URBAN LAND COVER DETECTION EMPLOYING A WINTER LANDSAT MSS DATA SET

Paul R. Baumann
Department of Geography
SUNY-College at Oneonta
Oneonta, New York 13820

Attempts to create land cover classifications of urban areas based on Landsat data have encountered several problems. These problems include: 1) a low level of accuracy in identifying urban land use patterns, 2) the separation of suburban patterns from surrounding non-urban areas especially agricultural lands, and 3) the recognition of only a few, general urban land cover patterns with specific patterns rarely being detected. Most studies using MSS data to investigate urban areas have been based on summer data sets. This study examines a winter scene covering Syracuse, New York and its environs and compares the findings from this scene to an early summer scene of the same area. The study relates the findings to the three problem areas previously indicated.

DATA SETS

A December 1976 MSS data set was used as the winter scene for this study. THe ground was clear of any snow cover and most of the deciduous trees had lost their leaves by this time. Consequently, mature deciduous trees with large crowns played a less important role in the detection of summer green areas such as parks and certain residential neighborhoods. Reducing the amount of green vegetation increased the dynamic ranges of reflectance values especially in channels 1 and 2. Most summer data sets covering areas in the Northeast have small dynamic ranges in channels 1 and 2. Except for channel 4 the winter scene's four channels produced larger dynamic ranges than the channels associated with the summer scene. Table 1 demonstrates this situation by showing the size of each dynamic range as a percent of a channel's potential data range. The overall standard deviation values for the channels, also shown in Table 1, further illustrate this point. One might expect that the larger dynamic ranges would decrease the possibility of finding good homogeneous surfaces needed to construct reliable spectral classes, but at the same time, increase the opportunity to detect more variety in the landscape, a condition generally associated with urban areas. The summer data set was taken in early May,

1977. Although a late winter snow storm hit the study area a few days after this scene was taken, the trees and bushes were in full bloom at this time and green vegetation dominated the landscape. Finding a mid-summer, cloud-free data set for either 1976 or 1977 was not possible. Also, to eliminate any variance resulting from different scanners both data sets need to be obtained from the same platform, which in this case was Landsat 2.

SPECTRAL CLASSES

SEARCH, a hybrid technique based on features common to both the supervised and the unsupervised approaches and developed by NASA's Earth Resources Laboratory, was employed to create spectral classes. SEARCH reduces the human bias aspect of producing spectral classes by automating the process; thus, both data sets were examined by the same procedures. SEARCH employs standard deviation bounds to find spectrally homogeneous training fields. These bounds were determined by using the overall standard deviation of each channel relative to the lowest deviation which was assigned the arbitrary bound of 1.26. Channels with large dynamic ranges had large standard deviations resulting in large deviation bounds. This approach was used in order to give each channel the same potential for having homogeneous surfaces to create training fields. A small deviation bound on a channel with a large dynamic range would decrease the channel's opportunity to demonstrate its homogeneous character; whereas, a large deviation bound on a channel with a small range would result in many non-homogeneous fields being accepted as training fields. Table 1 shows the bounds used for each channel. Forty-six spectral classes were generated for the winter scene in comparison to thirty-four classes for the summer scene. By generating twenty-five percent more spectral classes, the winter data set possessed the potential of either having more land cover classes than the summer data set or greater land cover accuracy.

SEARCH is limited to recording fifty spectral classes. Thus, as the program scans a data set with its window finding training fields, it must merge fields once the number exceeds fifty. Merging is accomplished by putting together the two fields with the smallest divergence. Once the entire data set has been scanned for fields, the fifty spectral classes are merged one final time and all classes below a set divergence bound are merged together. With forty-six classes, the winter data set experienced less merging through the final divergence process than the summer data set with thirty-four classes. Thus, the winter data set produced spectrally more distinct classes. This situation resulted from the larger dynamic ranges associated with the winter data set channels.

