Reprinted from

Symposium on Machine Processing of Remotely Sensed Data

June 27 - 29, 1979

The Laboratory for Applications of
Remote Sensing

Purdue University
West Lafayette
Indiana 47907 USA

IEEE Catalog No.
79CH1430-8 MPRSD

THE USE OF LANDSAT MULTISPECTRAL DATA TO DERIVE LAND COVER INFORMATION FOR THE LOCATION AND QUANTIFICATION OF NON-POINT SOURCE WATER POLLUTANTS

HENRY F. FOSTEL
Baltimore Regional Planning Council

JAMES E. MANLEY
OMNI Information Services

JAMES P. ORMSBY
Goddard Space Flight Center

I. INTRODUCTION

The Regional Planning Council (RPC) provides a clearinghouse and advisory service to the city and five surrounding counties of the Baltimore region (Anne Arundel, Baltimore, Carroll, Harford, and Howard counties and Baltimore City.) The RPC staff is divided into various technical divisions providing support in the numerous aspects of modern planning. A great portion of this work is in direct response to the requirements of various Federal programs and legislation. In early 1976, the RPC recieved a Water Quality Planning Grant under the Federal Clean Water Act Amendments of 1972 (PL-92-500, Section 208). The work program for this grant required a detailed, up-to-date knowledge of the land cover for use in water pollution models. In addition, the Section 208 Planning Grant was to be used to raise the level of local goverment expertise wherever practical in data collection procedures, use of different data sources, modeling, computer classification procedures, etc. Various methods of data processing, remote job entry (RJE) and computer interactive systems (Larsys, Elltab, Orser, GE Image-100, Bendix, etc.), available for processing LANDSAT and other sensor data, were evaluated to determine how well they could meet the Section 208 Project's needs and goals.

II. BACKGROUND AND OBJECTIVES

One of the most unique portions of the work to be done under Section 208 was the categorization, location and quantification of non-point source pollutants. These pollutants are, in fact, directly attributable to man's activity on land. Various pollutants carried from the land during and after rainfall events include fertilizers and pesticides from agricultural feed lots and overflowing septic tanks, oil and tar residues along road sides and curbs, and even salts used in winter-road deicing. In order to estimate the location and scale of activity, the landuse/cover must be identified.

The identification of landuse/cover by experienced staff members has been an important part of the RPC's past studies. A detailed landuse inventory was prepared for the years 1964, 1970, 1973, and 1975 (RPC, 1975 and RPC, 1976).[1,2] However, this inventory included only those landuses considered as developed. Previously, developed land had been the prime interest of the RPC. However, in order to estimate pollutants from all types of landuse/cover, a detailed inventory of both the developed and undeveloped land was required.

In addition, the RPC was in the process of preparing an automated inventory of the developed land uses. This new inventory would have the capability of being summarized by planning districts, census tracts, natural watersheds, sewersheds, or whatever boundaries were required. It was apparent that an inventory of undeveloped land should also be automated and have the similar capability of summarization. The RPC staff had various options to consider: (1) the inventory could be done by hand, at considerable time and expense, and coded into the automated systems in the manner that the developed land had been; (2) the Maryland State Planning Department's Maryland Automatic Geographic Information (MAGI) System could have been used to supply summaries of the undeveloped land in the region; (3) the inventory could have been developed from analysis of LANDSAT data on a commercial image classification system such as that of General Electric or Bendix; (4) the land cover inventory could be developed from analysis of LANDSAT data by RPC personnel, in cooperation with Intralab at Goddard Space Flight Center.

Judging by the time and effort that had been involved in coding the developed land both in the original inventory and the automated system, it was obvious that the RPC staff could not afford the time or the expense of similarly encoding the undeveloped land. The first option was eliminated.

The second option promised a quick and inexpensive inventory of the undeveloped land. However, the data of the MAGI System had two drawbacks. First, having been coded on a 91.8 acre grid, the data in many cases was larger than the Section 208 prototype sampling areas and many categories of the data were too generalized and lacked the detail necessary for the non-point source evaluation. Further, the MAGI data was 1973 vintage, whereas, the water quality sampling data to which it was to have been correlated with was undertaken in 1977. So, the age of the data and its generality precluded favorable acceptance.

The third option was evaluated and considered too costly. GE had completed a study of the region's water supply for the Baltimore City Public Works Department. The cost of that study, for an area slightly less than 20% of the region, was $16,000. The EPA Grant money was not available for such major expense because it had been already committed to the balance of the water quality program. This eliminated the third option.

