2C. Error Matrix for the Classification Result of ISODATA Classifier (200 Spectral Clusters)

Actual Land Use	Classified Land-Cover/Land-Use 1	2	3	4	5	6	7	8	9	10	11	12	13	14	Row Total	Omission Error (%)	Accuracy (%)
1	7														7	0.0	100.0
2	2	9											1		12	25.0	75.0
3			12												12	0.0	100.0
4			1	9			1								11	18.2	81.8
5					14										14	0.0	100.0
6						6									6	0.0	100.0
7				2			7								9	22.2	77.8
8								9							9	0.0	100.0
9									13						13	0.0	100.0
10										11	3	4		1	19	42.1	57.9
11											1				1	0.0	100.0
12												0			0	—	—
13													5		5	0.0	100.0
14										1				5	6	16.7	83.3
Column Total	9	9	13	11	14	6	8	9	13	12	4	4	6	6	124	Overall Accuracy 108/124 = 87.1%	
Commission Error (%)	22.2	0.0	7.7	18.2	0.0	0.0	12.5	0.0	0.0	8.3	75.0	100.0	16.7	16.7		kappa value = 0.857	

1. Water 2. Urban/Developed 3. Shrub 4. ManagedGrass 5. DenseGrass 6. SparseGrass 7. Allegheny Hardwood I (30–40years) 8. Allegheny Hardwood II (50–60years) 9. Allegheny Hardwood III (70–80years) 10. Hemlock 11. Mixed Conifer Plantation 12. Pine Plantation 13. Dry/Senesced Grass 14. Mixed Forest (White Pine/Allegheny Hardwood).

Table 2D. Error Matrix for the Classification Result of Unsupervised Spectral Angle Classifier (200 Spectral Clusters)

Actual Land Use	Classified Land-Cover/Land-Use																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Row Total	Omission Error (%)	Accuracy (%)
1	9														9	0.0	100.0
2		9											1		10	10.0	90.0
3			11												11	0.0	100.0
4			2	11			1								14	21.4	78.6
5					14										14	0.0	100.0
6						6									6	0.0	100.0
7							7								7	0.0	100.0
8								9							9	0.0	100.0
9									13						13	0.0	100.0
10										11					11	0.0	100.0
11											1				1	0.0	100.0
12											3	4			7	43.0	57.0
13													5		5	0.0	100.0
14										1				6	7	16.7	85.7
olumn Total	9	9	13	11	14	6	8	9	13	12	4	4	6	6	124	Overall Accuracy 116/124 = 93.5% kappa value = 0.929	
ommission rror (%)	0.0	0.0	15.4	0.0	0.0	0.0	12.5	0.0	0.0	8.3	75.0	0.0	16.7	0.0			

Water 2. Urban/Developed 3. Shrub 4. ManagedGrass 5. DenseGrass 6. SparseGrass 7. Allegheny Hardwood I (30–40years) 8. Allegheny Hardwood II (50–60years) 9. Allegheny Hardwood III (70–80years) 10. Hemlock 11. Mixed Conifer Plantation 12. Pine Plantation 13. Dry/Senesced Grass 14. Mixed Forest (White Pine/Allegheny Hardwood).

angular distances" are used, image pixels that have similar shape patterns will be classified together into the same cluster or information class. When the "distance" concept is used, pixels that are close together in feature space, regardless of the shape of the pattern, will be classified together into the same class based on the distribution pattern. Because the spectral angle classifier utilizes the shape of the pattern for clustering and classification of multispectral image data, the analyst's ability to relate field information to spectral characteristics and spectral shape patterns for different land-cover/land-use types is an important factor for acquiring accurate mapping results. The results of our experiment clearly suggest that both the supervised and unsupervised spectral angle classifiers generate more accurate classification results than do other classifiers, and that the spectral angle classifier can potentially be one of the most accurate classifiers.

The following is the summary of the test results:

- The cosine of the angle θ can be a robust metric for multispectral image classification and clustering due to the fact that spectral response patterns of the same type of surface objects are *approximately* linearly scaled variations of one another;
- When the same set of reference (training) signatures was provided, the supervised spectral angle classifier (SSAC) produced the most accurate mapping result with 95.2 percent overall measured accuracy followed by the minimum-distance classifier (93.5 percent) and the maximum-likelihood classifier (70.2 percent);
- For unsupervised classification, when the number of clusters generated increased from 100 to 200 clusters, the accuracy of the mapping result of the unsupervised spectral angle classifier (USAC) increased significantly, while that of the ISODATA classifier remained almost the same; and
- Despite the popular belief that the maximum-likelihood classifier is the most accurate classifier, according to our experimental results, the maximum-likelihood classifier was the least accurate of all. The maximum-likelihood classifier uses a parametric rule that requires data normal distribution and well-defined variances for each band in image data and each training class. When those assumptions are violated, its performance appears less than desirable. This reconfirms that we need to pay more attention to the image data characteristics such as data normality, data variances, etc., before choosing a classifier for image classification.

