(ECHOR)

E/S No. G-093

NASA CR-147457

E7.6-10216

APPLICATION OF SKYLAB EREP DATA FOR LAND USE MANAGEMENT

(E76-10216) APPLICATION OF SKYLAB EREP DATA FOR LAND USE MANAGEMENT Final Report (Earth Satellite Corp.) 306 p HC $9.75 CSCL 08B
N76-19519
Unclas
G3/43 00216

FINAL REPORT

Contract No. NAS9-13314

NASA, LYNDON B. JOHNSON SPACE CENTER
EARTH OBSERVATION DIVISION
HOUSTON, TEXAS 77058

JANUARY 1976

EARTH SATELLITE CORPORATION (EarthSat)

7222 47th St. (Chevy Chase), Washington, D.C. 20015

2150 Shattuck Avenue, Berkeley, California 94704

(301) 652-7130 (415) 845-5140

PREFACE

This report was prepared as the final report for NASA contract number NAS9-13314, EREP investigation number 520 entitled "SAMPLING STRATEGIES IN LAND USE MAPPING USING SKYLAB DATA". The investigation was conducted under the technical direction of the principal investigator Dr. David S. Simonett. Major technical contributions have been made by Robert Shotwell and Nathaniel Belknap with supporting contributions by Charles Drackett, Timothy Gregg, William Meyers, and Geoffrey Coleman. Early phases of the work were conducted by Wayne Rohde, Jack Bale, and Darryl Goehring. The illustrative materials have been prepared by Edmund Schantz.

The authors wish to express their thanks to the contract Technical Monitor Rigdon Joosten for his assistance throughout the entire contract period.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

As the effect of man on the environment in which he lives continues to increase, the importance of land and natural resource management should rise at an equal rate. Effective land planning requires an awareness of all human and natural resources that may be impacted by various management alternatives. Any land resource management decisions must be preceded by the collection of adequate and accurate information on which to base the decision process. Remote sensing offers powerful information gathering capabilities and can provide the accurate and detailed data needed by the planner in a timely and cost effective manner.

It is not the intent of this report to review the early history of remote sensing as a source of land use information or to describe the many applications in traditional urban planning programs. Works by Branch (1971), Westerlund (1972), and Estes and Senger (1974), all with numerous references, will provide the reader an excellent overview of remote sensing as it has been applied to land planning and related environmental analysis. These volumes indicate that aerial photography has significant potential for meeting data collection needs of land planners and managers.

Recently, aerial photography from high altitude aircraft has become available to planners. Such data have been found useful for detailed land use mappings over large areas. Vegas (1974) presents a methodology for the use of high altitude photography in land use classification. Similar techniques have been employed to map land use over the entire State of Maryland (Brooner and Wolf, 1974). Many other states and counties have completed land use surveys from high altitude photography. In recent years, state, county, and other regional planners have been increasingly faced with the need for region-wide land use and related data to update existing information, develop land use plans, and to monitor outcomes of the planning process. As new techniques are acquired, needed data are developed, tested, and become operationally available. Planners can adopt them as a means of meeting a part of their information needs. In addition to the use of high altitude aircraft imagery, the new imagery from the Earth Resources Technology Satellites, ERTS-1 and 2 (now called LANDSAT-I and II) and SKYLAB offer a potential source of data useful to the planner.

LANDSAT-I and II provide data from which changing land cover patterns over large regions can be rapidly mapped and monitored. The SKYLAB Orbital Manned Workshop launched in early 1973 provided color, color infrared, and multi-banded imagery over much of the United States which may be used for similar purposes.

A. GENERAL OVERVIEW OF INVESTIGATION

This investigation was initially designed to use both SKYLAB/EREP-A, S-192 multispectral scanner data, and S-190 multispectral photography acquired over the Washington, D.C.-Baltimore area and the Alice Springs, Central Australia area to examine those parameters influencing sampling strategy with respect to manual, semi-automated, or automated thematic land use mapping with spacecraft data. The intention was to merge multidate, 13-channel imagery and assess the various trade-offs between multiband and multidate concepts in improving land use identification to level III, and possibly even level IV, within a sampling format.

The work discussed in this final report is confined to the SL/3 Mission over the Washington, D.C.-Baltimore, Maryland urban areas and nearby rural areas in Virginia and Maryland, and the SL/4 Mission over the Alice Springs site.

As originally planned, the U.S. test site was to be covered twice and was to encompass an area some 159 miles long and 42 miles wide, extending northwest to southeast over the Virginia-Maryland area, including the Washington, D.C.-Baltimore, Maryland corridor.

Orbital changes early in the SL/2 Mission, however, provided imagery shifted to the southwest of the original track. When the SL/3 Mission reported on here was flown, the original SKYLAB track was used, leaving very few of the original SL/2 sample sites covered twice. This substantially negated the original study. A combination of local clouds and high cirrus within the twice-seen area further complicated the picture.

Similar problems occurred in the Central Australian site. The area was virtually totally cloud-covered during the SL/3 Mission. Even the SL/4 Mission was handicapped by clouds in portions of the test site. In addition, the S-192 multispectral scanner performance over the Australian site was such that after inspection of the data tapes it was decided to confine analysis to the S-190 A and S-190 B color and color infrared photography. Comparisons with previous analyses were carried out with Gemini photography incorporating field observations and low altitude aircraft flights provided by cooperating Australian scientists.

Of the original objectives therefore, those questions dealing with multidate imagery were set aside because an acceptable body of observations was not available. In addition, those questions dealing with sample design for land use ground truth sampling and geographic distance-decay function were also set aside: 1) because these questions, by the time of receipt of the SKYLAB digital tapes, had already been answered sufficiently in substance by several LANDSAT 1 investigations; and 2) the data set available with the S-192 tapes could not constitute an appropriate test of the hypotheses involved, nor could significant additional conclusions to those of the LANDSAT studies be expected.

In essence, as pointed out in the original proposal, ground truth sampling must be both randomized and systematic to avoid bias. LANDSAT investigations in which this has been observed include that by Von Steen and Wigton (1973). Similarly, geographic distance decay functions have been shown to be different for different environments, and to be dependent on the degree of spatial stratification of homogeneous land use regions and sub-regions employed in both sampling and prediction. Studies at the University of California, Berkeley (Nichols et al., 1973), and by Morain (1973) at the University of Kansas are relevant.

B. <u>WHAT IS TO BE LEARNED FROM THE INVESTIGATION</u>

The questions which were addressed are as follows:

<u>S-192: U.S. Site, Washington-Baltimore Region</u>

1. To what degree does the increased spectral coverage of the SKYLAB S-192 Multispectral Scanner (0.41 microns-12.5 microns) provide new and important information not currently available with the LANDSAT I and II MSS?
2. What are the optimal spectral bands to be used in discriminating various land use categories? How do the optimal spectral bands vary from category to category?

3. What accuracy levels are achievable when performing multi-spectral classification using SKYLAB S-192 data? Do these accuracy levels meet the needs of the land use planner?

4. How do the classes of data which are spectrally discriminable by the S-192 Multispectral Scanner relate to the land use categories of value to the land use planner?

5. What impact does the SKYLAB study have on future space platforms with regard to: 1) Spectral band (number, widths, and locations), 2) Sensitivity, 3) Spatial resolution, 4) Temporal coverage.

S-190 B: Alice Springs, Australia

6. What gains in grazing land category identification are there in comparison with either the S-190 A or earlier Gemini photography?

7. What improvements are there, and how significant are they, with respect to landscape boundary delineation as a result of the improved spatial resolution of the S-190 B?

8. How do these improvements compare with predictions made in 1969 on the value of resolutions of 50 to 100 feet for pasture and range category separation and delineation in Central Australia?

S-190 A and S-190 B: U.S. Site

Additional questions dealing with the value of S-190 A and S-190 B photography for land use mapping in comparison to aircraft high altitude color-infrared photography were also studied for the Washington-Baltimore region. These questions were addressed fully in an earlier report by Bale, Rohde, Goehring, and Simonett (1975) and will not be examined in detail here, although both the questions and the results are summarized in this final report.

The questions were:

9. Does the trend towards planning for larger areas with a regional perspective, now evident in the planning community, make satellite-derived imagery suitable for a planning base?

10. How many of the land uses mappable with high altitude aircraft imagery are derivable from SKYLAB S-190 B imagery? Can the S-190 B imagery substitute for the aircraft data?

The emphasis of this investigation was therefore twofold: first, to analyze digitally the S-192 Multispectral Scanner data and determine its utility in the context of land use planning and resource management; second, to investigate the utility of S-190 multispectral photography, again as related to the land use planning community.

C. SUMMARY OF MOST SIGNIFICANT RESULTS

S-192: U.S. Site, Washington-Baltimore Region

1. To what degree does the increased spectral coverage of the SKYLAB S-192 Multispectral Scanner (0.41 microns - 12.5 microns) provide new and important information not currently available with the LANDSAT I and II MSS?

    - Of the six spectral bands identified as providing the best spectral discrimination between land use categories (spectral band numbers 1, 3, 6, 9, 11, and 13), only two bands fell within the LANDSAT I and II spectral coverage region (0.5 microns - 1.1 microns).

- The best single spectral band identified for discriminating between land use categories (spectral band 9, 1.09 microns -1.19 microns) was slightly above the LANDSAT I and II spectral coverage region.

2. What are the optimal spectral bands to be used in discriminating various land use categories? How do the optimal spectral bands vary from category to category?

- Different sets of spectral bands were selected when analyzing different groups of level II land use categories. When analyzing groups of level II land use categories all belonging to the same level I land use category, the following sets of spectral bands were selected as providing the best discrimination:

  Urban - spectral band numbers 9, 3, 11, 6, 13
  Agricultural - spectral band numbers 9, 7, 1, 11, 6
  Forest - spectral band numbers 9, 3, 5, 4, 11
  Water - spectral band numbers 6, 1, 3, 8, 9
  Wetlands - spectral band numbers 10, 3, 1, 8, 6

- In an overall rating including all spectral bands and all land use categories, the optimal spectral bands identified included bands evenly distributed over all spectral regions:

  Visible - spectral bands 1, 3, 6
  Near infrared - spectral band 9
  Mid infrared - spectral band 11
  Thermal infrared - spectral band 13

- The optimal spectral bands identified for discriminating between general land use categories were significantly different from the spectral bands identified for discriminating within specific land use categories.

3. What accuracy levels are achievable when performing multispectral classification using SKYLAB S-192 data? Do these accuracy levels meet the needs of the land use planner?
   - Overall accuracy in classifying general level I land use categories was found to be about 70%.
   - Level II land use category classification accuracy varied significantly from category to category. In general, however, the land use level II accuracies were unacceptable for land use planning purposes.
   - Careful stratification of the data between training classes and test classes produced comparable accuracies for both sets of classes.
   - Traditional definitions of land use categories frequently do not produce spectrally separable classes of data. To obtain maximum benefits from remotely sensed data it may be necessary to establish land use category definitions on the basis of spectrally discriminable classes.
4. How do the classes of data which are spectrally discriminable by the S-192 Multispectral Scanner relate to the land use categories of value to the land use planner? It should be noted that this question was assessed through use of a clustering algorithm. Since such algorithms cluster the actual data, fully quantitative analysis would be misleading since the composition within the clusters is influenced by the number of

items in the sample as well as the clustering routine and thresholds employed. Consequently, the clustering is a general guide, not a fully quantitative assessment procedure.

- Acceptably clear groups may be established for water, including a split between estuarine and river water groups.
- High density industrial and commercial sites may be identified as a broadly discriminable group, principally because they are of low vegetation density.
- Transportation and industrial sites of very low vegetation cover are readily separable as a group.
- Land uses with deciduous trees present, including forest land and urban residential, are frequently confused on a one-time basis, but multidate imagery should enable the residential areas to be separated.
- Evergreen forest should also be discriminable with multidate imagery.
- The mixed group comprising urban and agricultural land may also be further separable with multidate imagery.

5. What impact does this study have on future space platforms with regard to: (1) Spectral bands (number, widths, and location) (2) Sensitivity (3) Spatial resolution (4) Temporal coverage.

- Repeatedly, the first four channels selected by the stepwise discriminant analysis came from the thermal IR, mid IR, near IR, and visible spectral regions. That is, three channels from the visible region were not selected as the first three channels by the analysis. This indicates that a wide variation of the spectral channel coverage may provide the greatest discriminability between data classes. This lends strong support to the general LANDSAT D approach of using channels from each of the four areas. Much of the discriminatory power of these channels for separation of land-use categories derives from the type and density of the vegetation present. LANDSAT D is intended primarily, though not exclusively, to meet vegetative sensing needs.

- The 1.09 - 1.19um channels proved to be very important, ranking high in nearly all tests with different land use groups (this channel ranked 1st or 2nd in 4 out of 6 tests). This channel is not at present available on LANDSAT I or II. It should be looked at very carefully indeed as a contender for a berth on LANDSAT D.
- Multidate imagery will be essential to achieve separation of some classes of importance to the state land use planner. Even using all channels at a single date, separation cannot be obtained at an acceptable level. Thus multidate imagery must be employed.
- Finer spatial resolution than in SKYLAB (if coupled with low noise and adequate sensitivity and spectral resolution) will enable additional discrimination by reducing the number of mixed pixels. The LANDSAT D design goal of 30 meters would yield additional improvements in identification.

- Sensor design characteristics of the S-192 Multispectral Scanner are found to have significant detrimental impact on the quality and utility of the scanner data. In particular, it is found that:
  (1) Small band-to-band misregistration is an important factor to be considered in the design of future sensors. Misregistration of as little as 1/2 pixel may seriously degrade classification accuracy.
  (2) Scan line straightening required because of the conical scan pattern introduces additional spatial misregistration.
  (3) Oversampling by the high rate detectors does not significantly improve spatial resolution.
- Sensitivity questions cannot properly be addressed because of the significant noise component in the SKYLAB tapes.