The fourth option was not immediately available to the RPC at the beginning of the Section 208 program. The decision to attempt to develop a LANDSAT inventory and utilize it in the non-point source correlation came nearly midway through the program. It was decided that the NASA LANDSAT program would be able to provide the land cover data needed, especially in the under-developed (rural) areas of the region. Investigations of documented uses of the LANDSAT sensor data revealed that there still were not real successful uses of LANDSAT data in identifying specific urban uses (other than gross densities of residential use, paved area, roof tops). Therefore, it was decided to wed the RPC information on developed land uses, from the air photo interpretation and the land cover from the LANDSAT sensors. It was anticipated that this would yield the most accurate result by utilizing the best from each technique. The fallback position still would be the data available in the State's MAGI System. Of additional importance in the final decision to utilize the LANDSAT data was the likelihood that other planning activities would also be able to utilize the data in local county planning, zoning, and permitting departments throughout the region.

III. GENERAL PROCEDURES IN THE ANALYSIS OF LANDSAT DATA

A. DATA

Summer August 3, 1975/5106-14543 and Fall November 19, 1975/5214-14480 LANDSAT-1 scenes of the Baltimore-Washington, D.C. region were selected for detailed analysis of the five counties and Baltimore City. These scenes were selected to take advantage of temporal changes in the forest canopy and agricultural lands, complete coverage of the region on both scenes, and availability of the data. Additional material included: (1) black and white aerial photographs at a scale of 1:24,000 taken on October 22, 1975 by NASA which had been previously obtained from Photo Science; (2) 9" x 9" color infrared (IR) aerial photographic transparencies at a scale of 1:36,000 taken during October, 1976 by NASA; (3) RPC maps at a scale of 1:24,000 and records classifying developed land uses; (4) USGS's 7½ Minute Series (topographic) Quadrangle maps of the region; and (5) individual maps of selected areas at a scale of 1:24,000 prepared by the RPC staff during field checks and source interviews.

B. PRELIMINARY PROCESSING

The original raw LANDSAT multispectral scanner subsystem (MSS) data tapes were sent to the Office for Remote Sensing of Earth Resources (ORSER) computer system via remote job entry (RJE) terminals at Intralab. The southern portion of Anne Arundel County was chosen for initial training purposes because of its extensive land/water boundary with Chesapeake Bay, thus providing easy identification of the area, close proximity to GSFC for field checking, Intralab staff familiarity with the area, larger parcels of homogenous land area compared with other areas of the region, and the location in the area of the Chesapeake Bay Center for Environmental Studies (CBCES) within the Rhodes River drainage basin. The CBCES later provided useful local inventories of past landuses and ground covers used to verify signatures. While GSFC staff and Operations Research, Inc. (ORI) at GSFC processed the two scenes into a geometrically rectified and rescaled image, the RPC staff underwent a training period of both the methods of signature identification and the computer techniques used to assist in classification using a copy of the original August scene.

Initial training consisted of becoming familiar with the RJE terminal commands and subsetting portions of the August scene. Principally, as a first cut, the scene training consisted of mapping areas of relatively uniform reflectance in all MSS bands (UMAP) and identifying the land cover of those areas through the use of aerial photographs and other ground truth. Once locations and descriptions were known, statistical descriptions of the spectral response of these areas, usually known as "signatures" were determined. The signatures are actually the average intensities of light as detected within the four bands or wave lengths (see Short, et al, 1976) for areas training sites, defined by the user to be similar. For example, various separate areas of tree cover would actually have four associated intensities within a certain range of values, the signature also included a description of the range of expected values. This is variously known as the "limit" or "standard deviation" of the signature. The test area was then mapped to show the overall occurrence of the identified signatures and checks were undertaken through field surveys and aerial photo investigations.

The early classified maps consisted primarily of water and trees (as those were the most readily and easily identified cover types) and unclassified areas. The unclassified areas were reexamined and new signatures developed through the use of more training sites and a statistical cluster analysis program (CLUS) until the blank, unclassified areas were only a small percentage of the test areas.

As signature identification continued, field checks continued and a detailed personal knowledge of the test area was developed. The list of signatures had grown and had been refined to about 20 district cover types and water types for the August scene.

When the rectified data was available, the processing steps were essentially repeated within the test area. The sample areas used in determing the single scene signatures were used to determine a set of temporal signatures based on both the August and November scenes.

