DIFFERENTIATING A SMALL URBAN AREA FROM OTHER LAND COVER CLASSES EMPLOYING LANDSAT MSS

Lon Arowesty
Department of Geography
SUNY-College at Oneonta
Oneonta, New York 13820

For over a decade, Landsat data have been successfully used for mapping, vegetation assessment, and a variety of other applications. One area of focus has been urban areas. The majority of research which employed Landsat Multi Spectral Scanner (MSS) for urban areas has involved large urban centers. Little attention has been given to small cities, especially those situated within rural environments. This paper will express some of the problems related to the classification (or misclassification as the case may be) of these small settlements when creating a general land cover map employing Landsat MSS. Oneonta, New York, a small city of approximately 20,000 people, and its environs were chosen as the study area because the urban area had land cover features which should be discernible within the Landsat MSS limited picture element (pixel) resolution of 56m x 79m.

STUDY AREA

The study area is a subscene based upon the City of Oneonta and its rural surroundings (See Figure 1). The data were taken from an October 11, 1972 Landsat-1 MSS image of Central New York State. The study area can be defined easily with reference to lakes and rivers. The eastern boundary of the image is the middle of Crumhorn Lake. To the south Treadwell Creek is the marker. The western border is the east facing slopes of the Otego Creek Valley. The northern border can be found at Little's Pond just northeast of Arnold Lake. The valleys in the study area are open land for the most part, with fringes of tree cover along the Susquehanna River and is tributaries. These valleys also have some wetland forests located on meander scars as well as around oxbow lakes. Going up the hills, the gentler slopes are generally occupied by cleared pasture while the steeper portions of the hills are predominantly woodlands. The steepest sections have rock outcrops as well as woodlands. Reversion of land to its natural land cover has been increasing with the movement of dairy farming downslope

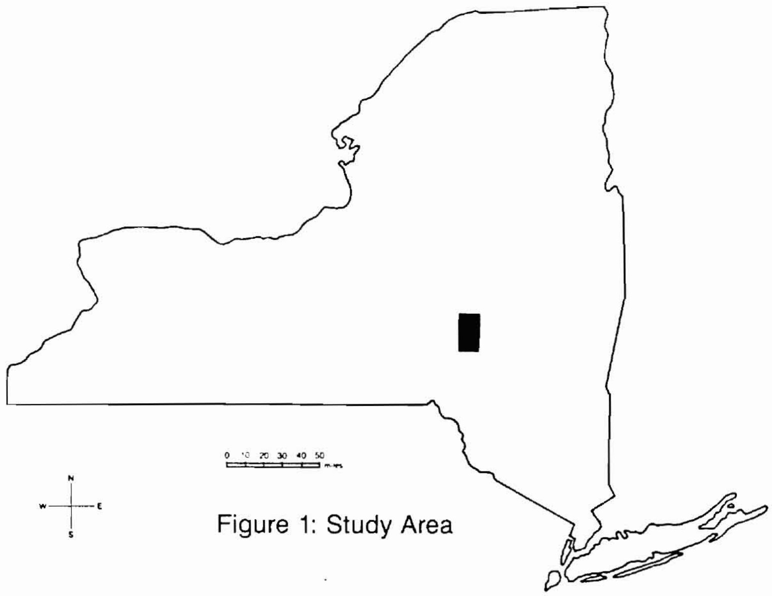

Figure 1: Study Area

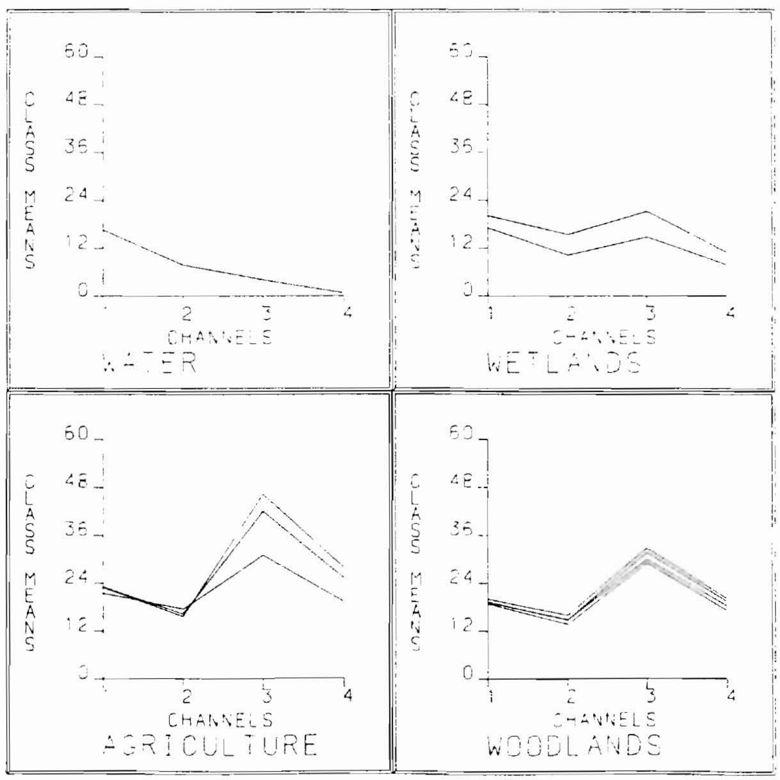

Figure 2: Spectral Signatures

and the abandoning of farms on the flatter uplands. Abandoned or unoccupied farms are numerous in the Otsego County area.