S-190 B: Alice Springs, Australia

6. What gains in grazing land category identification are there in comparison with either the S-190 A, or earlier Gemini photography, through use of the S-190 B system?
   - There are significant improvements in identification of grazing-class categories arising from the improved spatial resolution of the S-190 B system, particularly with respect to detecting differences between plant communities on the basis of internal spatial variability within a class. As examples, the Mulga (Acacia aneuva)

communities are clearly visible through observation of groves (small clumps of trees within the community generally). Melaleuca shrublands fringing small saline depressions to the east of Napperby Dry Lake are also readily observable. Detailed lithologic and related plant community differences in small areas of the Kunoth Paddock are now discriminable. On previous space imagery all three were blurred and not interpretable. These abilities were in fact predicted in 1969 by Simonett et al., and are important in making space photography more useable as part of a sampling frame. The more categories detected in the first stage of a multi-stage sample, the greater the improvement in the precision of an estimate of within and between plant community discrimination.

- The quality of category detection was such as to suggest errors in the initial ground truth mapping used for comparison with the space images.

7. What improvements are there, and how significant are they with respect to landscape boundary delineation as a result of the improved spatial resolution of the S-190 B in comparison with Gemini and LANDSAT imagery?

    - Paralleling the improvement in category identification, there is a sharp reduction in the uncertainties of boundary placement, and a considerable increase in the number of possible additional categories and boundaries at lower hierarchical levels of classification of pasture types, and of smaller units previously not detectable.

- Detection of boundaries of cultural features such as property and paddock boundaries, shown by differential grazing, and small cultural details such as narrow unpaved tracks, stock ponds, and overgrazed areas around wells, is greatly facilitated with the S-190 B. Thus, exact location of sample sites is much improved, reducing the need for a high level of second stage (aircraft) data. Indeed, a two-stage design using only spacecraft and ground observation is feasible.

8. How do these improvements compare with predictions made in 1969 on the value of resolutions of 50 to 100 feet for pasture and range category separation and delineation in Central Australia?

   - The improvements in pasture and range category separation and boundary delineation compare closely with predictions made in 1969, based on calculations of the number of entities present in a resolution cell (see Simonett et al., 1969, and Simonett and Coiner, 1971).
   - Spatial resolution, not additional bands, is the critical factor in such improvement. S-190 A imagery, with infrared response, but with resolution a factor of three poorer than S-190B is not adequate for additional separations.

S-190 A and S-190 B: Washington-Baltimore Site, U.S.

9. Does the trend towards planning for larger areas with a regional perspective, now evident in the planning community, make satellite-derived imagery suitable for a planning base?

- Yes; the shift from planning with an urban emphasis, with its very detailed, large-scale data needs, to regional and state planning will facilitate the use of satellite and spacecraft derived imagery.
- The needs of the planner concerned with large areas, with comprehensive environmental analysis, and with land resources management are roughly compatible with the scales, resolutions, and information content of space-derived images.

10. How many of the land uses mappable with high altitude aircraft imagery are derivable from SKYLAB S-190 A and S-190 B imagery? Can the S-190 B imagery substitute for the aircraft data?
    - Good quality information in map form can be expected from S-190 B imagery at both level I and level II as defined by U.S. Geological Survey Circular 671.
    - Variability in the results for tests of level III and level IV suggest that the spectral coverage of both color and color infrared film are needed for accurate identification and mapping in forest and agricultural classes.
    - The experience of the interpreter markedly influences the quality of interpretation of both S-190 A and S-190 B imagery to a higher degree than is true of high altitude aircraft imagery.
    - The information content of S-190 B is much more suitable for the land use and regional planner than is LANDSAT data in areas of high urban-rural density. In other areas with larger natural plant community groups and with less intensive settlement the LANDSAT data may well be adequate.

- S-190 B cannot acceptably substitute for all aircraft image uses by land use and regional planners. Their responsibilities require the coordination of smaller jurisdictions with much more detailed data needs.
- SKYLAB S-190 A and S-190 B imagery can be used to update existing aircraft-based land use maps at the county level at a substantial cost savings (see Rohde and Simonett, 1975).
- SKYLAB S-190 A and S-190 B imagery can be used to revise, update, and improve the delineation of forest type boundaries at the county level in Maryland (see Rohde and Simonett, 1975).

D. CONCLUSIONS OBTAINED FROM INVESTIGATION

In the preceding summary of the most significant results from the investigation, many of the major conclusions of the study are foreshadowed.

The major conclusions are given below, first with respect to the Multispectral Scanner and the implication for future spacecraft systems, and second with respect to the value of a high resolution photographic system. The conclusions for the multispectral scanner apply to both unmanned and manned systems (LANDSAT D, SHUTTLE), while the photographic systems apply mainly to the manned SPACE SHUTTLE.

Multispectral Scanner

The principal conclusions regarding the multispectral scanner deal with selection of spectral bands most useful for land use planning applications.

1. The 1.09-1.19 um band proved to be very valuable for discriminating a variety of land use categories, including within-category discrimination of agriculture, forest, and urban classes.
2. The thermal infrared channel proved useful for discriminating between urban and vegetated categories.
3. The 1.55-1.75 um band proved very useful in combination with the 1.09-1.19 um band.

Additionally, it was concluded that:

4. Misregistration between spectral bands, even by as little as 1/2 pixel, may degrade classification accuracy.
5. Accuracy of identification of boundary or border pixels was as much as 13% lower than the accuracy for identifying internal field pixels.
6. Multidate imagery may be necessary to accurately discriminate land use categories both at level I and level II.
7. In order of overall ranking, the most useful six spectral bands were found to be:
    1) Spectral band number 9 (1.09-1.19 um)
    2) Spectral band number 3 (0.52-0.56 um)
    3) Spectral band number 6 (0.68-0.76 um)
    4) Spectral band number 1 (0.41-0.46 um)
    5) Spectral band number 11 (1.55-1.75 um)
    6) Spectral band number 13 (10.2-12.5 um)

S-190 B Photographic System

The principal conclusion with respect to the S-190 B camera system is - as would be expected - that the higher resolution of the S-190 B system in comparison to previous space photography (Gemini, Apollo), to the S-190 A system (SKYLAB), and to LANDSAT imagery significantly increases the range of additional discrimination achievable. While evidence is available that the infrared layer of the S-190 A system and the IR bands of LANDSAT enables some identification not feasible (e.g. forest category separation, separation of small water bodies from forest, etc.), with color film in the S-190 B data the high resolution is more generally useful than additional or alternative bands.

There is no reason why future photographic systems could not include color infrared film in association with high resolution. The advantages of high resolution of the order of the S-190 B system are as follows:

1. More categories can be identified with lower ambiguity.
2. Boundary delineation and small-area field delineation is more precise and closely matches that obtainable with high altitude aircraft data.
3. Small cultural details and natural landscape details can be observed, thus reducing the need for an intermediate aerial photographic stage in some multi-stage sample designs.
4. Because high resolution in the first "spacecraft" stage of a multi-stage sample design enables better boundary delineation and category separation, greater precision is obtained in the estimate of natural production of volume (timber, grazing land, agricultural production, etc.) than with LANDSAT data.

5. The higher resolution of the S-190 B will greatly increase the analyst's confidence in reconnaissance land, soil association, and natural systems mapping.

6. For many applied areas the improved resolution brings the data interpreted from S-190 B images into the realm where existing data requirements of land-oriented agencies may be met. Alternatively, if they are not fully met, the degree of modification of their procedures in order to accept S-190 B data is likely to be generally acceptable.

In short, there is an important role to be played on SPACE SHUTTLE for wide-area coverage, high resolution camera systems in land use and other resource analysis. This conclusion is hardly startling - the same point has been made for a decade in the tables of resolution needs for various disciplines, and has previously been documented in studies by Wobber (1970), Colwell (1969), and Simonett et.al. (1969).

Finally, the results of these high-resolution studies lend strong support to the increased value of 30-meter resolution in the Thematic Mapper for LANDSAT D. Because of the parallel high sensitivity design goal along with the 30-meter resolution, much smaller features will be observed in the Thematic Mapper than with the LANDSAT MSS. They may indeed be not too dissimilar to those observed with the S-190 B system. This possibility suggests that these high resolution and high sensitivity design goals are justified.

1.0 SYSTEM DEVELOPMENT AND DATA PROCESSING

As indicated within the Executive Summary, the application discipline of land use planning and resource management provided the central focus for the investigations performed during this project. The two categories of investigations (i.e., digital analysis and photointerpretive analysis) will be discussed separately in this report. Analysis of the digital data required to address the fundamental questions of interest to this investigation required that a rather complex data processing/analysis system be designed, developed, and implemented. This System, referred to as the SKYLAB Data Processing/Analysis System, is the subject of the first two sections of this report. After a brief overview of the entire System (see Section 1.1), the remainder of Section 1 is devoted entirely towards describing the data processing phase of the SKYLAB Data Processing/Analysis System. As will be seen, a large percentage of the total effort required to analyze the multispectral scanner data was devoted towards pre-processing and manipulating the various input data products prior to, and in preparation for, the analysis phases of the investigation which followed. Section 2 of the report discusses the data analysis phase of the System. Analysis techniques, methodologies, as well as the results obtained from the analysis are presented in that section.

1.1 System Overview

Figure 1 shows a general flow diagram for the SKYLAB Data Processing/Analysis System. As indicated on this figure, each processing block is discussed separately in the report. The System was initially designed to allow the greatest possible flexibility during

SKYLAB DATA PROCESSING / ANALYSIS SYSTEM

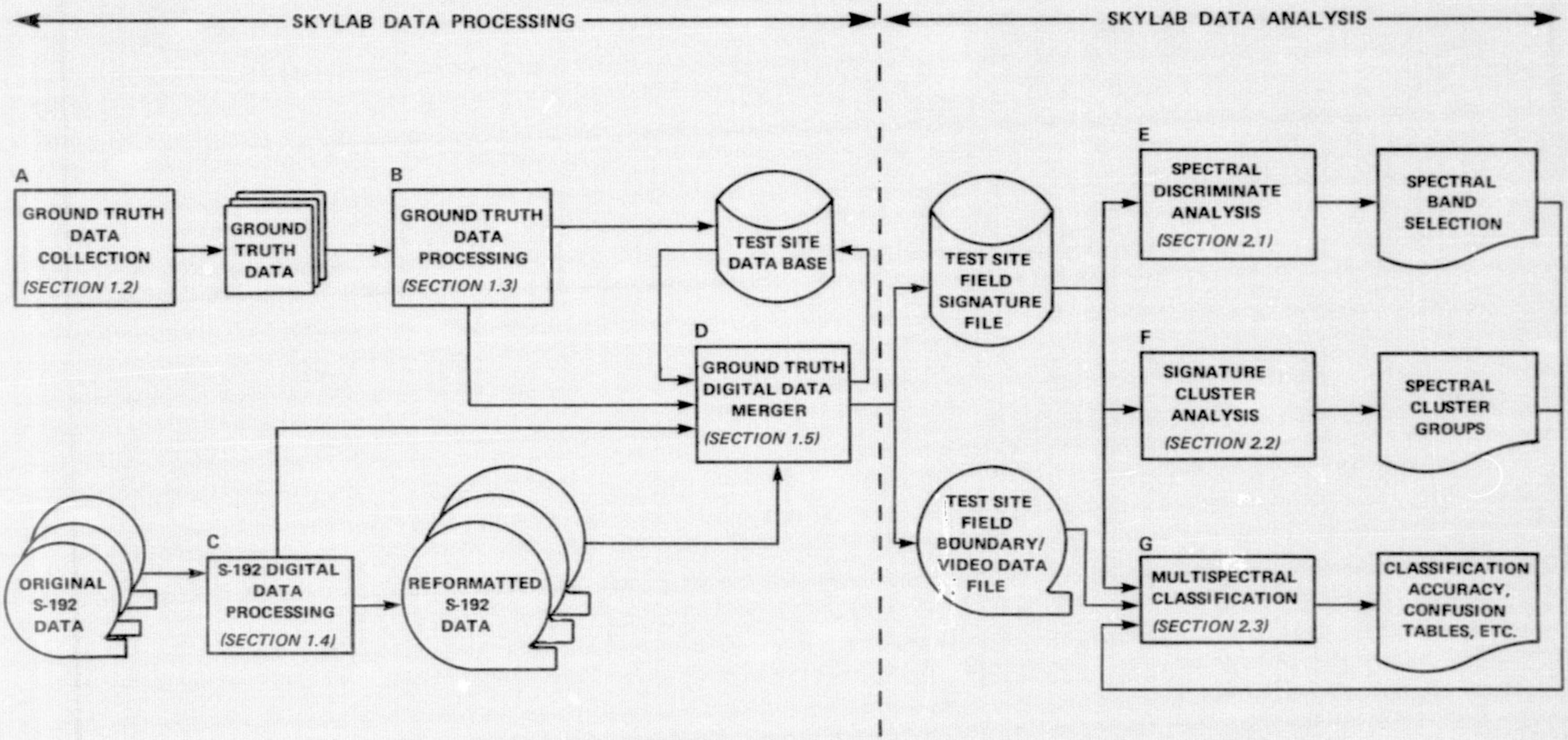


Figure 1.

the analysis phase of the investigation. As a result, it was possible to allow the results of a partial analysis of the data to direct and focus the remaining analysis steps without requiring a great deal of additional data processing. For example, the results of the spectral discriminant analysis (see processing block E on Figure 1) indicated that a multistaged, multispectral classification algorithm may provide the best classification technique. This technique was then implemented and tested during the multispectral classification analysis task (processing block G on Figure 1) with no additional data processing required.

1.1.1 SKYLAB Data Processing

As indicated by Figure 1, the principal task performed by the SKYLAB Data Processing component of the System was the integration of the ground truth data with the S-192 digital data to form a composite ground truth/digital data file to be used during the SKYLAB data analysis tasks. This task was basically a large data management problem. The magnitude of the data management problem is realized by noting that the test area coverage provided just by SKYLAB Mission SL/3 included nearly 5.5 million acres of ground. Even though an extensive ground truth collection effort was conducted, in the final analysis less than 8/10 of 1 percent of this total land area was directly correlated with ground truth data. Although this percentage figure is small, it represents nearly 40,000 acres of land for which ground truth data was correlated with S-192 digital data at the pixel level and subsequently analyzed.