Additional information about the historic land cover of 1975 was sought at this point. The county office of the state forester was consulted on the species and extent of the tree cover in the test area. The extension office of the county soil conservationist was consulted on the location and rotation practices of the agricultural land within the test area. Staff members of CBCES were consulted on the location of open lands and stage of regrowth being experienced in that portion of the test area. Several sources of data were examined for consistency with the classified LANDSAT mapping.

Further field checks and reevaluation of the signatures continued until the data sources had apparent consistency and congruency. There were now 43 signatures which classified most of the area. No measured comparison was available at this time.

C. REGIONAL EXTENSIONS

Using the signatures set developed in the Anne Arundel County test area, sample portions of the other four counties and Baltimore City were classified using the rectified temporal data. These were distributed to the localities for evaluation and comment. Two district problems were noted in this first regional extension.

First, there were additional unclassified areas representing land covers or signatures not encountered in the test area (i.e., different colored base soils, eutrophic water bodies, etc.). These areas were classified by the addition of new signatures developed for new training sites.

Second, several signatures had been assigned to unique land covers in the test area, which misclassified areas in the extension areas. In these cases, the misclassified areas were reevaluated from air photos and field checks, and new signatures developed to represent the areas. With the newly enlarged signature set, the test area in Anne Arundel County was reclassified to determine if the changes made elsewhere would affect the previous accuracies in the test area. As a result, there appeared to be several signatures in which the statistical descriptions were similar enough to confuse the classification and intermix the apparent land covers. This problem was overcome by comparing the signatures of the conflecting classifications with signatures of classifications which were apparently correctly identifying land covers. The averages representing each channel were adjusted slightly toward those of the "correct" signatures and the limit of each signature was reduced in stages until the classified map became more homogenous, both in the test areas and the extension areas. This represented a series of approximately 12 repetitive steps which may have a simpler replacement method, but the misclassification was eliminated and the signature set appeared to correctly classify both the test area and the extension areas, at least in location and extent of land cover types.

D. SIGNATURE CONSOLIDATION

Although the questions of level of detail and reliability had yet to be compared, preliminary results confirmed that without considerably greater effort, landuse breakdowns within the urbanized areas would be limited to two residential densities; tree cover; asphalt; concrete; grass; vacant; water; and building cover. Outside of the urbanized areas, the preliminary results indicated that there would be information on tree cover by coniferous and deciduous; corn fields; other agricultural field types; hay fields and pastures; scrub brush; sand and gravel pits; two more residential densities; and disturbed or bare ground. The signature set still did not adequately define all of the land covers that were required by Section 208. Many of the agricultural identifications were not specific enough. The classification did differentiate between pasture and hay and cropland, but there were several signatures which identified row crops, which while correct to that extent, were not able to satisfactorily separate the land covers into specific crops or crop practices which were identified as important aspects to identify sources of agricultural pesticides, fertilizers, and other runoff-related pollutants.

Working with the Baltimore County Soil District conservationist, the county Soil Conservation Extension office, and several members of the State Soil Conservation Service, a section of Baltimore County was chosen to check and expand the agricultural classifications. Meetings and extensive field checks, including interviews with local farmers, resulted in a base map of the best estimate of the 1975 agricultural coverage of this new section. During field checks and interviews the crops grown in the summer and fall of 1975 were indicated upon an aerial photo (at 1:24,000 scale) which was used as a base map. The crop rotation in effect was also noted.

Using these data a visual comparison was made between the LANDSAT classification and the field checks of the same area. Several of the previously identified "row crops" consistently appeared as individual crops. In fact, with only one change in a signature, the "row crops" were identified as distinctly different crop types or practices. Further by comparing graphs of the signatures, the rotation pattern conformed with the data almost perfectly. For example, several signatures identified as corn through the field checks appeared different in the Fall scene, while similar in the Summer scene. The difference was found to be due to the Winter cropping practices of the individual farmers. Some had planted a Winter cover, either barley or wheat over the corn before November 19, the time of the Fall scene. Others had plowed the corn into the ground after the harvest. While in others, the corn had been removed for silage, leaving only a bare stubble on the land. Each case appeared different and each would potentially affect the type of runoff related pollutants.