The merits of using spectral angle classifiers over other classifiers include the following:

- The spectral angle classifiers do not require the data to be normally distributed. They are insensitive to data variances and to the size of the training data set as well. According to our preliminary test results, the spectral angle classifier performs consistently well in different ecoregions, including biotic communities in semiarid desert areas.
- The spectral angle classifier is less sensitive to gain factors related to topographic illumination and atmospheric effects as suggested by Kruse *et al.* (1993). As a result, topographic illumination effects and atmospheric effects will likely be less problematic in multispectral image classifications. This may allow reference spectra collected from different scenes and different imaging systems to be used for classification of images as long as both reference and image spectra are corrected for atmospheric effects and converted to surface reflectance.
- The spectral angle classifier rests on the shape of the spectral pattern, while conventional classifiers rest on the statistical distribution pattern in feature space. So when the shape of the pattern is more important than the statistical distribution pattern, as in mapping land cover/land use using the multispectral satellite image data, the spectral angle classifier is expected to perform better classifications.

Acknowledgments

This study is mostly supported by the NASA Land Cover-Land Use Change (LCLUC) program through grant NAG5-9645. The authors wish to offer special thanks to the LCLUC program manager, Dr. Garik Gutman. Even though the idea about the new classifiers was developed earlier, this work would not have been completed without encouragement from Dr. Gutman and the financial support from the NASA LCLUC program. The authors also would like to thank Dr. Charles Williams in the Department of Biology at Clarion University, Pennsylvania, for helping the authors with field identification of forest types, and Dr. Andrew Miller in the Department of Geography & Environmental Systems at UMBC and anonymous reviewers for reviewing the manuscript.

References

Abrams, M.D., and C.M. Ruffner, 1995. Physiographic analysis of witness-tree distribution (1765–1798) and present forest cover through North Central Pennsylvania, *Canadian Journal of Forest Research*, 25:659–668.

Campbell, J.B, 1996. *Introduction to Remote Sensing, Second Edition*, Guilford Press, New York, N.Y., 622 p.

Davis, A.J., 1887. *History of Clarion County, Pennsylvania.* Reprinted by Record Press (1968), Rimersburg, Pennsylvania, 936 p.

ENVI, 1999. *User's Guide*, Research Systems, Boulder, Colorado, 864 p.

Hough, A.F., 1965. A twenty-year record of understory vegetational change in a virgin Pennsylvania forest, *Ecology*, 46:370–373.

Hough, A.F., and R.D. Forbes, 1943. The ecology and silvics of forests in the high plateau of Pennsylvania, *Ecological Monographs*, 13:299–320.

Kruse, F.A., A.B. Lakeoff, J.B. Boardman, K.B. Heidebrecht, A.T. Shapiro, P.J. Barloon, and A.F.H. Goetz, 1993. The Spectral Image Processing System (SIPS)—Interactive visualization and analysis of imaging spectrometer data, *Remote Sensing of Environment*, 44:145–163.

Marquis, D.A., 1975. *The Allegheny Hardwood Forests of Pennsylvania*, General Technical Report NE-15, USDA Forest Service, Northeastern Forest Experiment Station, Broomall, Pennsylvania, 32 p.

Miller, E.W., 1949. Strip mining and land utilization in western Pennsylvania, *Scientific Monthly*, 69:94–103.

Pao, Y-H., 1989. *Adaptive Pattern Recognition and Neural Networks*, Addison-Wesley. Reading, Massachusetts, 309 p.

Parker, D.C., and M.F. Wolff, 1965. Remote sensing, *International Science and Technology*, 43:20–31.

Puglio, D.G., 1983. Production, distribution, and reserves of bituminous coal in Pennsylvania, *Pennsylvania Coal: Resources, Technology, and Utilization* (K.S. Majumdar and E.W. Miller, editors), Pennsylvania Academy of Science, Philipsburg, New Jersey, pp. 25–42.

Rooney, T.P., and W.J. Dress, 1997. Patterns of plant diversity in overbrowsed primary and mature secondary hemlock-northern hardwood stands, *Journal of the Torrey Botanical Society*, 124:43–51.

Sohn, Y., and R.M. McCoy, 1997. Mapping desert shrub rangeland using spectral unmixing and modeling spectral mixtures with TM data, *Photogrammetric Engineering & Remote Sensing*, 63(6):707–716.

Sohn, Y., E. Moran, and F. Gurri, 1999. Deforestation in north-central Yucatan (1985–1995): Mapping secondary succession of forest and agricultural land use in Sotuta using the cosine of the angle concept, *Photogrammetric Engineering & Remote Sensing*, 65(8):947–958.

Whitney, G.G., 1990. The history and status of hemlock-hardwood forests of the Allegheny Plateau, *Journal of Ecology*, 78:443–458.

Zarichansky, J., W.J. Steputis, J.R. Cerutti, H.R. Harwell, L. Seglin, and J. Hixson, 1964. *Soil Survey of Jefferson County, Pennsylvania*, USDA Soil Conservation Service, U.S. Government Printing Office, Wasington, D.C., 91 p. + 52 maps.

(Received 03 May 2001; accepted 18 May 2002; revised 21 May 2002)