The ground truth data processing tasks therefore provided another important function. In addition to integrating the various data sources it also performed a data compression so that subsequent analysis tasks could be performed using manageably sized data sets. For example, the common practice of performing multispectral classification on areas for which no ground truth data is available is an interesting, but generally unproductive process when the goal of the investigation is to determine the achievable accuracy levels. Without ground truth information it is impossible to know whether the results are correct or incorrect. The data files generated through the SKYLAB data processing tasks allowed multispectral classification to be easily performed using only those pixels for which ground truth is available. This type of data structure therefore allowed a more meaningful investment of the resources available to the investigation.

Processing blocks A and B on Figure 1 represent the collection and pre-processing of the ground truth data necessary to prepare for merger with the S-192 digital data (processing block E). Processing block C represents the processing necessary to prepare the S-192 digital data for merger with the ground truth data. Two data files were produced during the ground truth data/ digital data merger. The first file contained test site field boundary maps in the corresponding 13-channel digital data for all of the test sites covered by a particular SKYLAB overpass. The second file was produced by calculating 13-channel spectral signatures for all of the ground truth fields identified on the first file. These two data files then provided all of the information needed to perform the SKYLAB data analysis tasks.

1.1.2 SKYLAB Data Analysis

The only data inputs required to perform the SKYLAB data analysis tasks, shown as blocks E, F, and G on Figure 1, were the two files produced by the SKYLAB data processing tasks. The spectral discriminate analysis (processing block E) used the test site field signature file to analytically determine which spectral bands provided the greatest spectral separability between various land use categories. This information was subsequently incorporated in the design of the multispectral classification algorithm. The signature cluster analysis (processing block F) performed an unsupervised cluster analysis again using the test site field signature file. The results of this analysis indicated which land use categories displayed similar spectral characteristics. This analysis provided useful information about the types of land use categories which could, or could not, be accurately identified by a multispectral classification analysis. The results of this analysis were used in the selection and stratification of test and training sets for the multispectral classification analysis (processing block G). The multispectral classification analysis used the test site field signature file to generate the training statistics used in the analysis and accessed the test site field boundary/video data file to perform the actual analysis. As previously indicated, the results of the spectral discriminate analysis and the signature cluster analysis were used to help design the multispectral classification algorithm. The intention was to attempt to optimize the trade-offs between the classification accuracy and computational cost.

1.2 GROUND TRUTH DATA COLLECTION

The ground truth data collection in the Washington - Baltimore and vicinity test site was undertaken to provide a statistically acceptable data set, well-distributed over the diverse landscapes from the eastern edges of the Appalachian Mountain system to the very flat coastal areas of the lower Delmarva Peninsula.

The data set so obtained was designed not only to provide complete ground truth for the sample, but also the sample was intended to be large enough that training and prediction could be partitioned within the large spatially - distributed sample.

1.2.1 Test Area Selection

The area was chosen because of proximity to the investigator's offices - and hence to keep field work costs within bounds - and because the area was already under investigation by the U.S.G.S. in their CARETS program, and much data was being accumulated for the area. In addition, LANDSAT studies with the Department of State Planning of Maryland being carried out by Earth Satellite Corporation were also underway and were expected to provide additional support to the investigation. Finally, the area is a region of considerable diversity and complexity, with relatively small units of mixed rural and urban landscapes which would provide a significant, and demanding, test of the various instruments on SKYLAB.

Included in the 6,678-square mile Washington-Baltimore test site are representative areas of various morphological features, including the piedmont section of Virginia and western Maryland, the deeply dissected coastal plain area of the western shore of the Chesapeake Bay, and the very flat coastal areas of the lower Delmarva Peninsula. Portions of Chesapeake Bay, as well as estuaries of the Potomac and Patuxent Rivers are also within the test site.

The test site exhibits considerable diversity in land use. Most of Washington, D.C. metropolitan area is included in the site, which is an area of great urban-rural complexity, expecially in the urban fringe areas where suburban development is encroaching on agricultural lands. The areas outside the Washington metropolitan area are largely agriculture-based, and contain large areas of forest and wetland plant communities.

1.2.1.1. Selection Criteria

The criteria for selection of this site, in addition to those already noted include:

1) Preliminary testing of aircraft and spacecraft - based land use classification schemes had already been carried out in this area.
2) The Washington D.C. - Baltimore area were both instituted as census cities in the U.S.G.S. program of high altitude aircraft flights during the summer of 1970, to coincide with the National Census. Thus there was aircraft data, comparable to that obtained by NASA underflights after time of SKYLAB passage, but dating back to 1970.
3) Mapping of the entire test site area by the U.S.G.S. at a scale of 1:100,000 was underway using the 1970 high altitude aircraft data. This employed the Anderson, Hardy, and Roach (1972) land use classification used as the base for the present study.

1.2.1.2. Geographic Location

The area covered in the S-192 strip for SL/3 is bounded by the following co-ordinates:

Northwest corner:	77^{o} 2' W, 39^{o} 23' N
Southwest corner:	75^{o} 49' W, 38^{o} 6' N
Northeast corner:	77^{o} 31' W, 39^{o} 53' N
Southeast corner:	75^{o} 17' W, 38^{o} 37' N

1.2.1.3. SL/2 - SL/3 Overlap and Orbit Changes

As noted in the Executive Summary the shift in orbit between SL/2 and SL/3 was most unfortunate, in that the shift was of sufficient magnitude to largely negate the multidate overlay intended as part of the experiment. This was further exacerbated by cirrus and other clouds in portion of the overlap strip, and by the inability to precisely identify some sample sites on the alphanumeric print outs of the S-192 tapes. As a result of these composite effects, the number of identified cloud - free samples in the overlap zone dropped to a level where most land use classes were so underrepresented (some were not represented at all) that no valid comparisons could be made. Figure 2 shows the area covered by the SL/3 pass.

It was considered desirable to obtain additional field data in the overlap strip and in fact a considerable number of additional samples were obtained, which at first were expected to be enough for multi date comparisons. However, it was not realized at the time how thin cirrus and patchy clouds would seriously degrade the sample, nor was it realized, in the more bland, undifferentiated or very segmented areas, lacking water bodies, how difficult - indeed infeasible - it would be to make accurate identifications of locations. With many small fields and small urban sites a 1 pixel misregistration was a serious concern

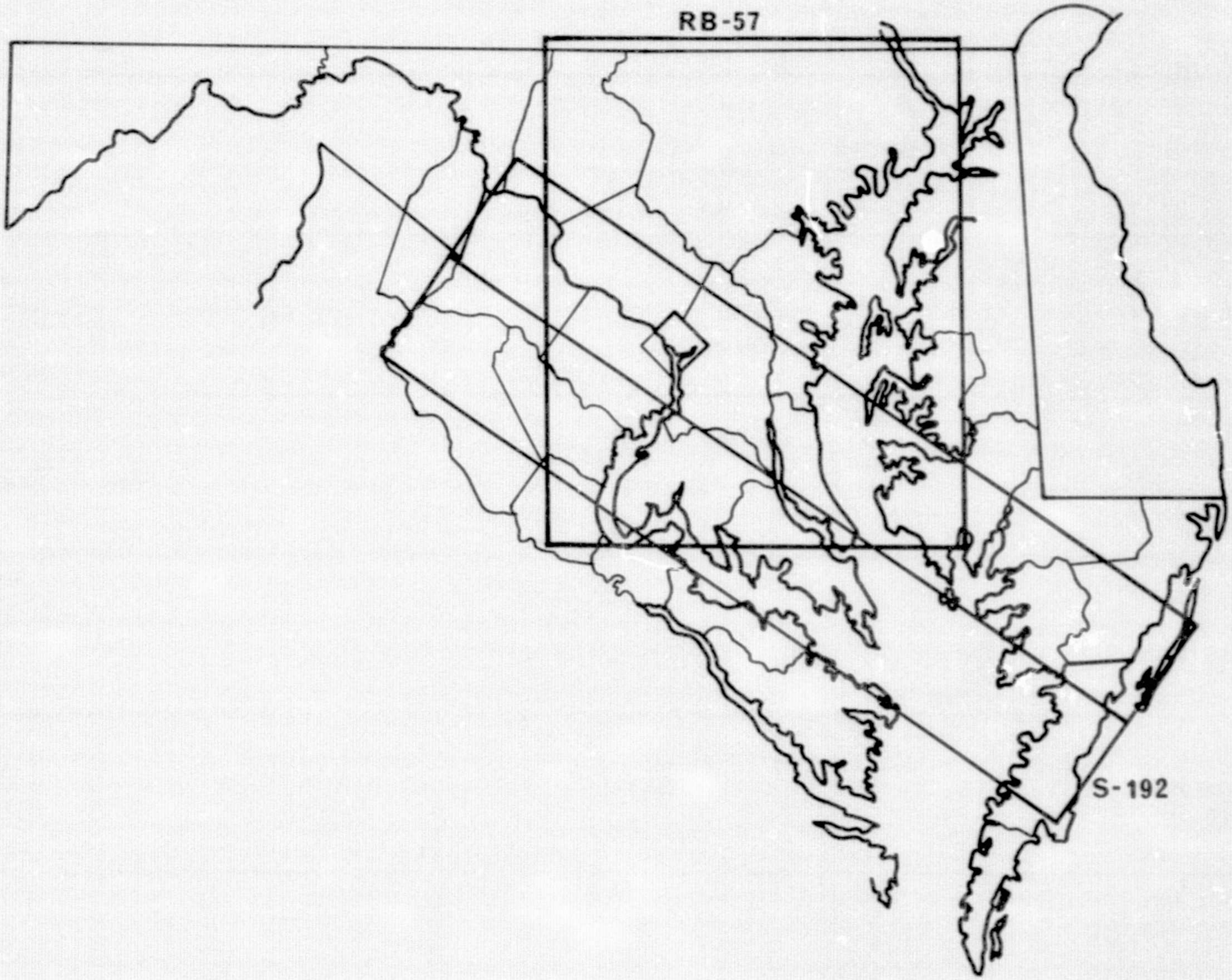


Figure 2. WASHINGTON, D. C. - BALTIMORE SKYLAB TEST SITE

between dates. In numerous sample sites it was not possible to locate areas any closer than within 2 to 5 pixels of the correct location. Data misregistered to this degree could not be merged and provide a proper test of multi-date value.

Multi-date S-192 comparisons were reluctantly dropped as a consequence.

1.2.2. Test Area Sampling

Since only the SL/3 mission is reported on here, neither the SL/2 mission ground truth collection, nor the overlap sites will be considered.

A sampling frame procedure was developed for the Washington, D. C. test site to provide a 6.25% sample of the entire area for ground truth data collection. A systematic, unaligned, locationally stratified, random design which has proved highly successful in land use studies (Holmes, 1967; Holmes, 1970) was selected. The systematic grid with cells of constant size provided locational spread, equal selection probability, and a statistically representative approach to randomness. The test site was divided into grid cells of approximately 16 square miles each on 1:250,000 map sheets. Each grid cell was then subdivided into 16 one-square mile units, and a random number table was used to generate the selection of a one-square mile site within the 16-square mile grid. This automatically gave a 6.25% sample, well distributed throughout the area and with both systematic and random components to minimize sampling error.

After the selection of the sample plots, the one-square mile areas were transferred from the 1:250,000 map sheets to existing 1:120,000 RC-10 color infrared pnotography. A polaroid MP-3 copy camera was then used for black and white enlargements of the individual test sites for the field work. The photo enlargements provided a valuable tool in locating sample plots, defining land use boundaries, and simplifying the ground truth data collection effort.

The sample sites actually used in the investigation are listed by occurrence in the counties of Virginia and Maryland and Baltimore and Washington in Table 1. It will be seen later that in the event a much smaller useable sample than 6.25% was used in this study. In fact the final sample remaining after a whole series of exclusions totalled slightly less than 1%. This is, still a very substantial sample equivalent to some 63 square miles. The exclusions took place because of inadequacies in the ground truth collection, banks of deep cloud cover in the southeastern portion of the area, patchy clouds, cirrus veils and trails, inability to identify exact locations of some sample sites, and inability within a sample site to identify boundaries of categories, and also because of elimination of many very small fields as being unsuitable - through an extreme mixed-pixel effect - for either training or prediction.

1.2.2.1 <u>Land Use Categorization</u>

The level of generalization of land use categories (i.e., gross, moderate, and detailed) originally proposed for field work in the Washington, D.C. area corresponds with levels I, II and III of the Anderson, et al. classification (1972). In conjunction with other Earth Satellite Corporation land use projects, a revised land use inventory and classification scheme was adopted which provides information needs that are responsive to many federal and state agencies involved in land use planning. The land use classification scheme

TABLE 1. WASHINGTON, D.C. - BALTIMORE, MARYLAND SKYLAB TEST SITE

List Of Counties In Which Ground Truth Was Collected For The SL/2 and SL/3 Passes.

County Name	State	# of Testsites SL/2 Only	SL/3 Only	SL/2 & SL/3	Total #
Accomack	VA.	25	0	0	25
Anne Arundel	MD.	1	26	2	29
Arlington	VA.	3	0	0	3
Baltimore City	MD.	0	3	0	3
Baltimore County	MD.	0	14	0	14
Calvert	MD.	7	0	4	11
Carroll	MD.	0	9	0	9
Charles	MD.	25	0	0	25
Dorchester	MD.	20	0	6	26
Fairfax	VA.	21	0	0	21
Fauquier	VA.	20	0	0	20
Howard	MD.	0	15	0	15
King George	VA.	1	0	0	1
Loudoun	VA.	29	0	0	29
Montgomery	MD.	2	11	16	29
Prince Georges	MD.	11	9	11	31
Prince William	VA.	19	0	0	19
Somerset	MD.	18	0	7	25
Stafford	VA.	1	0	0	1
St. Mary's	MS.	26	0	0	26
Talbot	MD.	0	0	1	1
Washington	DC.	0	0	2	2
Washington	MD.	1	0	0	1
Westmoreland	VA.	2	0	0	2
Wicomico	MD.	0	0	6	6
Worchester	MD.	3	0	13	16
Totals		235	87	68	390

(levels I and II) used in the study thus was based directly on that developed by the U.S. Geological Survey (Anderson, et al., 1972), while levels III and IV were devised in conjunction with the Maryland Department of State Planning, and met their needs in land use mapping.