The final problem was the extent of the crops. In nearly every identified field, LANDSAT identified or located the ocurrence of the crop type, but it did not accurately indicate the extent. This could be seen in simple visual comparisons of the base map and the LANDSAT map. Upon review of the signatures, there appeared to be a reason. All of the crop signatures were very similar, but had different limits. The variable limits were causing over and under classification of the various crops. Again by a very rudimentary method, the signatures whose limits appeared to be too large were reduced, and those whose limits appeared to be causing under-classification were increased. This was done progressively through nearly 30 iterations until the extent of the LANDSAT classification and the base map appeared consistent. Now consistent for location and extent, and as congruent as the gridded data could be made to the actual field shapes, the Baltimore County study was concluded.

The revised signature set was then rerun on the Anne Arundel County test area. The new classification matched the previous one with only isolated reclassifications, and was considered satisfactory. One additional problem arose in the description of some of the crop types, particularly the green leafy crops. Row crops that had been defined as soybeans in Baltimore County appeared as tobacco and stringbeans in Anne Arundel County. Reviewing signatures derived from both test areas led to the conclusion that these crops could not be adequately separated with the chosen scenes. Also, the relation and response of the crops in runoff related pollutants were not dissimilar. So further separation was judged unnecessary.

IV. EVALUATION OF LAND COVER INFORMATION FROM THE CLASSIFIED IMAGES

A. SECTION 208 REQUIREMENTS

Collection and analysis of the stream water quality sampling data represented approximately 25% of the total Section 208 Grant expenditures. Even at this level of expenditure, the on-site sampling was not extensive enough to cover the entire region with the confidence necessary to implement and enforce an effective pollution control plan. An alternate to full regional sampling had been assumed from the onset of the Section 208 work.

The 208 Work Program had assumed several correlation methods would be developed or applied using land cover information to extrapolate the stream-sampling results from the areas of intensive sampling to the balance of the region. In fact, it had been assumed that these methods would have produced results sufficiently valid for generalizing the location of water quality problems throughout the region and for enforcing control measures to control these problems.

It has been emphasized that classification and identification of land cover proceeded through most of the progress of this work without precise measurements of the cover type associated in the study. The degree of accuracy required for the eventual use of the classifications in this further Section 208 work was determined to be greater than 90 percent.

B. LANDSAT VERTIFICATION

The methods used by the RPC staff represented the best judgement of the staff during the project. Subsequent review of the steps involved has shown some redundancy and some unnecessary efforts. However, most of this unnecessary effort consisted of expanding the verification and accuracy comparisons of the LANDSAT classifications to include may local and state governmental staffs. This was, again, to further the understanding of these staffs, but principally to gather first-hand observers into the process. A faster, more efficient method may have excluded these potential users, and may have reduced the effective usefulness of the data for these users.

The LANDSAT data were statistically processed through a first-hand knowledge of the area to be classified based on a variety of land cover types. This knowledge consisted of walking the site (for smaller areas), windshield surveys, air photo interpretation, and local personal knowledge of residents or experts.

All of these methods were used in obtaining ground truth for verification of the LANDSAT classification. The images were classified into various land cover types and categories which appeared consistent and relatively accurate. Relatively accurate in that no actual measured data on a large scale was available for comparison until nearly the end of the classification process.

C. INITIAL MEASURED RESULTS

Upon completing a detailed land cover information inventory for roughly 2,000 acres in the Rhode River watershed in southern Anne Arundel County, the CBCES of the Smithsonian Institute made a detailed comparison of its inventory versus a LANDSAT classification of the same area completed prior to the agricultural work in Baltimore. This comparison showed an initial weighted category difference of 31% (actual total difference between CBCES "ground truth" and LANDSAT for each category divided by the total CBCES "ground truth" area) and an overall weighted difference of 17%. The comparison was made on summaries of nine small watersheds, ranging from 15 to 625 acres each.

D. TYPES OF POTENTIAL PROBLEMS OR ERRORS

Major problems resulting in such a difference or error have been discussed by Alexander et al (1975) and are listed here only to format the following discussion:[3]

1. The mixture of different land use categories within a small area, which is the minimum-size mapping unit...resolution difficulties.

2. The generalization of land surface types into units covering larger areas, as in lower-resolution sensors such as LANDSAT...problems with heterogeneous study areas.

3. Errors due to imperfect registration of boundaries between categories on the maps being compared...boundary location problems.