LAND COVER CLASSIFICATION

A maximum likelihood classifier was used to assign each pixel to a spectral class and each spectral class based on its spatial pattern and spectral signature was associated with a particular land cover condition. Some spectral classes possessed well-defined spatial patterns making it easy to identify the related surface features. Although such classes might have clear spatial patterns, they would be associated with more than one surface feature. Other spectral classes displayed what might be referred to as a "pepper" pattern, a sprinkling of individual pixels across the scene. Determining the exact locations of individual pixels on the support photography was nearly impossible; consequently, these "pepper" pattern classes were very difficult to relate with surface features. Only a few spectral classes were identified with land cover conditions based solely on their spatial characteristics.

The land cover classification for most spectral classes was determined by both spatial patterns and spectral signatures. In most cases a signature was used to make the final determination of how a class should be identified. Figures 1 and 2 show sample signatures of six land cover classes for the winter and summer data sets respectively. Each data set produced six land cover classes; consequently, the winter scene did not create any more classes than the summer scene. However, the signatures associated with the land cover classes for the winter scene displayed a tighter correlation than the signatures for the summer scene. The signatures in Figures 1 and 2 associated with commercial land cover illustrate this point. The category designated as "open area" was of particular interest. This class has been frequently used to deal with the confusion related to separating agricultural fields from suburban patterns. In Figure 2, the summer scene, the signatures for open areas show considerable variance and some signatures are very similar to those related to residential and woodland conditions. In comparison, the signatures in Figure 1, the winter scene, for open areas demonstrate little variance and are markedly different from the signatures of the other classes. Thus, based on signatures, the winter scene appears to be better suited for separating residential and agricultural land covers. Also, the winter scene's signatures show a strong correlation within the specific land cover classes and a marked separation between the classes. Relative to the summer data set, the winter scene produced more distinguishable land cover signatures.

THRESHOLD ACCURACY

In testing for classification accuracy a variety of issues occurred, especially with determining spatial accuracy. This study deals with the accuracy issues by reclassifying the data sets through the maximum likelihood classifier using a three standard deviation threshold level on each spectral class. Pixels falling outside the level were not classified. In other words, pixels which were not within a three standard deviation range of any of the spectral classes were left unclassified. The winter scene had 43.41 percent of its pixels not classified; whereas, the summer scene had 62.31 percent. Both data sets had 87,500 pixels. One would assume that with small dynamic data ranges in two of its four channels and a great amount of merging of spectral classes which frequently results in large standard deviations, the summer data set would have more pixels being classified using a threshold level than the winter data set. With a higher percent of pixels being classified, the winter data set's spectral classes appear to be better fitted to the data than the spectral classes associated with the summer data set. Table 2 shows the percent of pixels classified for each spectral class by data set and the land cover for each class. The winter scene has seventeen spectral classes with percentage values above 60 percent in comparison to eight such classes in the summer scene. Five classes in the summer scene have below 10 percent; whereas, the winter scene has none. When initially assigning land cover classes, seven of the top ten spectral classes in the winter scene were easily related to land cover patterns. Only three of the top ten classes in the summer scene related easily to specific land covers. In general, the winter scene produced better accuracy results than the summer scene. Also, many of the spectral classes with high threshold percentages were associated with urban land cover patterns.

As a footnote, the water and mining classes had well-defined spatial patterns indicating a high accuracy level. However, Table 2 shows both classes in both data sets as having moderate to low threshold values. This contradiction between having a high spatial accuracy and a low threshold accuracy might relate to both classes possessing noticeably different spectral conditions in four dimensional statistical space than the other land cover classes. Figures 1 and 2 show the water and mining classes as having unique signature patterns. The SEARCH procedures for merging of spectral classes and looking for homogeneous surfaces might not have found either enough separation between training fields for these land cover classes or enough homogeneous training fields to represent these classes. In either case, the threshold accuracy test is still a valid procedure.

FINDINGS

The winter scene gave better separation of classes as indicated by the spectral signatures. The residential areas showed more separation from the open areas in the winter scene than in the summer image. However, the urban open areas such as parks and golf courses did not separate well from the non-urban open areas such as pastures and fields. The winter scene demonstrated much better accuracy based on the threshold test than the summer scene. The better separation and accuracy results found with the winter scene apparently related to the large dynamic ranges associated with the four winter channels. In turn, the large dynamic ranges related to the decrease in green vegetation surfaces. Both data sets produced the same number of land cover classes. Some indication did exist that the summer data set could have more land cover classes but the accuracy of these classes would be in question. Considerably more work needs to be done on winter scenes but this study did indicate promising results.