The classification scheme was developed on the assumption that different levels of classification would be derived from different sources of information, and in general the relationship can be shown as:

Classification Level

I. Satellite imagery

II. High altitude and satellite imagery combined with topographic maps.

III. Medium altitude remote sensing (1:20,000) combined with detailed topographic maps and substantial amounts of supplemental information.

IV. Low-altitude imagery, most of the information derived from supplemental sources, including ground observations.

Field work sheets were then prepared for level III and level IV information for agricultural and urban areas. These provided a consistent reporting form for the field teams and can be easily transferred into computer readable format.

The land use classification employed is given in Table #2.

1.2.2.2. Impact of Stratification on Results

Prior to the launch of LANDSAT the collection of ground truth was very much subject to the procedures adopted by individual investigators. Any suggestion of an explicit, statistically random design was quite rare. In practice any convenient location tended to be chosen with little

Table 2. Land Use Classification Employed in this Study

1000-Urban & Built-up Land
- 1100- Residential
  - 1110-Single-Family Household Units
    - 1111-Rural, Low Density, With Trees
    - 1112-Rural, Low Density, Without Trees
    - 1113-Urban, High Density, With Trees
    - 1114-Urban, High Density, Without Trees
  - 1120-Multi-Family Household Units
    - 1121-Low Density, With Trees
    - 1122-Low Density, Without Trees
    - 1123-High Density, With Trees
    - 1124-High Density, Without Trees
  - 1130-Group Quarters (Rooming Houses, Lodgings, Homes, Camps...)
  - 1140-Residential Hotels
  - 1150-Mobile Home Parks or Courts
    - 1151-Low, With Trees
    - 1152-Low, Without Trees
    - 1153-High, With Trees
    - 1154-High, Without Trees
  - 1160-Transient Lodging
  - 1190-Other
- 1200-Commercial & Services
  - 1210-Wholesale Trade Areas
    - 1211-Motor Vehicles and Automotive Equipment
    - 1212-Drugs, Chemicals, and Allied
    - 1213-Drygoods, Apparel, and Footwear
    - 1214-Groceries and Related Products
    - 1215-Farm Products (Raw Materials)
    - 1216-Electrical Goods
    - 1217-Hardware, Plumbing, Heating Equipment, and Supplies
  - 1220-Retail Trade Areas (Business Dist., Shopping Centers, Comm.)
    - 1221-Building Materials, Hardware and Farm Equipment
    - 1222-General Merchandise
    - 1223-Food
    - 1224-Automotive, Marine, Aircraft Accessories
    - 1225-Apparel and Accessories
    - 1226-Furniture, Home Furnishings, and Equipment
    - 1227-Eating and Drinking Establishments
    - 1228-Other Retail (Drugs, Liquor, etc.)
  - 1230-Business, Professional, & Personal Services
    - 1231-Finance, Insurance, Real Estate
    - 1232-Personal Services
    - 1233-Business (Advertising, Employment Services, Storage, etc.)
    - 1234-Repair Services
    - 1235-Professional Services
    - 1236-Contract Construction Services
    - 1237-Utilities
  - 1240-Cultural, Entertainment & Recreational Facilities
    - 1241-Cultural and Nature (Libraries, Galleries, Monuments, etc.)
    - 1242-Entertainment (Theatres, Amphitheatres, Drive-ins, etc.)
    - 1243-Sports (Stadium, Arenas, Racetracks, Other)
    - 1244-Public Assembly (Auditorium, Exhibition Halls)
    - 1245-Amusement (Fairgrounds, Amusement Parks, Miniature Golf)
    - 1246-Recreational (Tennis, Ice Skating, Stables, Play Areas...)
  - 1290-Other

Table 2. (Continued)

1300-Industrial
  1310-Mechanical Processing
  1320-Heat Processing
  1330-Chemical Processing
  1340-Fabrication & Assembly
  1350-Food Processing
  1360-Other
1400-Extractive
  1410-Stone Quarries
  1420-Sand & Gravel Pits
  1430-Open Pit or Strip Mining
    1431-Type
  1440-Oil, Gas, Sulphur, Salt, & Other Wells
    1441-Type
  1450-Shaft Mining
    1451-Type
  1490-Other
1500-Transportation, Communication & Utilities
  1510-Highways, Auto Parking, Bus Terminals, etc.
    1511-Highways
    1512-Parking Areas
    1513-Bus Terminals
    1514-Motor Frieght
  1520-Railroads & Associated Facilities
  1530-Airports & Associated Facilities
    1531-Commercial
    1532-General
    1533-Military
  1540-Marine Craft Facilities
    1541-Dredged Channel
    1542-Jetty
    1543-Combination
    1544-Port Facilities
  1550-Telecommunications, Radio & Television Facilities
  1560-Electric, Gas, Water, Sewage Disposal, Solid Waste, Util.
  1590-Other
1600-Institutional
  1610-Educational Facilities
    1611-Primary
    1612-Secondary
    1613-Junior College
    1614-College
    1615-University
    1616-Other
  1620-Medical & Health Facilities
    1621-Hospitals
    1622-Sanitariums
    1623-Other
  1630-Religious Facilities
    1631-Churches
    1632-Church Affiliated Buildings
    1633-Other

Table 2 . (Continued)

- 1640-Military Areas
  - 1641-Housing
  - 1642-Administration Buildings
  - 1643-Storage Areas
  - 1644-Training Areas
  - 1645-Other
- 1650-Correctional
  - 1651-Local
  - 1652-County
  - 1653-State
  - 1654-Federal
- 1660-Government & Administrative Offices
  - 1661-Type
- 1690-Other

1700-Strip & Clustered Settelment

1800-Mixed

1900-Open & Other
- 1910-Improved
  - 1911-Golf Courses
  - 1912-Cemeteries
  - 1913-Park
  - 1914-Parking Lots
- 1920-Unimproved
- 1990-Other

2000-Agriculture Land
- 2100-Cropland and Pasture
  - 2110-Active Cropland
    - 2111-Fallow
    - 2112-Bare--Recently Plowed
    - 2113-Growing Crop Present
    - 2114-Harvested
  - 2120-Abandoned
  - 2130-
  - 2140-Pasture
    - 2141-Wooded
    - 2142-Improved
    - 2143-Unimproved
  - 2190-Other
- 2200-Orchards, Groves, Bush Fruits, Vineyards, Horticulture
  - 2210-Fruit and Nut Trees
    - 2211-Apple
    - 2212-Peach
  - 2220-Bush Fruit
    - 2221-Type
  - 2230-Vineyard
    - 2231-Type
  - 2240-Nurseries and Floricultural Areas
    - 2241-Type
  - 2250-Turf Farm
    - 2290-Other
- 2300-Feeding Operations
  - 2310-Cattle Feed Lots (incl. Holding Lots for Dairy Animals)
  - 2320-Poultry and Egg Houses
  - 2330-Hog Feed Lots
  - 2390-Other

Table 2. (Continued)

3000-Rangeland
- 3100-Grass
- 3200-Savannas (Palmetto Prairies)
- 3300-Desert Shrub

4000-Forestland
- 4100-Deciduous
- 4200-Evergreen (Coniferous & Other)
- 4300-Mixed
- 4400-Upland Shrubs
- 4500-Lowland Shrubs
- 4600-Lumbering

5000-Water
- 5100-Streams and Waterways
  - 5110-Natural (Rivers & Creeks)
  - 5120-Man-Made (Canals, Ditches, & Aquaducts)
- 5200-Lake
  - 5210-Natural Lakes & Ponds
  - 5220-Man-Made Lakes & Ponds
- 5300-Reservoirs
- 5400-Bays & Estuaries
  - 5410-Bays
  - 5420-Estuaries
- 5900-Other

6000-Wetlands
- 6100-Vegetated Wetlands
  - 6110-Brackish Marsh
  - 6120-Fresh Water Marsh
  - 6190-Other
- 6200-Forested Wetlands

7000-Barren Land
- 7100-Salt Flats
- 7200-Beaches
  - 7210-Sandy Beaches
  - 7220-Gravelly, Rocky Beaches
  - 7230-Mud Shorelines
- 7300-Sand Other Than Beaches
- 7400-Bare Exposed Rock
- 7500-Disturbed Land
- 7900-Other

8000-Tundra

9000-Permanent Snow & Ice Fields

consideration of its representativeness or otherwise. The same pattern also prevailed in the early years of LANDSAT, and only in rare instances- such as in the study by Von Steen and Wigton (1973) - was there a conscious effort to randomize ground truth site selection, either for developing a training set for mapping, or - as in the present study - to provide a sufficient data base for analysis of band selection and clustering.

Except in an area of complete uniformity, the degree of spatial variability encountered in most moderate sized areas - such as the 6,678 square mile region investigated in this study - demands some form of systematic, yet random sample design. These are the principal factors which play some role in necessitating a systematic/random sampling procedure:

1) Differences in physical land types bring parallel differences in the mix of land use types in each land type unit. Land types as used here signifies such items as:

   Plateau - sandstone
   Valley(I) - limestone
   Mountain - acid igneous intrusive
   Valley(II) - glacial debris
   Coastal Plain - high marine terrace, etc.

   There should be a sufficient number of samples in each significantly different physical land type unit, normally about 20 to 30 cases.

2) In some areas, different physical land types are finely interfingered, necesitating a random yet systematic spatial sampling procedure to ensure that each unit is adequately represented.

3) Finally, there may be sufficient variability within each natural unit stratum, that at least 20 and preferably 30 cases of each unit should be obtained.

These cases themselves should contain enough examples of the land uses to constitute a proper sample.

These factors when taken together led to a choice, as noted earlier of a 6.25 % sample (one square mile in every 16) as the initial sampling percentage.

The method of distributing these samples spatially, again as noted earlier in 1.2.2, was the method employed by Berry and extensively analyzed as to efficiency by Holmes (1967; 1970) - namely a systematic, unaligned, locationally stratified, random design. This design is very efficient in dealing with spatial variability and the systematic grid with cells of constant size (one square mile) provided the necessary locational spread, equal and high (6.25%) selection probability, and a statistically representative approach to randomness through use of a random number table for each 16 square mile unit.

In practice the sample size dropped to 1% or less, for the reasons outlined earlier, and the spatial distribution tended to cluster in the central part of the area near Washington & Baltimore (the most cloud-free area!). This stratification procedure was important nevertheless in retaining some measure of randomness in the analysis, and ensuring that there would be representation of the various cloud-free physical environments in the data set used in the analysis.

1.2.3. Ground Truth Data Collection Procedure

The collection of ground truth data was divided into two sections: 1) changing environments, i.e., agricultural areas; and 2) static environments, i.e., urban, forest, and wetlands. Field work was conducted initially in the rural areas of change, and more leisurely in the urban and other slowly changing areas. Teams were in the field for a period of approximately two weeks around the time of the SL/3 overpass (August 5, 1973). Data was collected for 390 test sites, one square mile each. Of the 390 test sites, 303 were covered by the SL/2 overpass, 155 by the SL/3. The sample plots were located in 26 counties in Maryland and Virginia, and the District of Columbia (Table 1). Teams of two persons covered the various sites during each overpass 2-week period. A flow chart of the ground truth operations is shown in Figure 3.

Field work was conducted principally to obtain detailed (level III and IV) land use information in agricultural and urban areas. Detailed forest/vegetation maps were supplied through a cooperative arrangement with the Maryland Department of State Planning. Detailed wetlands data were provided through cooperative arrangements with ERTS-1 investigators in the area who had done extensive field work with mapping of the coastal wetlands in Maryland and Virginia.

The RC-10 photography flown on June 14, 1973 was used to establish the test site boundaries for the SL/3 mission which took place on August 5, 1973. The original intention, when the random number generator was used, was to collect ground truth on the entire one-square mile site. Data was collected over a larger area than one-square mile, but as will be seen later, a much smaller area than one square mile was actually used in these studies. Polaroid pictures were taken of the site and surrounding area to facilitate location

GROUND TRUTH DATA COLLECTION PROCEDURES

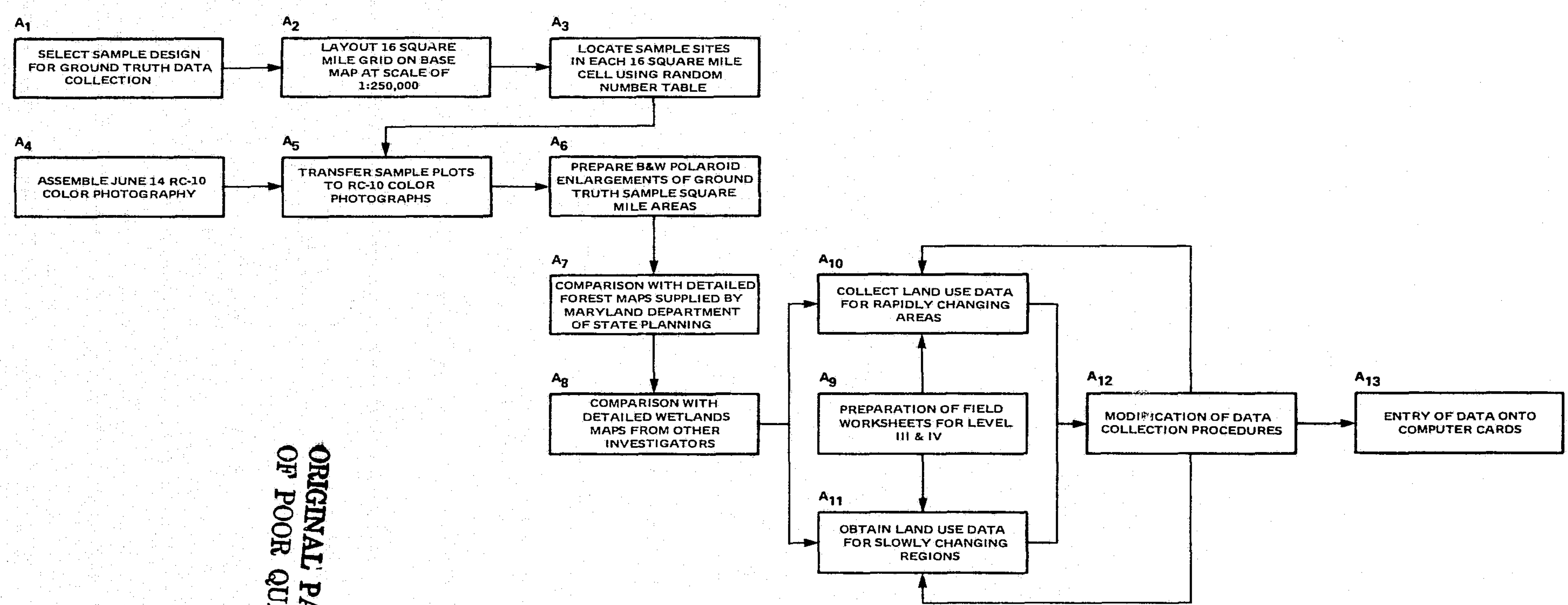


FIGURE 3.

and collection efforts by the ground truth team. Since small plots of ground truth were of little use when testing with a one-acre resolution system, only land use features larger than about nine acres and along a passable road were accepted as suitable for ground truth collection.