4. Errors due to generalization from larger map scales to smaller map scales ...boundary edge problems.

5. Errors due to differences in interpreter applications of the classification system... definitional differences.

6. Errors due to interpreter misclassification...verification data error.

7. Errors due to change between the times of the gathering of the two data sets ...actual change.

Discussions with the CBCES staff about their interpretation methods, field checking and scale of data negated the possibility of many potential problems. Their land cover had been collected on 50' scale maps through extensive field checking, principally on foot, throughout a period of several years, using an extremely detailed classification scheme. Types 6 and 7 error were virtually eliminated by the nature of their data.

Types 3 and 4 error were difficult to access because in overlaying base maps on the LANDSAT image at 1:24,000, a one-tenth of an inch uncertainty in location resulted in the inclusion or exclusion of a strip of LANDSAT pixels from the comparison summary. This problem was ignored initially and will be subsequently discussed. However, the measured acreage and LANDSAT pixel acreage were significantly different. In some of the sub-watersheds, the difference in area was as great as the weighted category difference. Rechecking of the handclassified data yielded several summarization mistakes, a series of clerical errors, and some statistical errors, causing the weighted category difference to decrease to 30% while the overall weighted difference decreased to 16%.

E. FINAL SIGNATURE CALIBRATION

The first most obvious source of error was in the LANDSAT classification itself. In fact, the version of the LANDSAT classification used in the initial CBCES comparison was a version completed prior to the agricultural reclassification. The subsequent work in Baltimore County had revised several signatures and numerous signature limits and should have been the basis of the comparison to reduce any residual error.

Second, knowing the heterogeneous nature of the land cover in the Baltimore region, the Types 1 (resolution difficulties) and 2 error (heterogeneous study area) were anticipated. Each signature as it was developed was given a descriptive narrative of what was actually observed in the training sites. This early description included the approximate average percentage of water, tree cover, grasses/brush, grass/pasture, corn, alfalfa/hay, other crops, urban build-surface, and bare soil, with the percentages based on evaluation of the aerial photographs and other ground truth in the training sites. Table 1 is the final matrix of these observed percentages of the signatures developed for the region.

When CBCES staff had aggregated the LANDSAT acreage, they had done so by grouping the signatures by the signature names. The early matrix of signature cover types should have provided a finer aggregation with less Type 5 error (definitation differences). A new comparison was made using the most recent signatures and the cover matrix to develop the aggregate acreages. This comparison showed that the weighted category difference decreased to 18% and the overall weighted difference decreased to 8%.

Using an iterative process (22 stages) of increasing or decreasing the percentages in the cover matrix as necessary to reduce the individual watershed category differences, a revised matrix (Table 1) was developed. This approach was the best available to reduce the Type 5 error. The difference eventually stabilized during the iteration until no further improvement was evident. The weighted category difference decreased eventually to 16% and the overall weighted difference decreased to 7%.

The error that remained appeared to be either Types 3 and 4 (boundary location and edge problems) which had been ignored, or some unexplainable residual. The LANDSAT image was shifted to see if improvement, indicating better match, would result. A one cell shift to the west, north, and south resulted in an average increase in the difference, while a one cell shift to the east reduced the difference to almost half the previous comparison. This eastward shift was incorporated in the final comparison, where the weighted category difference decreased to 11% and the overall weighted difference decreased to 5%.

F. REGIONAL FOLLOW-UP

Following the verification work done in Anne Arundel County with the CBCES inventory, four individual test sites of between 800 and 1,000 acres each were chosen in the other counties in the region. The County Soil Conservationist and County Extension agents were contacted and requested to provide detailed land cover acreage for verification of the satellite data. These data were measured and summarized by county staffs under the direction of the RPC 208 staff. The apparent difference of initial comparison of these data for the four sites varied from 46% to 57%.

The RPC staff reviewed the ground truth and LANDSAT data again to determine the probable cause of the difference. The difference in the comparisons was shown to have been a Type 5 (definitional differences) or a Type 7 (actual change) error.

In the case of the Howard County site, large areas of brush were designated as trees, pasture as brush, and idle agricultural land as fallow land and successional fields. The Type 5 error was so gross that total reinterpretation was necessary.

For the Carroll County site, the interpretation excluded nearly 200 acres, a third of the site, from the summary, because a landowner refused to provide data. The comparison was still made by county staff without excluding this portion of the site. Additional interpretation of the exluded portion would have been necessary to make an accurate comparison, but access to prior years' information remained unavailable.