	Dynamic Range As Percent of Potential Range	Overall Standard Deviation	Standard Deviation Bound Used in SEARCH
May, 1977 Data Set			
Channel 1	38.5	5.25	1.26
Channel 2	48.0	7.37	1.73
Channel 3	64.5	11.89	2.84
Channel 4	79.3	7.54	1.80
Dec., 1976 Data Set			
Channel 1	51.2	12.68	3.32
Channel 2	74.8	17.28	4.12
Channel 3	74.8	16.94	4.04
Channel 4	61.9	6.31	1.50

Table 1: Comparative MSS Channel Data

	Winter Scene			Summer Scene		
	Threshold Percent	Spectral Class	Land Cover Class	Threshold Percent	Spectral Class	Land Cover Class
1	88.75	(26)	Residential	78.77	(6)	Woods-Wetlands
2	74.35	(17)	Wooded Areas	76.20	(27)	Residential
3	73.29	(5)	Residential	76.09	(4)	Open Areas
4	72.02	(13)	Open Areas	69.23	(10)	Residential
5	70.30	(11)	Residential	68.08	(13)	Woods-Wetlands
6	69.55	(7)	Open Areas	68.07	(11)	Woods-Wetlands
7	69.03	(10)	Open Areas	66.90	(9)	Residential
8	68.86	(25)	Commercial	60.48	(20)	Woods-Wetlands
9	67.91	(24)	Commercial	59.20	(14)	Woods-Wetlands
10	67.21	(8)	Residential	58.21	(32)	Open Areas
11	66.13	(3)	Commercial	54.76	(28)	Woods-Wetlands
12	65.69	(40)	Wooded Areas	51.33	(3)	Woods-Wetlands
13	65.62	(18)	Open Areas	51.06	(1)	Open Areas
14	63.17	(43)	Wooded Areas	49.42	(19)	Woods-Wetlands
15	63.12	(14)	Residential	48.17	(33)	Residential
16	63.02	(45)	Wooded Areas	47.02	(5)	Open Areas
17	62.34	(12)	Residential	44.79	(18)	Water
18	58.99	(9)	Residential	41.06	(26)	Residential
19	58.82	(35)	Wooded Areas	39.27	(17)	Open Areas
20	57.86	(21)	Open Areas	31.96	(12)	Residential
21	57.64	(46)	Wooded Areas	31.95	(34)	Open Areas
22	57.52	(31)	Water	29.14	(16)	Woods-Wetlands
23	56.88	(19)	Open Areas	28.82	(2)	Woods-Wetlands
24	56.72	(16)	Residential	24.72	(21)	Residential
25	55.47	(1)	Open Areas	23.14	(7)	Open Areas
26	55.32	(4)	Residential	21.76	(31)	Residential
27	55.29	(41)	Commercial	21.04	(15)	Open Areas
28	54.94	(2)	Open Areas	19.70	(30)	Woods-Wetlands
29	52.07	(29)	Commercial	11.54	(23)	Mining Area
30	51.54	(42)	Commercial	8.29	(22)	Water
31	51.49	(6)	Residential	8.02	(25)	Commercial
32	49.67	(32)	Wooded Areas	6.72	(24)	Commercial
33	48.92	(20)	Open Areas	3.33	(8)	Commercial
34	48.55	(44)	Wooded Areas	2.44	(29)	Open Areas
35	48.45	(34)	Open Areas			
36	48.21	(15)	Open Areas			
37	46.81	(27)	Open Areas			
38	45.49	(28)	Residential			
39	42.98	(37)	Residential			
40	39.50	(22)	Open Areas			
41	33.48	(23)	Open Areas			
42	31.23	(30)	Water			
43	30.06	(33)	Open Areas			
44	20.11	(39)	Open Areas			
45	20.00	(38)	Mining Area			
46	15.21	(36)	Mining Area			