To assure maximum use of the gorund truth data, very detailed data was collected where possible. For example, in urban residential categories data was collected on the width of building material in streets and sidewalks, and type of rooftop material. This type of information was collected in the expectation that in a complex urban environment, numerous materials which contribute to the signal may occur within one acre as viewed by the scanner. This detailed data was judged, for the most part, to be unusable for digital analysis, (it was, however, employed in various ways in photo-interpretation studies).

In lieu of this, a scheme was devised for urban areas at level III and IV which took into account the relative density of the residential area and whether the area had greater than 50% cover of trees as determined from the RC-10 color infrared photography. These two factors are the most important ones to consider for numerous reasons. First, no matter what type of building materials are used on houses and roads, if the area is tree-covered and scanner imagery is being used the IR reflectance will bloom relative to other components of the signal. If no tree cover exists, it is more important to know how the buildings are spaced than to know the building material. If housing density is low, what occurs between the houses, yards, swimming pools, etc., becomes very important in determining the sensor signal. If the housing density is high, the resultant signal is a ratio of all signals within the acre; namely, roofs, sidewalks, driveways, swimming pools, yards, etc.. The task of weighting the contribution of each and correlating it to the sensor signal would be highly complex and fruitless in its results.

Coordination and organization of the ground truth effort was complicated by the uncertainty of where the SKYLAB system would collect data. SL/2 collected data approximately 60 miles south of the original target site. When SL/3 collected data over the site on August 5, the SKYLAB path had been corrected. As a result, there was an overlap zone of only 10 miles in width. A third SKYLAB overpass was also anticipated at right angles to the two passes, running directly over the Baltimore-Washington urban areas. Thus, the 10% overlap area and the Baltimore-Washington cities became the focus of the ground truth effort in SL/3 in anticipation of obtaining three-date imagery. In fact it was not feasible to obtain the third imagery pass, and thus much wasted effort was involved in data collection.

In anticipation of three-time small-area coverage the ground truth data collection went through an evolution until an efficient procedure evolved. As noted earlier, it was intended to collect ground truth for each entire one-square mile site in order to meet the requirements of the sample design. This was found to be far too time consuming and impractical, since many areas were inaccessible by car or were private property. The second phase involved collection of ground truth data for <u>all</u> features along the roads within the test site photograph. This also was judged to be impractical, since very many land use features in this region occurred in units of one acre or less. Assignment of a pixel to a specific area on the ground can be done to within one pixel, only in areas of high contrast and sharp edges. Furthermore, because a training field of less than about nine acres (or pixels) would have numerous boundary pixels associated with it, the location of those pixels with even a one-pixel tolerance would result in a high boundary pixel error in signature calculation, in classification, or the use of clustering algorithms. For all of these reasons only the larger land use areas exceeding ten acres were finally utilized.

It is clear from this discussion that an "ideal" sampling frame would in practice be extraordinarily difficult to achieve. In this area it was infeasible. The experience was most instructive on the considerable difficulties which will be encountered in developing and using any spatially well-distributed sample frame involving the merging of multi-date images.

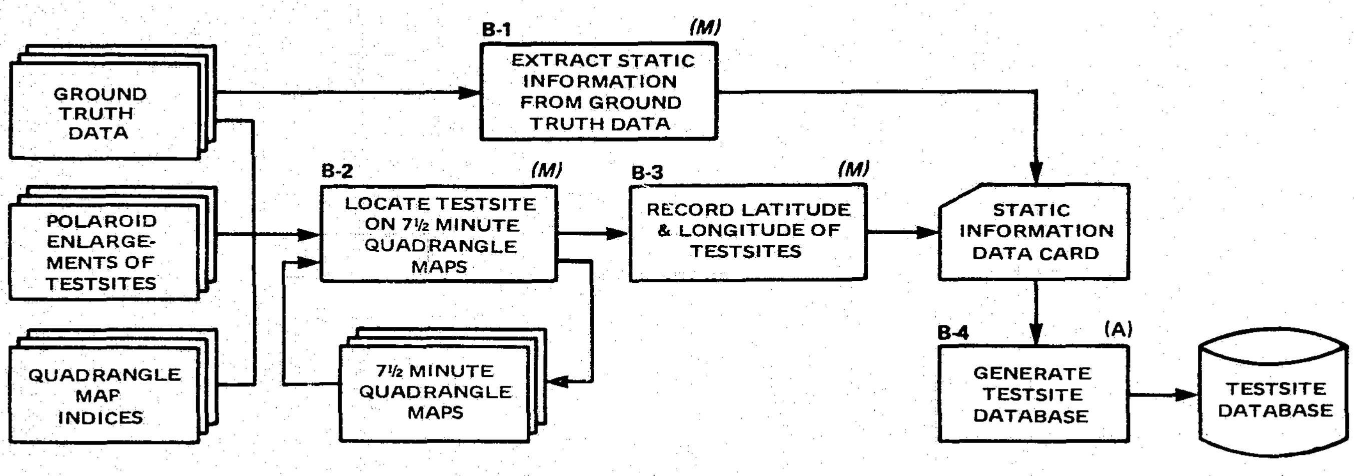
1.3 Ground Truth Data Processing

The data collected by the ground truth collection teams provided primary inputs to the investigations which followed. Through the use of the ground truth data it was possible to identify, on the S-192 digital data tapes, individual fields within each test site. The ground truth data was used to assign specific land use categories to these fields and thus provided a basis on which the S-192 multispectral scanner data could be categorized and stored. In order to perform these tasks, it was first necessary to record the ground truth data in a computer compatible form and then to establish a relationship between the ground truth test sites and the corresponding digital data. Processing of the ground truth data required to establish this relationship was both manual and computerized. An expanded task flow diagram for processing block B on Figure 1 is shown in Figure 4. The processing procedures and methodologies applied to the ground truth data will be discussed in this section.

1.3.1 Delineation Of Test Sites

The ground truth data collected by the collection teams provided information about selected fields within each test site. In addition, the specific fields were marked and labeled on the black and white polaroid enlargements. In order to incorporate these data into the analysis, it was necessary to identify the individual pixels of the S-192 digital imagery corresponding to the fields selected by the collection teams. The first step in this procedure was to accurately determine the geographic location of each test site.

GROUND TRUTH DATA PROCESSING



(A) INDICATES PROCESSING TASKS INVOLVING AUTOMATED PROCESSING

(M) INDICATES PROCESSING TASKS INVOLVING MANUAL PROCESSING

Figure 4.

1.3.1.1 Identification Of Test Site Geographic Locations

The final identification between individual fields and their corresponding pixels on the S-192 digital imagery had to be accurate to within a pixel. The accuracy limits required for the identification of the test sites were less severe. The most convenient means of specifying the location of the test sites was through the identification of the latitude and longitude.

The approximate location of each test site had been marked on county maps for use by the ground truth collection teams. These county maps were used to locate the test site areas on U.S. Geological Survey 1:500,000 topographic index maps. The U.S. Geological Survey topographic index maps specified the index map names for U.S. Geological Survey quadrangle maps covering 7 1/2 minutes of latitude and longitude. U.S. Geological Surveys 7 1/2 minute quadrangle maps were obtained for all areas which included the Skylab test sites.

Utilizing the 7 1/2 minute U.S. Geological Survey quadrangle maps, the polaroid black and white enlargements and the ground truth data, the location for each test site was identified and marked on the quadrangle maps (see processing block B-2 on Figure 4). Street patterns, state highway route numbers, ponds, streams and other distinguishable geographic characteristics aided greatly in the identification procedure.

1.3.1.2 Recording Test Site Latitude and Longitude

After a specific test site had been located on a 7 1/2 minute quadrangle map, a rectangular area 2 1/2 minutes of latitude and 2 1/2 minutes of longitude was centered about the test site. The rectangular areas were drawn in on the 7 1/2 minute quadrangle maps and the latitude and longitude of the northwest corner of the rectangles were recorded (processing block B-3). The latitudes and longitudes recorded for the northwest corners of the 2 1/2 minute rectangles provided the bases for identifying the test site areas on the S-192 digital imagery. The latitudes and longitudes were keypunched on computer cards and then checked for accuracy by a different photo interpreter to minimize the chance of error.

1.3.2 Test Site Data Base

It soon became apparent that a great deal of bookkeeping would be necessary to keep track of the current processing status of each test site. In order to minimize the bookkeeping effort, a computerized random access test site data base was designed and implemented. In the early phases of the ground truth data processing, the test site data base was simply used to keep record of what information was available for each test site and which test sites had been located on the U.S. Geological Survey quadrangle maps. In the later phases of the test site data processing more extensive use was made of the test site data base. The general characteristics of the test site data base will be described in this section.

1.3.2.1 Test Site Data Base Generation

The test site data base was designed to contain both static and dynamic information. Static information refers to basic information about each test site available at the time the test site data base was created. Dynamic information refers to information about the individual test sites available only after the completion of a particular processing step.

Static information about each test site was recorded on computer cards (processing block B-1). These data included:

1) test site number
2) county number
3) test site Skylab coverage (SL/2, SL/3, or both)
4) RC-10 roll number
5) RC-10 frame number
6) availability of ground truth data for test site
7) availability of polaroid enlargements of test site

The random access test site data base was generated from the computer cards containing this static test site information (processing block B-4).

1.3.2.2 Test Site Data Base Updating

Once the test site data base had been generated, dynamic information was added to the direct access data base through a read/write, update in place, access mode. Two types of data base updating were performed. First, as the test sites were located on the 7 1/2 minute quadrangle maps, the latitude and longitude of the 2 1/2 minute rectangular areas were added to the computer cards containing the static test site information. The test site data base was updated with the latitude and longitude information directly from these computer cards.

The second type of data base updating was performed automatically by the subsequent processing programs. As a particular processing step was completed for an individual test site, the processing programs set flags in the test site data base to indicate completion of that step. Pertinent information calculated at each step about the test sites was also entered into the data base.

Through this procedure it was possible to use the test site data base not only as a record keeping mechanism but also as a source of input information through the processing programs. A detailed description of the data items entered into the data base automatically by the processing programs will be provided as the individual processing steps are discussed.

1.3.2.3 Test Site Data Base Information Retrieval

The random access nature of the test site data base provided flexibility in the procedures available for retrieving information from the file. At the time the data base was generated, one record per test site was entered into the data. The records were keyed to the test site numbers so that record number "N" corresponded to test site number "N". Two types of information retrieval keys were added to the test site records as they were entered into the files. These keys allowed information retrieval by county or by Skylab coverage. The county key on a particular record pointed to the next test site in the file located within the same county. The Skylab coverage key on that record pointed to the next test site in the file included on the same mission or missions as the current test site. By using the retrieval keys on the test site records to define the next test site to process the data base could be scanned to process only test sites within a specific county or covered by a specific Skylab mission. A retrieval key of zero terminated a particular sequence of test sites.

In order to utilize the retrieval keys, it was necessary to have a starting point for each retrieval sequence. The first record in the file, record number zero was used for that purpose. The information stored on this record included the first test site number in the file from each of the 26 counties as well as the first test site number covered by Skylab missions SL/2, SL/3, or both SL/2 and SL/3.

1.4 S-192 Digital Data Processing

Scanline straightened S-192 digital data tapes containing output from all 22 Scientific Data Output (SDO) channels for coverage of the Washington-Baltimore test site area were received for SKYLAB missions SL/2 (overpass date June 12, 1973) and SL/3 (overpass date August 5, 1973). The data format of these tapes was sufficiently different from the standard data format accepted by the digital data processing software available at Earth Satellite Corporation that it was necessary to reformat the digital data to a form compatible with existing software. During the reformatting process, information contained within the header records and the ancillary data blocks of the original S-192 digital data tapes were extracted for later use in the data processing procedure.

The ephemeris data contained within the ancillary data blocks was used to establish a transformation between latitude and longitude and the individual pixel row, column coordinates. After reformatting the data tapes, computer line printer shade prints or gray maps were produced from all 13 spectral bands for 3 of the test site areas. Using these shade prints, project photointerpreters selected two spectral bands for each pass which appeared to provide the greatest visual discrimination between land use categories.

Techniques were developed to "mass produce" shade prints of the test site areas. The shade print program developed under this contract required only the latitude and longitude of the upper lefthand corner of the test site in order to produce a contrast-enhanced, geometrically rectified, shade print of that test site area. Figure 5 shows an expanded task flow diagram for the S-192 digital data processing.