In the Harford County site, a large partially developed residential subdivision was interpreted as wholly residential. In fact, the subdivision was sparsely developed, including large tracts of trees, brush, old fields, and some remaining cropland. The developed portion of the subdivision was also in large lot parcels, 2 acres/house, which meant that each building site was actually larger than the LANDSAT data cell size. The allowed undeveloped portions of even the developed lots, grass and trees, to be classified as pasture or trees by LANDSAT. Table 2 shows the step by step improvement in the average weighted category error for the Harford County site. Column I is the initial data comparison. Column II is correcting for a Type 5 error in the residential interpretation. Column III is correcting for a Type 7 error in the similar categorization of hay as pasture in both data sources, not leaving hay as a crop in the hand interpretation and as pasture in the LANDSAT.

As can be seen from Table 2, the initial review between the county land cover information and the LANDSAT classification resulted in a 51% difference. Inquiry into the nature of the differences revealed problems with the county provided information and resulted in a better understanding of temporal classification.

1. The found truth land cover information provided by the Soil Conservationist and County Extension agents included as residential an entire residential

TABLE 1 MATRIX OF LANDCOVER PERCENTAGES ASSOCIATED WITH LANDSAT CLASSIFICATION CATEGORIES - SUMMARY OF GROUNDTRUTH DATA FOR VERIFICATION SITES (DATA SOURCE: RPC)

SIGNATURE... KEY:	DESCRIPTION:	WATER	TREE CROWN	GRASS/BRUSH	GRASS/PASTURE	CORN	HAY/ALFAFA	OTHER CROPS	URBAN	BARE SOIL
WATER TYPE.....										
W	DEEP WATER	100								
T	SHALLOW WATER	100								
S	WATER EDGE	50	35	10	5					
%	LAND EDGE	30	50	20						
$	SILTY WATER	100								
N	TIDAL GRASS	100 (1)			100 (1)					
Z	TIDAL FLAT	100 (1)								100 (1)
TREE TYPE.....										
G	GUM/POPLAR		80.5	3.5	7					
O	WHITE OAK		100							
M	MAPLE/GUM OAK		95	1.5	3.5					
P	CONIFERS		75	20	5					
Q	CONIFERS		75	20	5					
R	CONIFERS		75	20	5					
H	SYCAMORE/ASH		90.5	4	5.5					
B	PINE PLANTATION		85	15						
OPEN LAND.....										
9	GRASS/BRUSH			100						
-	GRASS/PASTURE				87	6				7
6	GRASSES		15		85					
L	GRASS				100					
E	GRASS/PASTURE				100 (2)					100 (2)
>	GRASS/BRUSH			63	27					
/	GRASS/BRUSH			100						
1	GRASSES		10		90					
CROP LAND....										
C	CORN				10	80				10
7	CORN/BARLEY				11	84				5
2	HAY/ALFAFA				5		95			
3	ROW CROPS		10					90		
4	ROW CROPS				50			50		
5	ROW CROPS				5			95		
J	ROW CROPS							100		
K	ROW CROPS		10					90		
F	ROW CROPS							100		
URBAN LAND.....										
A	VL.D. URBAN		65		10				25	
D	VL.D. URBAN		65		10				25	
#	L.D. URBAN		45		5				45	5
<	L.D. URBAN		15	20	20			10	45	
"	M.D. URBAN		5		30				65	
=	M.D. URBAN				30				70	
X	MD.D. URBAN		15	10	5				70	
V	H.D. URBAN		5	10	5				80	
U	H.D. URBAN		5	5	5				85	
Y	VH.D. URBAN								90	10
BARE SOIL.....										
@	BARE SOIL		10	10	5					75
+	BARE SOIL			15						85
•	BARE SOIL			5						95
8	GRAVEL QUARRY									100

NOTES:
(1) TEMPORAL SCENES INCLUDED TWO TIDAL STAGES
(2) DAIRY FARM PASTURE SEVERELY OVERGRAZED IN NOVEMBER SCENE

TABLE 2 HARFORD COUNTY VERIFICATION SITE GROUNDTRUTH ACREAGE COMPARISONS (DATA SOURCE: RPC)