Table 2: Spectral Class Threshold Level

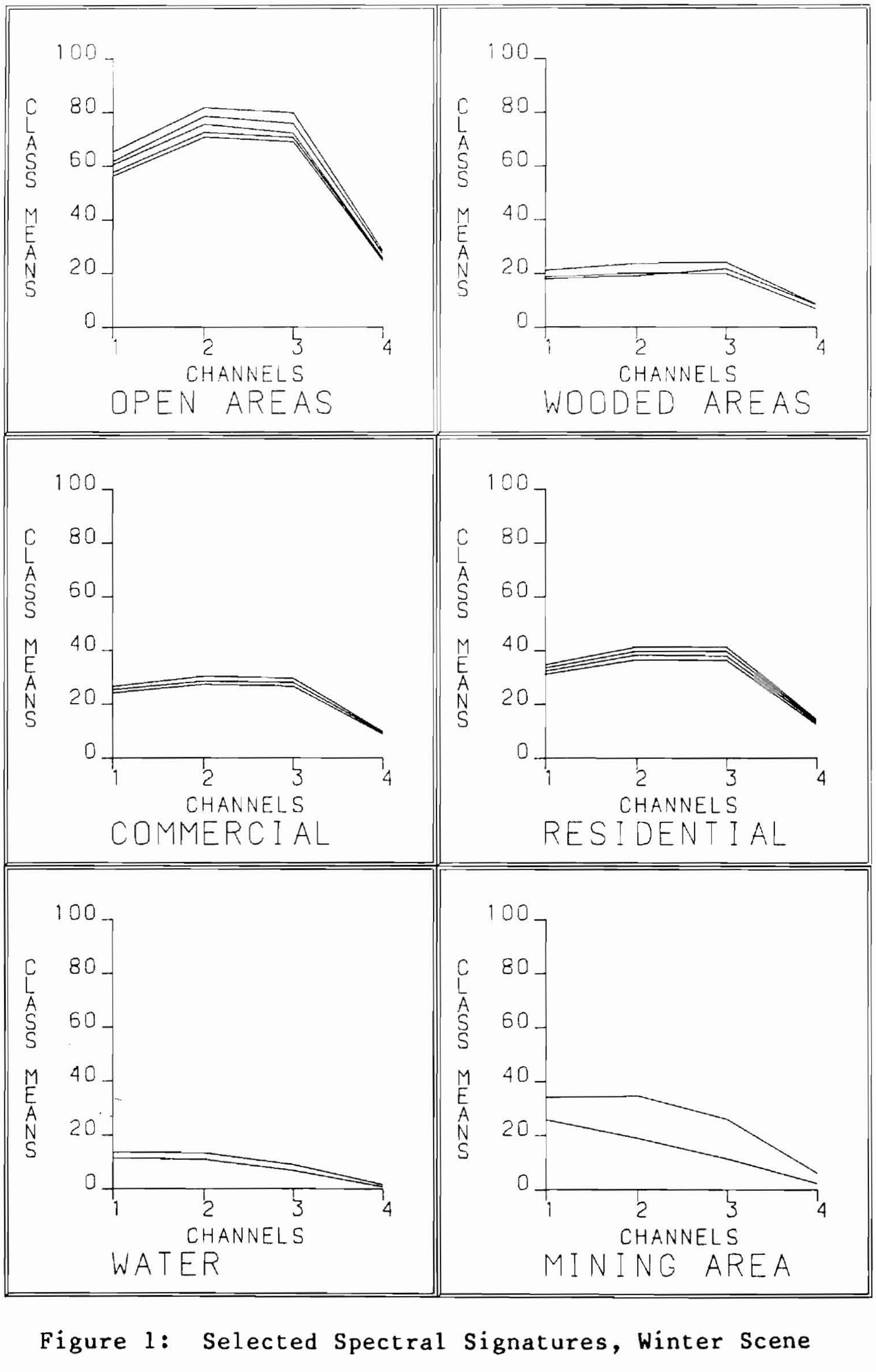

Figure 1: Selected Spectral