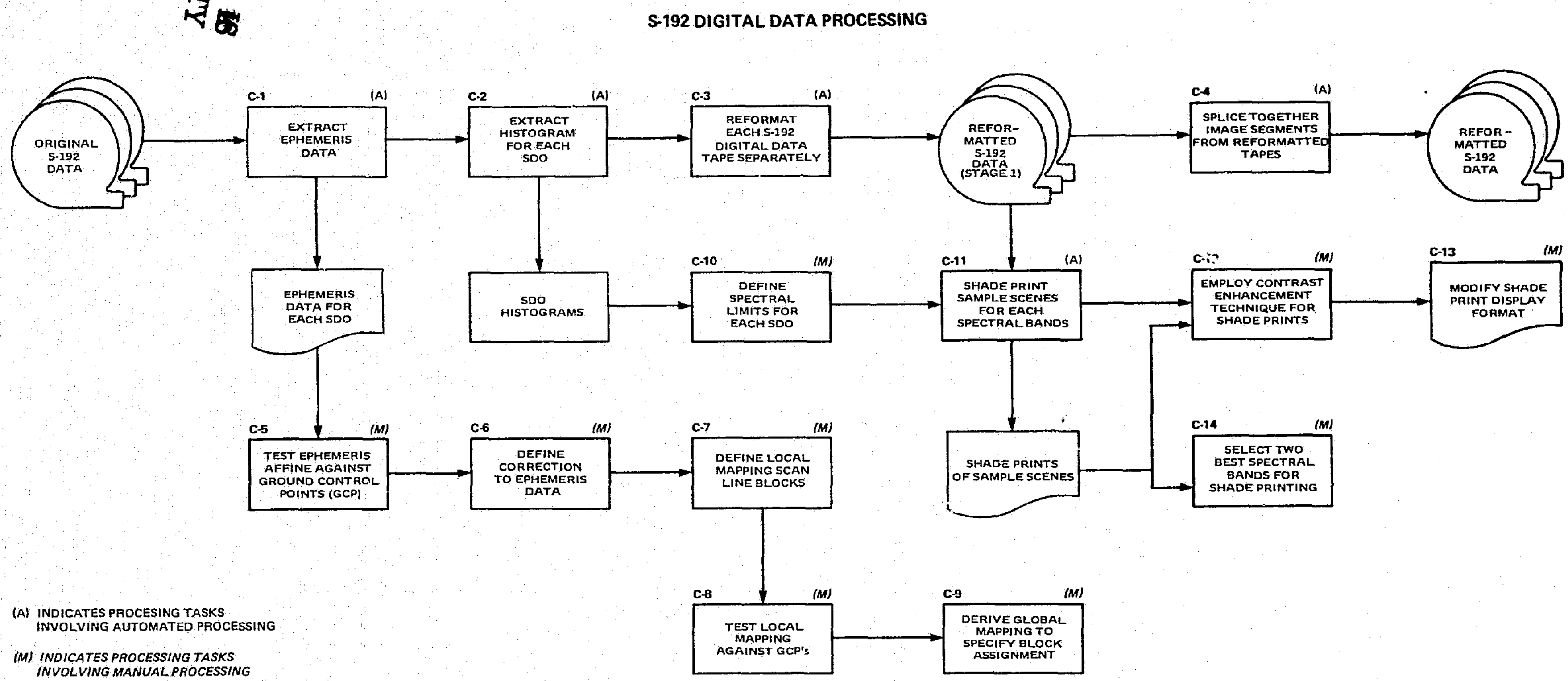


Figure 5. S-192 Digital Data Processing

1.4.1 <u>Data Tape Reformatting</u>

The SKYLAB S-192 data tapes received for analysis under this contract were line straightened data tapes containing all 22 Scientific Data Output (SDO) channels. The data tapes had been processed by NASA to remove systematic "noise" from the data. The format of these data tapes is thoroughly discussed in section 6.1 of NASA document PHO-TR543 entitled "Earth Resources Data Format Control Book." It is, however, useful to briefly describe the data tape format here.

The data tape format was an adaptation of the "Imagery Data Universal Tape Format" for line-straigntened data tapes. Because of the generality in design of the universal tape format, the first record on each tape was a header record used to describe contents and data structure of the following video data. The video data was segmented into "data sets." Each data set contained all of the video information from all channels present for a single scanline. The number and structure of the physical records within a data set depended on:

1. The type of video data (conical scanline, line-straightened, etc.)
2. The number of channels (SDO's) present.
3. The number of pixels per scanline.

The normal arrangement of pixels within a given scanline was by <u>channel</u>. The first physical record in each data set contained a block of ancillary data which provided information pertinent to that particular scanline.

The header record on each tape completely described the contents of the data format of the remainder of the data tape. Among other information, the header record specified:

1. Number of SDO channels present.
2. SDO channel numbers.
3. Spectral coverage for each SDO channel.
4. The number of bytes of video data per scan.
5. The number of physical records per data set.
6. The number of bytes of data per physical record.
7. The number of SDO channels in the first record of a data set.
8. The number of SDO channels in all other records of the data set.
9. Number of bytes in ancillary block.
10. Byte address of the first block of video data.
11. Byte address of the calibration data.
12. Number of bytes of calibration data.

The ancillary data block for each scanline included among other information a tape-line number, and the latitude and longitude of the first, center, and last pixel of that particular line. Each physical record on the data tape was 3,060 bytes long. If the data tape contained all 22 SDO channels, the maximum number of scanlines which could be stored on a single 2,400-foot 9-track tape was approximately 591. Multiple tapes were required to provide all of the data for the Washington, D.C.-Baltimore, Maryland area test site. Six data tapes were received for the SL/2 overpass, and seven tapes were received for the SL/3 overpass. The data tape line numbers in the ancillary data block were numbered sequentially throughout the overpass, so that the scan-line segments could be "spliced" together to form a continuous image for the entire test area. At one point in the SL/3 image, the scene segments from two consecutive tapes overlapped. This presented little problem; however, on the SL/2 image where there was a 103 scanline gap between two scene segments, blank data records had to be inserted into the SL/2 scene to compensate for this loss of data.

1.4.1.1 Original Data Reformatting

The planned utilization of the S-192 digital data required that all of the scanlines for the entire test area sampled by a particular SDO channel be arranged sequentially on a single computer tape. It was decided that the most efficient procedure for reformatting the data to this form was to perform a two-stage data reformat. The first stage data reformat rearranged each individual original S-192 data tape to the desired format. The second stage spliced the scene segments from all of the first stage reformatted data tapes to form a continuous image for each separate SDO channel of the entire area.

During the first stage data reformat process the ephemeris data from the ancillary data block was extracted from the data tapes and a printer listing of that data was produced (see processing block C-1 in Figure 5). In addition, each SDO channel was histogrammed and a histogram card deck was produced (see processing block C-2). These data were useful in the later phases of the S-192 digital data processing.

The format of the reformatted S-192 data tapes was designed to allow for convenient access and manipulation of the data while maintaining an efficient storage structure for the data. The physical records for the reformatted S-192 data tapes were all 4,992 bytes long. The first 32 bytes of data on each physical record contained header identification information. The remaining 4,960 bytes were segmented into four blocks each 1,240 bytes long. Each block contained video data from a single SDO channel. The following information was stored within the 32-byte header of each physical record:

1) SDO numbers for the four SDO scans contained on that record;
2) tape-line number for that particular scanline;
3) sequential line number (starting at one for the first record on the tape);
4) the total number of scans in that particular image segment;
5) the number of bytes of video data for each SDO (maximum of 1,240);
6) the input tape volume serial number.

The physical records were arranged on the reformatted tapes according to SDO channels. Suppose that there were 560 scanlines contained on a particular original S-192 data tape. The first 560 records on the reformatted tape would then contain all the video data for SDO numbers 1 through 4. The five-hundred sixty-first record was a separator record. Record numbers 562 through 1,121 would contain all the video data for SDO numbers 5 through 8. This sequence continued until all 22 SDO channels were transferred to the reformatted tape. Each reformatted tape contains six blocks of physical records. Each block contained output from four SDO channels. The last block of physical records of course contained only two SDO channel outputs.

1.4.1.2 Reformatted Data Splices

After each individual S-192 digital data tape had been reformatted, the second stage of the reformat sequence spliced the image segments contained on the separate tapes to a single image segment (see processing block C-4). The format of the spliced data tapes was the same as the format of the individual reformatted tapes except that each spliced digital tape contained only four SDO channels. Six spliced tapes were generated by the splicing program. The first tape contained all of the scanlines sequentially ordered from each of the separate reformatted tapes for SDO channels 1 through 4. The second tape contained SDO channels 5 through 8, and so on. The advantage of the reformatted spliced tapes was that all the scanlines from a particular channel were present on a single tape in sequential order. The shade print process was greatly simplified because of this data structure. All of the video data sampled by a particular SDO channel during a single mission overpass was contained on a single data tape. Once a particular test site location had been identified the shade print program could be assured of "finding" that particular area on a single data tape.

1.4.2 Development of the Affine Transformation

Essential to the photo interpretation effort was the accurate location of the test sites on the digital data. The accuracy of the latitude-longitude transformation to column-row coordinates determined the size of the shade prints required to present the data to the analyst. With an approximate size of 2-1/2 x 2-1/2 N. miles, the test site area would occupy about 80 x 80 pixels. In order to present the test site on one computer page, the maximum mapping error would be about 20 pixels for any point within the SL/2 or SL/3 areas. Using these constraints, a two-stage affine transformation algorithm was derived and tested.

1.4.2.1 Original Transformation Using Ephemeris Data

Initial mapping efforts involved comparing a globally defined affine transform to a transform developed by utilizing the ephemeris or footprint data which listed latitude-longitude for the first, center, and last pixel in each scan line.

The globally defined transformation was developed using the latitude-longitude of four ground control points (GCPs) determined from 7-1/2-minute U.S. Geological Survey quadrangle maps and matched to column-row indices indentified on shade prints of the S-192 digital data (see processing block C-5). Using three of these points, an exact affine transform was derived and then tested on the remaining GCP to estimate the magnitude of mapping error which might be expected from this process. The results indicated that, for this one point, errors would be in excess of 45 pixels. This result was expected in view of the non-linearities introduced by earth rotation during the long Skylab overpass.

Similar comparisons were made using "local" affine transforms derived from the ephemeris data. The SL/3 pass was divided into groups of 55 scan lines with each group represented by 1 scan line. For each such group, the affine coefficients were derived using the latitude-longitude values of that group and the previous group. The appropriate coefficients were used for each of the four GCPs and estimates of their column-row indices made. The resultant errors were of the same magnitude as the global affine, but the error appeared to be systematic.

1.4.2.2 Corrections to Ephemeris Data

In order to determine whether the ephemeris data could be corrected and subsequently used in the mapping program, the local affine transforms for each of the GCPs were inverted (see processing block C-6). That is, instead of mapping latitude-longitude into column-row, the transform was rederived from group ephemeris data to map column-row into latitude-longitude. Using the known column-row indices of the GCPs and their local transform coefficients, latitude and longitude coordinates of each GCP were estimated.

The resultant pairs of latitude-longitude measurements were input to a least-squares affine program which calculated the corrections necessary for the ephemeris data to be used in mapping. Let LAT be the actual latitude and LAT' be the latitude estimate derived from the local ephemeris affine transform. Similarly, define LON and LON'. The correction terms which were derived are:

$$LAT = 1.0225 \cdot LAT' - .0118855 \cdot LON' + .05774$$

$$LON = .0411892 \cdot LAT' + .977546 \cdot LON' + .118378$$

The calculated errors associated with the 4 GCPs were below 100m. Further tests were run which indicated that if the approximate row or block of the unknown point could be found, the local affine coefficients, based on corrected latitude-longitude of the ephemeris data, would yield results accurate to within 10 pixels.

1.4.2.3 Two-Stage Affine Transformation

The two-stage transform which evolved from the previous work utilized a global transform to estimate the row of the unknown point, and then refined that estimate using a local transform to obtain the column and row of the unknown point. The global transform was the affine transform obtained using all four GCPs - Tilghman Island, the mouth of the Blackwater River, Ferry Point, and Paton Island. This transform mapped the latitude and longitude of the unknown point into a row estimate. This row estimate was used to select the local block for which a set of coefficients existed which would complete the lat-lon to column-row transformation. The row estimator which was used is:

$$ROW = 111627 - (913.748 \cdot LAT) - (961.836 \cdot LON)$$

Subtracting the row bias (the row number of the first row in the pass minus 1), an estimate of the relative row number is available.

The entire SL/3 pass was broken into a sequence of blocks, each representing 220 rows of data (see processing block C-7). The selection of this size was primarily an estimate of the numbers of blocks needed to maintain accuracy without requiring large storage. The affine coefficients of the ephemeris data revealed an 11-line cycle in the latitude-longitude report. Specifically, latitude and longitude values repeated themselves at three- and four-line intervals. Thus, for the purposes of mapping, the same relative line within this cycle was selected to provide the ephemeris data. Every 55th row was used, and groups of 4 combined to derive the local affine coefficients for that block. After a particular block was selected by the global affine, the local coefficients for that block were used to estimate the column and row of the unknown point.

1.4.2.4 SL/2 Corrections

The procedure outlined in the previous section is dependent upon the existence of ephemeris data for the first and last pixels of a row. While this is satisfactory for the SL/3 pass, only the latitude-longitude of the first pixel exists for the SL/2 data used in this project. To bypass this limitation, the unknown point was mapped using the SL/3 two-stage transform. If the assumption can be made that the differences between the two passes can be explained by linear relationships, then it is possible to map an SL/2 point using the best SL/3 transform, and then map the column-row position on the SL/3 pass to the column-row coordinates of the SL/2 data. Using the equations shown below, this procedure was implemented by the mapping program.

$$COL_2 = 1.00333 \cdot COL_3 + .0081412 \cdot ROW_3 + 713.66$$

$$ROW_2 = -.0490658 \cdot COL_3 + .995038 \cdot ROW_3 - 670.139$$

1.4.3 Shade Prints of the Test Sites

The purpose of shade printing of each test was to present the digital data in a form that facilitated the matching of ground truth fields with the corresponding sets of pixels on the S-192 digital data. In order to do this successfully, three criteria must be be met:

1) The spectral bands displayed must differentiate the fields representing the land use codes,
2) each shade print should possess maximum contrast in the range of gray levels displayed, and
3) The geometric differences between the shade prints and the quadrangle maps had to be minimized.