Oneonta is located within the Catskill Region of New York on the banks of the Susquehanna River. The city is nearly midway between Albany, NY, and Binghamton, NY. The urban area of Oneonta comprises approximately 3.5 square miles and has a population of nearly 20,000 7,000 of which are college students. Within the city is a large old residential section which has many miles of tree covered streets. Most of the housing stock was built prior to 1939. On the north and west boundaries of the city lie two college campuses, State University College at Oneonta and Hartwick College, respectively. On the southern border of the city are the Susquehanna River and the D & H Railroad yard. The river is not more than 200 feet wide in most places in the area. It should be noted that during the time period the image was taken the river path was being altered by the construction of Interstate 88 which parallels the river throughout the area.

METHODS AND MATERIALS

A Landsat MSS scene of Central New York State, recorded on October 11, 1972, was selected because of the excellent clarity of the image and the inclusion of a vast rural landscape. A subscene was then selected as a study area. The subscene had an area of 325 pixels x 348 pixels (236 sq.mi.). Support material included a set of aerial photographs at the scale of 1:14000 and USGS topographic sheets at a scale of 1:24000. The aerial photography was taken April 23, 1973, with little change occurring from the date of the image until the following spring. The same cannot be said for the topographic sheets. Five sheets covering the area were used. The dates of these maps ranged from a 1943 topographic sheet for Oneonta to the most up-to-date 1982 series. The 1982 series was photo-revised yet not topographically revised. Much change has occurred during the time lapse. Also cartographic generalization renders the topographic sheets less valuable a tool as far as land cover characteristics are concerned. Still, the topographic sheets were a very useful tool for establishing general land cover conditions.

The statistical analysis of the MSS data was accomplished through the use of the software package termed LAP (Landsat Analysis Package). LAP is a series of programs which develops statistics based on the spectral reflectances of pixels and then classifies the pixels. First, frequency counts and histograms were generated for the four spectral channels. The purpose was to determine the dynamic ranges and obtain the means and standard deviations of each channel (see figure 2). Of the four spectral bands, channel 3 had the largest dynamic range and standard deviation. An index was created by dividing channel 3's standard deviation into the standard deviation of each of the other channels and then multiplying each value obtained by a constant.

Next, a technique developed by NASA's Earth Resource Laboratory termed SEARCH was implemented. SEARCH is a mathematically supervised procedure that produces unsupervised spectral classes. SEARCH works by moving a 6 pixel x 6 pixel window over the study area searching for homogeneous training fields. Homogeneity is defined by three statistical measures. The first measure is the acceptable standard deviation range for each channel which is delimited by what is referred to as the lower and upper bounds. The bounds were derived from the index procedure previously discussed. The second measure of homogeneity was the acceptable coefficient of variation (SD/mean). The final measure was divergence or scaled distance depending upon the choice of the operator and the size of the window. SEARCH has a second window size option of 3 pixels x 3 pixels. Because of the topography and the heterogeneous nature of the study area the smaller window option was used. Opting for the smaller window forced the use of scaled distance as the third measure because the option of divergence is not available for the smaller window. The use of SEARCH enabled quick generation of spectral classes from the homogeneous training fields found. The output from this process was stored in a file termed the statistical file.

The individual pixels were then classified taking the created statistical file and employing it with a maximum-likelihood-classifier program. The process created a classified data file, by assigning individual pixels to spectral classes. A line printer map was made from the classified data file. By combining spatial relationships and spectral signatures, represented by the statistics generated by SEARCH, grouping of spectral classes into land cover classes occurred.

With water features being the easy items to recognize, this class was located first on the line printer map. Water bodies readily stand out when using Landsat MSS because these features have low reflectance values in channel 4. From these valuable reference points, further classification was performed. Next, other water related features were brought out. Wetlands also have relatively low reflectance values in channel 4. The wetland class was used to find the Susquehanna River in the valley. By going from the valleys up the slopes of the hills classification was performed. The reason for this method for classification was that is seemed consistent with the land use of the settlement patterns in the region.

The support data, aerial photography and USGS topographic sheets, were used as reference materials and for the selection of training fields. Referencing between the line printer map and the support material was cumbersome and laborious. Thus, the information stored in the statistical file and the classified data file on the mainframe were transferred to a microcomputer with a high resolution monitor. This transfer allowed for the classified file to be viewed in an interactive graphics mode. The final classifying of the spectral classes was performed on the monitor by examining spatial relationships and spectral signatures of the classes.

A final procedure was an accuracy measurement. Accuracy was measured by placing a threshold of three standard deviations in the maximum likelihood classifier and merging the statistical file with the classifier. The accuracy is determined by the percentage of pixels that get classified into one of the spectral classes when a threshold is applied.