Signatures, Winter Scene

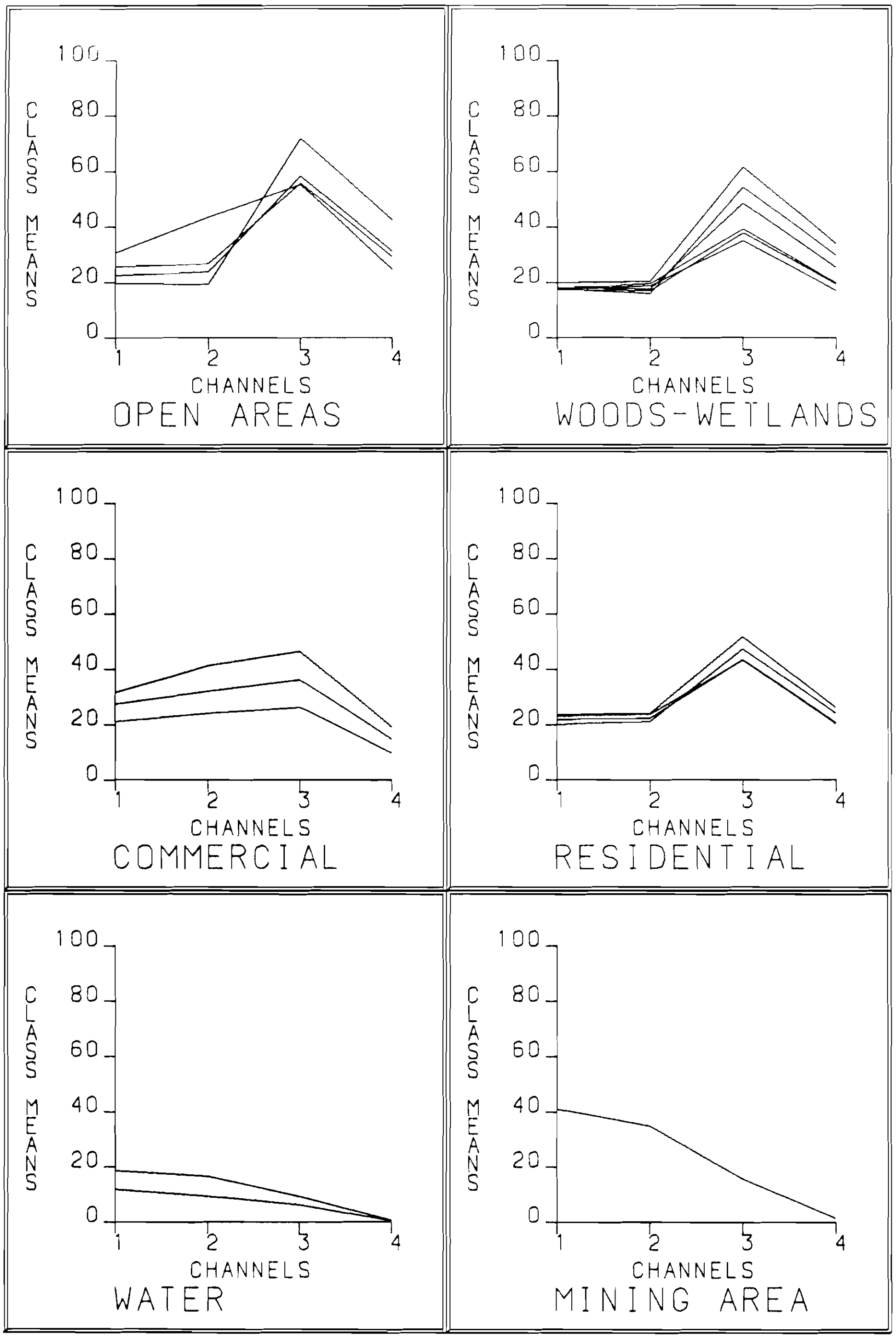

Figure 2: Selected Spectral Signatures, Summer Scene