LANDCOVER CATEGORY		INITIAL COMPARISON (I)	RESIDENTIAL CORRECTED (II)	HAY/PASTURE CORRECTED (III)
TREES	(1)	176.3	258.4	258.4
	(2)	240.6	240.6	240.6
	(3)	.36	.07	.07
OLD FIELDS	(1)	40.0	153.9	153.9
	(2)	154.0	153.9	153.9
	(3)	2.85	.00	.00
CROPS	(1)	250.7	250.7	176.0
	(2)	172.9	172.9	172.9
	(3)	(.31)	(.31)	.02
PASTURE	(1)	259.0	259.0	333.7
	(2)	295.2	295.2	323.6
	(3)	.14	.14	.03
RESIDENTIAL	(1)	219.9	23.9	23.9
	(2)	28.5	28.5	28.5
	(3)	(.87)	.19	.19
OTHER (5)	(1)	-	-	
	(2)	28.4	28.4	
	(3)	-	-	
TOTAL	(1)	945.9	945.9	945.9
	(2)	919.5	919.5	919.5
	(3)	(.03)	(.03)	(.03)
	(4)	.51	.14	.04

NOTES:

(1) HARFORD CO. SCS GROUNDTRUTH - RECHECKED: (A)
(2) ORSER/RPC LANDSAT - JULY 1977 VERSION 2: (B)
(3) PROPORTION DIFFERENCE: DIFF = (A-B)/A
(4) WEIGHTED CATEGORY DIFFERENCE:

$$DIFF = \frac{SUM_K (A_K - B_K)}{TOTAL_A}$$

WHERE K = CATEGORY

(5) BARE SOIL

subdivision which included woodland, brush and grass, and undeveloped building lots. The LANDSAT classification differentiated the area into the land covers mentioned above.

It was thus necessary to reinterpret the "residential" ground truth data. A more detailed study revealed two main types of development within the subdivision: open residential lots and wooded residential lots, with minimal lawns. Open residential lots totaled 90 acres with 8.2 acres being actually houses (built-surface), while the remaining 81.8 acres were a combination of grass and brush which would have been considered a "successional" field if the house had not been present.

The wooded residential area totaled 62.9 acres, four acres of which were in houses, with the remaining 58.9 acres being trees.

The remaining residential areas were reevaluated based on the information. Column I of Table 2 was revised to include these changes causing the weighted percentage difference to decrease to 14.0% (Table 2, Column II).

2. Further adjustment was made when it was discovered that 74.7 acres of hay had been defined as an agricultural field. For the purposes of the LANDSAT investigations, "hay" had always been included with pasture land/low grasslands. To be consistent, the "hay" acreage was transferred to the pasture category (Table 2, Column III).

3. Finally, it was determined that the LANDSAT category of 28.4 acres of "other" was a bare soil signature. The signature was one similar to other grass signatures in the summer, but more similar to bare soil than grass in the late fall. Reviewing the ground truth information again revealed a period in which pasture was, in fact, frequently overgrazed in the fall. The "other" category was transferred to pasture. Table 2, Column III shows the revised figures and now a weighted difference of only 4.0%.

G. CAVEAT ON THE MATRIX APPROACH

The verification described in this report was conducted on an area basis. That is, within known boundaries certain land cover relationships, either acreages or proportions of acres were known. The results of the LANDSAT classification were compared with these area summaries, and the accuracies relate to those summaries. The nature of the heterogeneous signature results in distinct comparison problems when specific locations are compared with known data.

Previous reference to the method of developing signatures from the training sites indicated that the training sites were not completely homogeneous. An example is the signature for "grasses" represented by symbol "6" on Table 1. The training area was principally grass and pasture (85%), but a significant portion (15%) was isolated trees and hedgerows. Thus, the matrix indicates that mix of cover. But to be able to investigate an individual pixel classified as this signature and expect to find this particular mix is unlikely. More probably, 15% of the pixels classified as this signature would be trees and the balance of the pixels, grass. This was the nature of the training site and should be the expected nature of the classification.

Because a "one-to-one" classification was not required for the data, the chosen approach was not developed further to give such a classification. And again, the actual nature of the land cover may not have allowed classification into even 60 separate classes when such a broad multitude of different cover combinations existed on a pixel scale.

V. UTILIZATION OF THE CLASSIFIED IMAGES

The reliabilities that were achieved were considered sufficient to allow utilization of the classified data for the RPC Water Quality Program and other potential RPC activities.

The data has been formated into a computerized Regional Land Classification Atlas which can be mapped in section or in whole and which can be utilized in the RPC Polygon-Overlay system. The Polygon-Overlay system allows the data to be incorporated in other gridded data bases and to be summarized by the various boundary and data variable conditions which have been encoded into the system.