RESULTS

SEARCH found 1,132 training fields which were then condensed into the thirty-seven spectral classes used to classify the Landsat MSS data. These spectral classes, generated by SEARCH, were grouped into four land cover categories: water, agriculture, wetland, and woodland. The overall accuracy of the product, measured by placing a threshold of three standard deviations with the maximum-likelihood classifier, was low. Only 38 percent of the pixels were assigned to a spectral class when the threshold was used. Woodlands had the best accuracy of all categories. Spectral classes 22 and 26, both woodland, had 84 percent and 82 percent accuracies respectively. Still, urban areas were grouped into wetlands and some conifers were also classified as wetland.

DISCUSSION

One might expect that relatively few problems exist with such a simple task as discriminating between an urban area and its rural surroundings. Some might state that either the area is urban or it is not, while others might want a specific type of land use. The purpose of this research was to relay problems related to the use of Landsat MSS data in differentiating a small urban setting from its rural surroundings.

Much of the inaccuracy of the results can be attributed to one of the following: the maximum-likelihood classifier, Landsat MSS pixel resolution, spatial aspects of an urban place, and/or morning shadows. First, the overall accuracy of the product must be examined. The large number of pixels not classified was caused by reflectance values of pixels falling outside of the threshold of the equiprobability contours on the scattered diagram used by the maximum likelihood classifier. The lack of a limiting factor such as a threshold, forced all pixels to be classified into a "best fit" situation. This is not to say the use of a threshold will guarantee accuracy but rather challenge the actual results.

The resolution of the MSS system used was definitely a limiting factor. The 59m x 79m pixel, nearly 1.1 acres, is very useful when dealing with large homogeneous areas. When trying to establish spectral characteristics for a heterogeneous area such as an urban place, better resolution would be beneficial. The system employed collected spectrally averaged data. Being spectrally averaged, individual pixels often contained more than one land cover type for which reflectance information was recorded. The averaged data created problems especially in the downtown area of Oneonta. The reason for these problems will be discussed later. Other areas where averaging of reflectance values occurred were along the Susquehanna River where land and water as well as mud surfaces were combined, on the shores of the lakes in the area, and at other transitional zones.

For a variety of reasons settlement patterns tend to be located on flat land. Oneonta was a major railway town in the late nineteenth century with the D&H line as its principal railroad. The rolling hills of the area forced the tracks to be laid in the flat valley bottoms since steam engines could only tolerate a maximum gradient of 2 percent. The traprock which the rails and ties are laid upon is grey. The grey shade of this basalt type material has a high reflectance value. The combination of this high reflectance surface with existing wetland conditions lining the tracks causes an urban feature to be misclassified.

Roads also are laid in the valleys as well as on the terraces lining the valleys. Interstate 88 (I-88) was still under construction when the study area was photographed. From the southwestern border of the image to the Town of Emmons, neighboring Oneonta to the east, construction was occurring. The river's natural course was diverted so the interstate could be built on the old river bed. In the downtown section of the city is an old mill run which in the past had generated energy for some of the mills. The mill run parallels the railroad tracks. With the river straightened by the diversion, four linear features were aligned in the valley; the railroad tracks, the mill run, a stretch of the Susquehanna River and I-88. These features are physically close enough in certain sections of the valley to produce a combined reflectance value that gets classified as wetlands. Thus, the bare ground of I-88 with high spectral reflectance gets averaged down by the low reflectance value of the Susquehanna River when the two features are within the same pixel. The mill run/railroad tracks area incorporates a wider variety of land cover than the other situation. This area is not just grass, water, and bare ground, but rather it contains tree covered wetlands, the mill run, bare rocks, and various urban patterns.

The railroad yard stands out on the classified scene. Of the three features expected to be found within the city, the railroad yard was the only one recognized. The area was easily discerned spatially on the color monitor, yet the spectral reflectances of the pixels falling within the yard were more characteristic of wetlands. Coal dust from many years of rail transport of the material has covered the tracks as well as the areas in between rendering the ground black. These low reflectance values forced the pixels in the yard to be classified as wetland. The misclassification of the yard as a wetland added to the inaccuracy.

Another situation occurred that created error. Morning shadows on the north and northwest facing slopes lowered reflectance values of the land cover. These misclassifications involved conifers, usually hemlock, being included in the wetland class. Areas that had this error were east of

Goodyear Lake. The finding of wetlands on such sloped area further confused the issue of classifying wetland environments.

CONCLUSION

Heterogeneous areas, especially urban developments, are extremely difficult to classify correctly into land use categories when using Landsat MSS data. Yet, it should be possible to delineate small urban areas with such data. The misclassification of urban areas as wetlands was the main problem encountered with the research conducted. The limitations of the system and the internal spatial arrangement of Oneonta combined to create the problem.

ACKNOWLEDGEMENT

The author would like to thank Mr. Paul R. Baumann, Chairman of the Department of Geography, for providing the necessary guidance and support during the research.