1.4.3.1 Identification of Spectral Bands

Selection of the spectral bands to be used in the analysis was made by the project photointerpreters from a complete set of shade prints for all spectral bands covering several distinctive test sites (see processing block C-11). The initial shade prints were 125 X 125 matrices generated using a single overstrike character set and assignment keyed to the histogram for each spectral band (see processing block C-10). The test sites displayed were all from the SL/3 pass and included Tilghman Island and the Pimlico track area of Baltimore. Using these displays, it was possible for photointerpreters to select preferred spectral bands as well as suggest improvements in the display format which would aid in the interpretative effort (see processing block C-14).

Selection of the two spectral bands to be used on SL/3 sites was done by comparison of shade prints from all spectral bands for three sites: one urban, one rural, and one on the shore where definition between water and land becomes an important locator. Through this process spectral band 4 (.56 - .61 microns in the green-yellow portion of the spectrum) and spectral band 8 (.98 - 1.03 microns in the infrared portion of the spectrum) were chosen because in conjunction they appeared to give the best visual discrimination between both urban and non-urban categories. Spectral band 4 gives best visual discrimination between urban categories but offers poor discrimination between agricultural and forest categories. Spectral band 8, on the other hand, gives little visual discrimination between urban categories but good discrimination between agricultural field types and between forest and agriculture. Since spectral band 8 is in the infrared portion of the spectrum, it also provides some penetration of haze which spectral band 4 does not afford. Examples of shade prints produced these two spectral channels are shown later in this report (see section 1.5.1.2).

1.4.3.2 Contrast Enhancement

While the spectral band selection was based on a prespecified gray-level transformation, the operational character set assignment was determined by the gray-levels found within the areas to be displayed. The advantage to this approach was that the data for each test site would be displayed with maximum local contrast. A technique which is widely used and referred to in the literature as histogram equalization or histogram flattening (Eberlein, 1974) was employed. Essentially, this procedure defines a non-linear many-to-one mapping of input gray levels to output gray levels or print characters. The net effect is to spread the input data across the whole dynamic range of the output device. Figure 6 shows the affect of applying the histogram equalization technique to a typical input histogram. Note that the output histogram has been inverted left to right. This inversion was necessary in order to produce a positive rather than negative photo product.

The assignment of output gray levels or print characters was a function of the frequency of occurrence of each input gray level and of the number of output values available for both SL/2 and SL/3 test sites. It was desirable to limit the input gray levels used to determine the mapping. The Tilghman Island test site shade print, for example, contained 60% water data points. The technique described above would normally use these data points in the mapping calculation despite the fact that they were all "noise" to the photointerpreter. The resulting effect on the data of interest would be to greatly reduce the visual contrast in the image. To circument this problem, each spectral band was evaluated to determine a gray level interval within which the land use data

CHANNEL 12 BAND 11 1.55 - 1.75 um MID I. R.

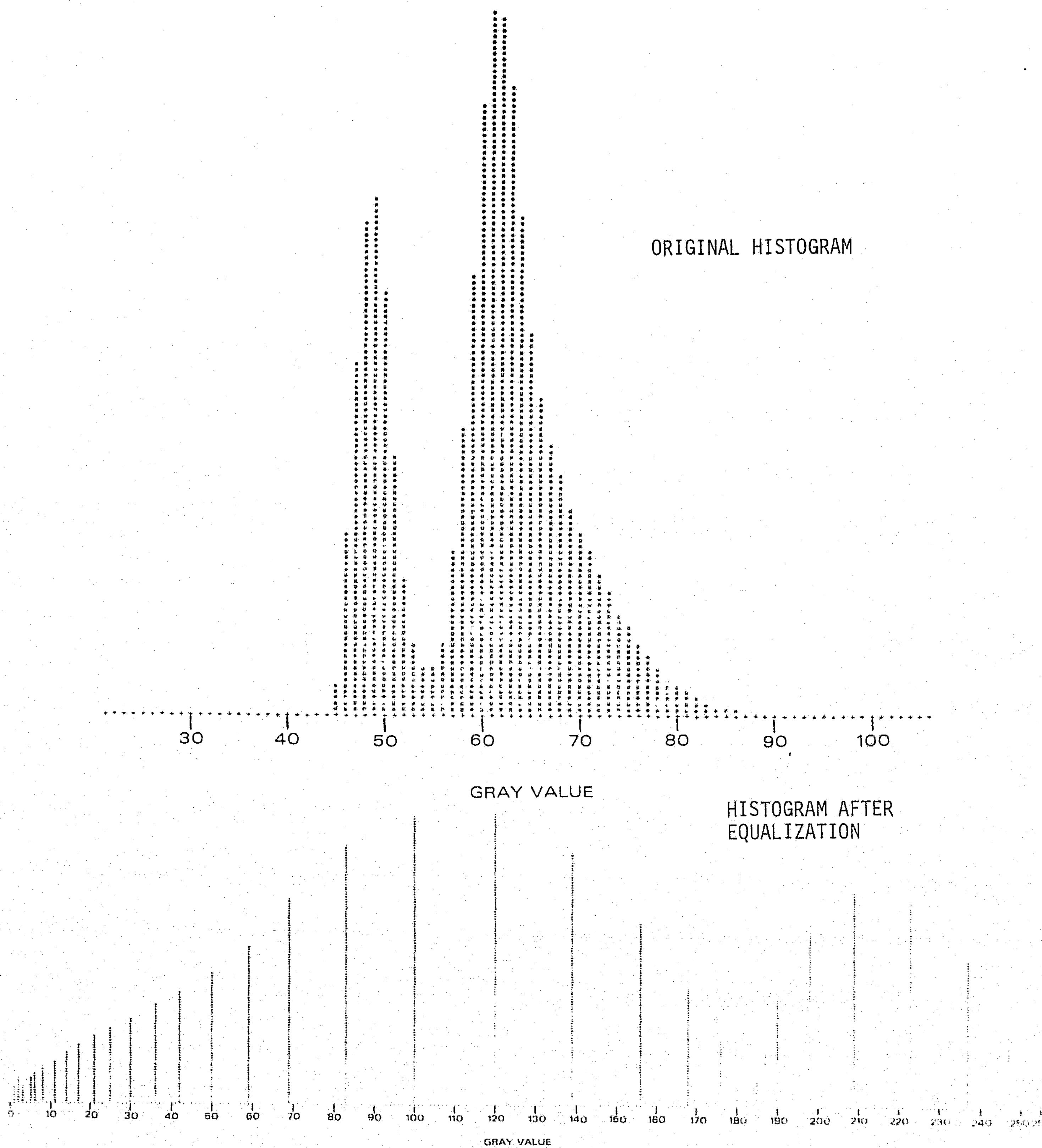


Figure 6. Histogram Equalization Technique

of interest would normally be found. The tails of each histogram were clipped and only the gray level interval of interest was contrast enhanced (see processing block C-12).

1.4.3.3 Geometric Matching

To aid in correlating the ground truth data with the digital data, shade prints were presented in a format which matched the U.S. Geological Survey 7-1/2 minute quadrangle maps used by the photointerpreters. A physical constraint was introduced by the use of the line printer. The printer produces ten columns per inch and six or eight rows per inch, thus, a 125 X 125 matrix which represented a square land area, would result in a print which would be longer than it was wide, making the photointerpretive effort more complex. To bypass this problem every fifth line was dropped, thus when printed at 8 lines per inch a 125 X 125 matrix would be displayed in a square measuring 12.5 inches on side. This represents a scale of about 1:31,500 compared to the 7-1/2' U.S. Geological Survey map scale of 1:24,000. To further match the two data sources, a transparency of the shade print was made and enlarged photographically by a factor of 1.3. Consequently, transparency can be used as a direct overlay to the quadrangle maps. Additionally, the four corners of the test site were numbered on the shade prints to further aid in locating the data (see processing block C-13).

1.5 Ground Truth - Digital Data Merger

The culmination of the data processing tasks described in the previous three sections was accomplished when the ground truth data and the S-192 digital data were integrated to form the principal data sets used in the analysis phase of the investigation. The ground truth-digital data merger, shown as processing block D on Figure 1, will be described in this section.

The processing tasks performed at this step were involved and highly interrelated. Figure 7 shows an expanded task flow diagram for the processing tasks performed at this step. The principal input data to this step included:

1) the reformatted S-192 digital data tapes (Section 1.4),
2) the test site data base (Section 1.3), and
3) the ground truth data (Section 1.2).

The output products of this processing task included two data files. The first file contained the 13-channel spectral signatures for all of the fields identified from ground truth data. The second file contained digital field boundary maps of the test sites as well as 13-channel video data extracted from the reformatted S-192 data tapes. These two data files provided all of the information needed to perform the analysis phases of the investigation.

1.5.1 Identification of Ground Truth Fields

The first principal task performed at this step required the accurate identification of ground truth fields on the video data. This task was of critical importance to the analysis that followed. The identification of ground truth field boundaries

GROUND TRUTH - DIGITAL DATA MERGER

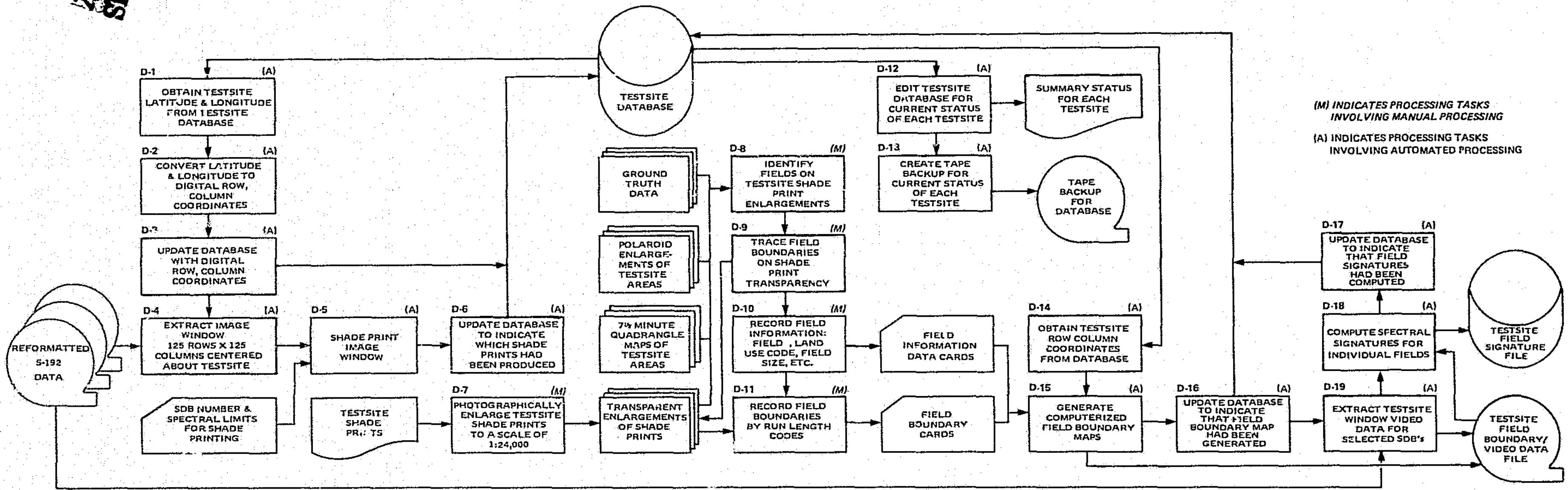


Figure 7. Ground Truth - Digital Data Merger

had to be accurate to within a pixel. If the spectral signatures were calculated for fields in which the field boundaries were incorrectly specified, the resulting signature errors could bias the results obtained from the signature analysis.

After an investigation of several alternative methods for identifying ground truth field boundaries, it was decided that the most accurate procedure would be through the use of computer printer gray maps or shade prints produced from the digital data. As was discussed in Section 1.4.3, it was possible to produce shade prints which were geometrically rectified. Enlarged transparencies could then be produced from the digital shade prints at a scale of 1:24,000. These transparencies could be directly overlayed on the U.S. Geological Survey 7-1/2-minute quadrangle maps, and manually registered to these maps.

1.5.1.1 Test Site Shade Print Production

Processing blocks D-1 through D-7 on Figure 7 show the processing steps required to produce shade prints of the test site areas. The latitude and longitude for a particular test site was retrieved from the test site data base. Using the affine transformation developed to relate latitude and longitude to the corresponding digital row, column coordinates, the approximate location of the individual test site areas on the video data was determined. A digital window 125 columns wide and 125 rows long centered about the test site was identified. The row, column coordinates of the upper lefthand corner of the 125 x 125 window were entered into the test site data base for future use (see processing block D-3). Card input data was used to specify which spectral channel was to be shade printed. The exact location of the 125 x 125 video window on the reformatted S-192 data tapes requested was calculated and the appropriate S-192 digital data tape was searched to locate that window. The header information on each record of the S-192 reformatted data tapes was used to verify that the exact location had been found. At that point, the processing program extracted the 125 x 125 video window from the digital data tape, and copied the window onto temporary direct access storage. During the copying process, a histogram of the video information was calculated. The extracted 125 x 125 video window and the histogram of that window were passed to the shade print program described in Section 1.4.3. The shade print program used the histogram to develop a gray value transformation designed to provide maximum contrast on the final shade print. The shade print was then produced by transforming the gray values given in the 125 x 125 video window. The final step in the procedure was to update the test site data base to indicate which shade print had been produced (processing block D-6). The updating procedure served two purposes. First, it provided a permanent record of which shade prints had been produced for each test site. Second, and equally important, it avoided the problem of reproducing shade prints which had already been generated. The shade print processing program automatically checked the test site data base to determine whether a particular shade print had already been produced for a specific test site. If no shade print had been produced, then it would generate one. If that particular shade print already existed, then it would go on and check the next test site in the file. This simplified the operation greatly. The user only had to specify which spectral channel he was interested in and how many shade prints should be generated within a specific run. The processing program would determine which test sites should be shade printed. Once the test site shade prints were generated they were photographically enlarged (see processing block D-7). The next step in the ground truth-digital data merger was to identify the ground truth field boundaries on the enlarged transparencies.