Table 3 is an example of summarization for the Gunpowder River Basin within Baltimore County. The classified image is summarized by each of the 45 Section 208 Water Quality sampling segments within the basin and aggregated by subbasins. A summary of this scale, approximately a 312,000 acres summary, allowed comparisons that previous hand summaries did not allow. Table 3 is also an initial comparison of the ORSER classified image and the results of a study done by General Electric using the G.E. Image 100 for the Department of Public Works, Bureau of Operations, of the City of Baltimore.

These comparisons are well within the previous range of difference that was encountered prior to an analysis of the source of the difference. Considering the difference of seasons and the two years occurring between the scenes, the comparison still shows major similarity.

VI. RECOMMENDATIONS

In addition to the present and future uses in the Section 208 work, the presently classified image has opened a broad category of activity to the RPC:

1. Detection of inconsistencies in existing data bases, particularly the interpreted densities of the developed land inventory.

TABLE 3 COMPARISON OF IMAGE CLASSIFICATION BETWEEN ORSER/RPC AND GE IMAGE-100 FOR THE LOCH RAVEN RESERVIOR WATERSHED -- GUNPOWDER FALLS BASIN -- (DATA SOURCE: BALTO.CITY DPW, RPC)

ORSER/RPC (1) ACRES	ORSER/RPC (1) %	COMMON LANDCOVER CATEGORY	IMAGE-100 (2) ACRES	IMAGE-100 (2) %	PERCENT DIFFERENCE
(A)			(B)		(A-B)/A
3071	2.2	WATER	2159	1.5	29.7
49221 (3)	34.8	TREES MIXED FOREST	12464	8.9	(28.5)
		DECIDIOUS FOREST	48417	34.6	
		CONIFEROUS FOREST	2390	1.7	
15735	11.1	BRUSH/OPENLANDS	50939	36.3	(3.6)
33455	21.0	CROPS/BARE FIELDS	15382	11.0	55.0
4501	3.2	BARE SOIL			
5638	4.0	RESIDENTIAL	5411	3.9	4.0
		URBAN	2164	1.6	
240	0.2	EXTRACTIVE	217	0.2	-
225	0.2	UNDEFINED	259	0.2	-
141522		TOTAL	139871		1.2

NOTES:

(1) AUGUST AND NOVEMBER 1975 TEMPORAL IMAGE
(2) JULY 1973 SINGLE IMAGE
(3) DETAILED ORSER TREE CLASSIFICATION HAS BEEN COMBINED IN ONE CATEGORY BY POLYGON/OVERLAY ROUTINES

2. Natural resource evaluations, particularly the quantification of forest lands, croplands, and open lands.

3. Wildlife association studies dependent on undeveloped land inventories.

4. Vegetation species identification.

5. Identification of sources of windblown fugitive dust, (particulate matter) from bare soil and urban areas for 1979 Maryland State Air Quality Implementation Plan.

6. Application of runoff coefficients relating to land cover for 1977-79 Soil Conservation Service Patapsco Basin Flooding Study.

Future image classifications have been discussed as a method of urban growth change detection to supplement current permitting procedures.

BIBLIOGRAPHY

1. Regional Planning Council Technical Memorandum Number 17 and 17A, 1973 Landuse and Zoning Patterns in the Baltimore Region, April, 1975.

2. Regional Planning Council Technical Memorandum Number 19, Land Development Changes, January, 1976.

3. Land Use and Environmental Assessment in the Mid-Atlantic Region, June, 1975, Robert H. Alexander, Katherine Fitzpatrick, Harry P. Lins, and Herbert K. McGinty III, NASA Earth Resources Survey Symposium, Lyndon B. Johnson Space Center, Houston, Texas.

4. Short, Nicholas M., Paul D. Lowman, Jr., Stanley C. Freden, and William A. Finch, Jr., 1976, Mission to Earth: LANDSAT Views the World, Scientific and Technical Information Office, National Aeronautics and Space Administration, Washington, D.C.

Henry F. Fostel:

B.Architecture, Carnegie Mellen University

M.S. in Environmental Science, Johns Hopkins University

M.L.A.U.D., Harvard University

Currently...Environmental Planning Consultant, Baltimore Regional Planning Council

James E. Manley:

B.A. in Geography, University of Maryland

Graduate Study in Geography, University of Maryland

Currently...Information Systems Consultant, OMNI Information Services

James P. Ormsby:

B.S. in Education, SUNY at Fredonia

M.S. in Earth Science, Cornell University

PhD. in Meteorology, Cornell University

Currently...Research Hydrologist, Hydrospheric Science Branch, Code 913, Lab for Atmospheric Science, NASA/GFSC