1.5.1.2 Photointerpretation of Shade Prints

Photointerpretation of the test site shade prints was a long and involved process. First, the photointerpreters had to manually register the transparent enlargements of the test site shade prints to the U.S. Geological Survey 7-1/2-minute quadrangle maps. Even though the photointerpreters had shade print enlargements from two spectral bands to work from, this process was frequently difficult and occasionally impossible. Test site areas which contained easily identifiable characteristics such as major roadway intersections or streams and rivers were easily registered. Test sites which did not contain these type of distinguishable characteristic were much more difficult to register. It soon became apparent that it would be necessary to distinguish between test site fields which could be delineated with a high degree of confidence from those fields which were identified with a lesser degree of confidence. A field quality rating (FQR) procedure was established for this purpose. Each field identified on the shade prints was given a rating between 1 and 5. The field quality rating was designed to incorporate various aspects of the individual fields including field size, visual uniformity of gray values in the area covered by the field, and confidence in the delineation of the field boundaries. The following description of the various ratings was given to the photointerpreters:

FQR 5 - Excellent, large fields (greater than 25 pixels) visually uniform in grey tone, with accurate delineation of field boundaries visible on shade prints.

FQR 4 - Very good, same as 5 except for field size constraints.

FQR 3 - Average, size, consistency, and accuracy of location may limit the usefulness of the field.

FQR 2 - Poor quality, field may not provide accurate spectral description of particular land use.

FQR 1 - Not usable.

Initially, it was intended that each field for which ground truth data was available would be identified on the shade prints. It was quickly found that this was a very long and time-consuming process. Additionally, the advantage of identifying fields with quality ratings 1 and 2 was questionable. It was decided that all fields which had a quality rating of 4 or 5 would be selected, and the lower field quality rating fields would only be selected if they represented land use areas generally not available with higher quality ratings.

Using the Polaroid prints of the test sites in conjunction with shade print enlargements and 7-1/2-minute U.S. Geological Survey Quadrangle Maps, an experienced photointerpreter could, under optimal conditions, select all of the 4 and 5 quality fields within a test site in 20 to 25 minutes.

In addition to the Polaroid enlargements, the photointerpreter had access to high-altitude color-infrared photography obtained on June 14, 1973. This photography was an invaluable aid to the photointerpreters for verification of the ground truth data and selection of land use categories not contained in the ground truth.

The photographic data was particularly useful for types of wetlands, forest, as well as classification of density and tree cover in residential areas.

As the fields were located on the transparent shade print enlargements they were outlined directly on the enlargements. The ground truth forms were consulted to determine the specific land use code assigned to that field by the ground truth collection team. For each field the photointerpreter coded on a computer form the following information (See processing block D-10):

1) Test site number
2) Field number
3) Land use code
4) Estimate of field size in pixels
5) The field quality rating

The estimate of the field size was used as a check to verify that the field information was associated with the correct field area. This information (field information cards on Figure 7) was then keypunched and verified for use in subsequent steps in the process. Figures 8 and 9 show the shade prints produced for test site 168. Figure 8 shows spectral band number 4 (SDO number 3) while Figure 9 shows spectral band number 8 (SDO number 19). The fields identified for test site number 168 are shown on Figure 9.

1.5.1.3 Recording Field Boundary Data

After the field boundaries had been outlined on the shade print enlargements, it was necessary to transfer this information to a computer compatible form. This task required considerable manual processing.

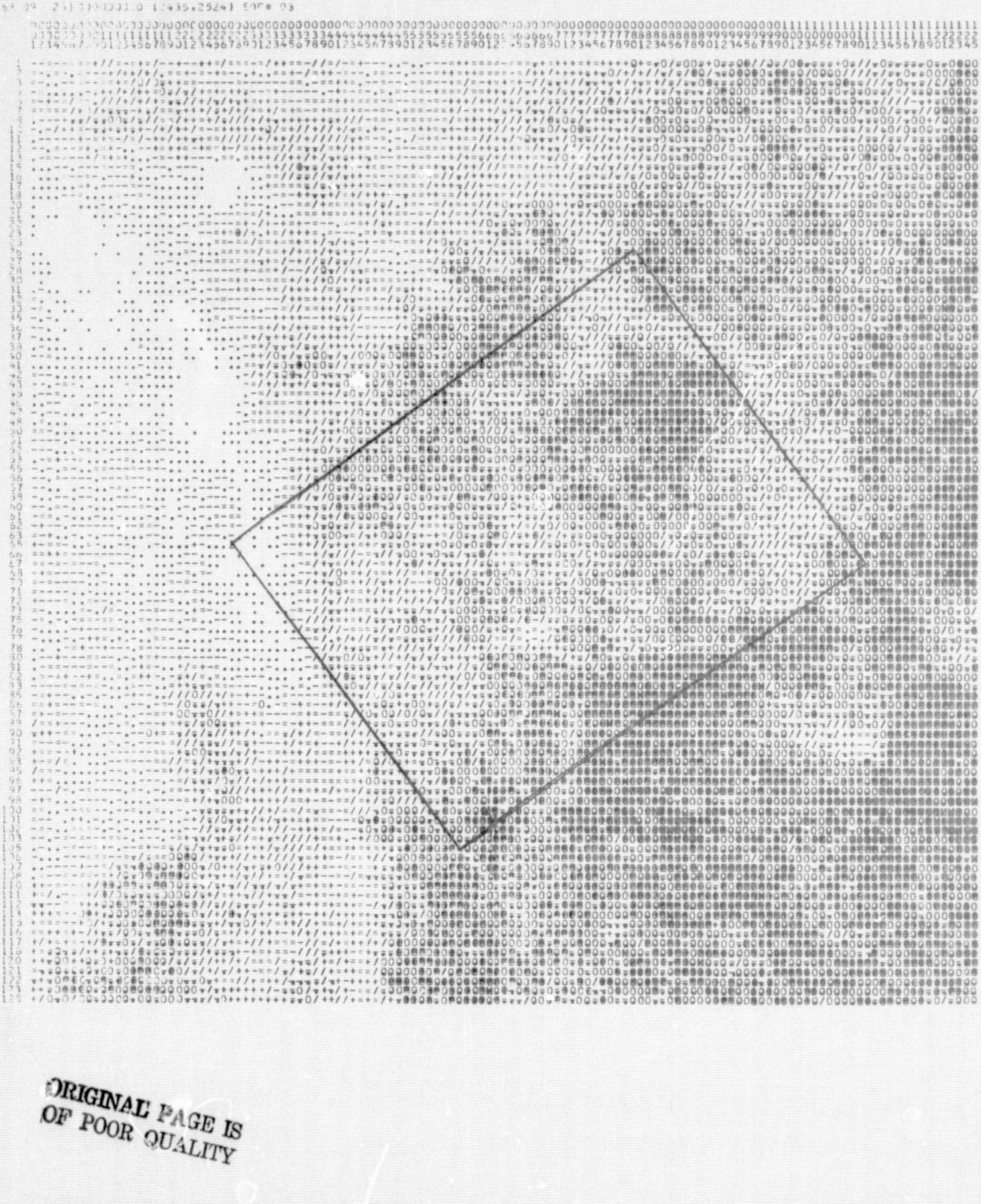


Figure 8. Shade Print of Test Site 168, Spectral Band 4

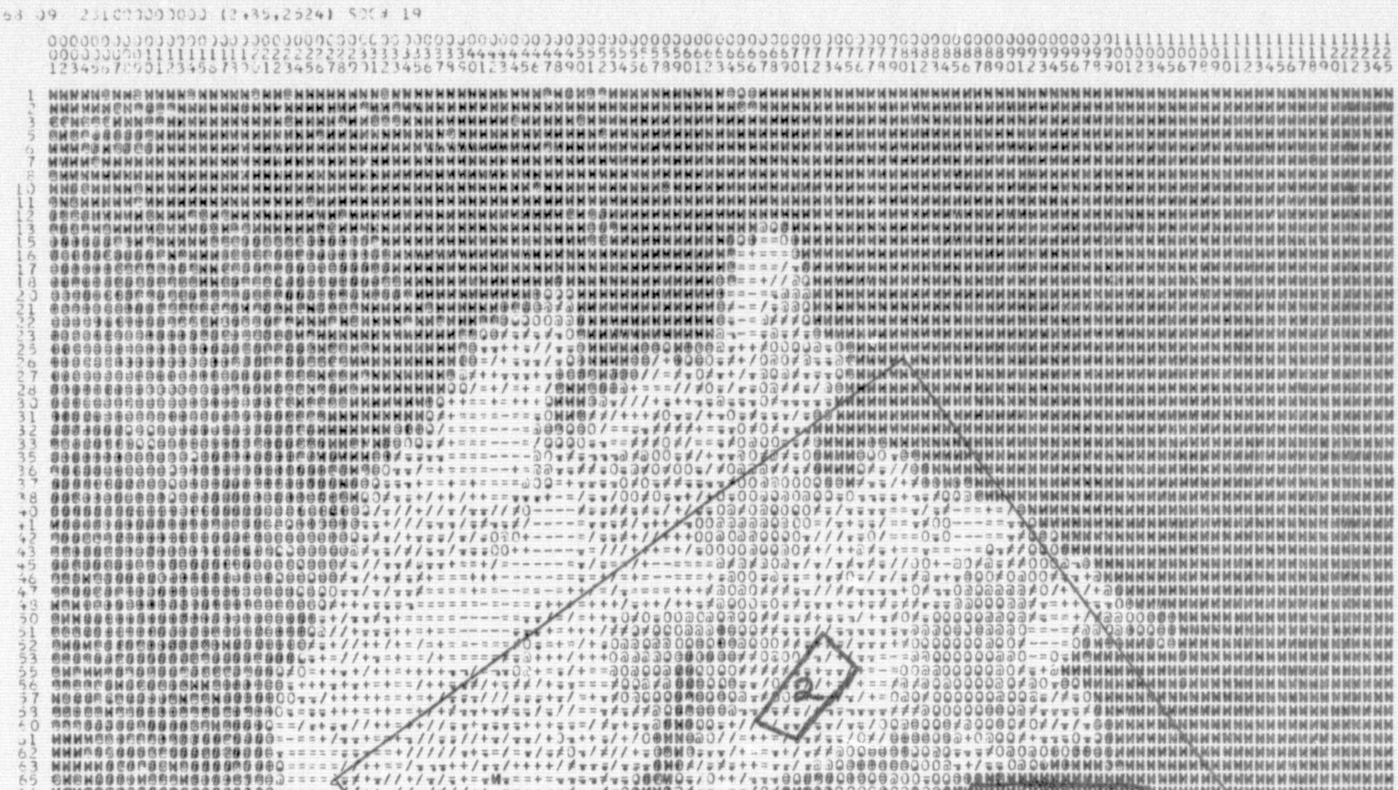


Figure 9. Shade Print of Test Site 168, Spectral Band 8

Within each shade print, the ground truth fields which had been delineated covered an area much smaller than the full 125 x 125 window. In order to reduce subsequent computer storage requirements, it was decided that a smaller sub-window 80 columns wide and 80 rows long, centered to include all of the test site fields, would be used in the subsequent analysis. Using the shade print transparencies, the 80 x 80 sub-window was positioned and the row and column offsets of the upper left hand corner of the 80 x 80 sub-window relative to the 125 x 125 window were recorded. The individual pixels assigned to each field were then recorded by run-length codes. Starting at the upper left hand corner of a specific field, the row, column coordinates (relative to the 125 x 125 window) were recorded. That information was followed by the number of columns within that specific row which were assigned to the same field. The pixels assigned to the same field in the next row were recorded by specifying a column offset and row length. A column offset of "0" implied that the left-hand boundary of the field in row "i" fell in the same column as the left-hand boundary in row "i-1". Negative offsets implied that the boundary moved to the left while positive offsets implied that the boundary moved to the right. The offsets were always recorded relative to the row immediately preceding the current row. This type of information coding allowed arbitrarily shaped fields to be recorded with a minimum of effort. In some situations where fields curled back on themselves (such as a U shaped field), it was necessary to record the field boundary as though there were actually two distinct fields assigned to the same field number. The following information was recorded for each field:

(1) Test site number

(2) Field number

(3) Row, column offset of upper left-hand corner

(4) Column offset and row length pairs for each row having pixels assigned to that field

All of this information was keypunched and verified (see field boundary data cards on figure 7). An experienced "boundary coder" could completely process an average test site in about 30 minutes. The next step in the process was to convert the run-length field boundary data into computer field boundary maps.

1.5.2 Test Site Field Boundary/Video Data File

The next step in the ground truth-digital data merger process, was to generate a data file containing the following information about each test site:

(1) Field boundary maps

(2) Land use boundary maps

(3) Thirteen channel spectral video data windows

This single file would then contain all the information required to perform the subsequent analysis. The generation of this file required four data inputs:

(1) The test site data base

(2) The reformatted S-192 digital data tapes

(3) The field information data cards

(4) The field boundary data cards

The processing required to generate the file will be described in this section.

1.5.2.1 Test Site Boundary/Video Data File Structure

The test site field boundary/video data file once created, provided all of the critical information required by the Skylab data analysis tasks. The structure of the file was relatively simple. Each record in the file was 1,225 bytes long and there were 80 records per test site. The first 25 bytes of each record contained identification information such as:

(1) Test site number,

(2) County number,

(3) Skylab mission number,

(4) SDO numbers for the video windows contained in the file.

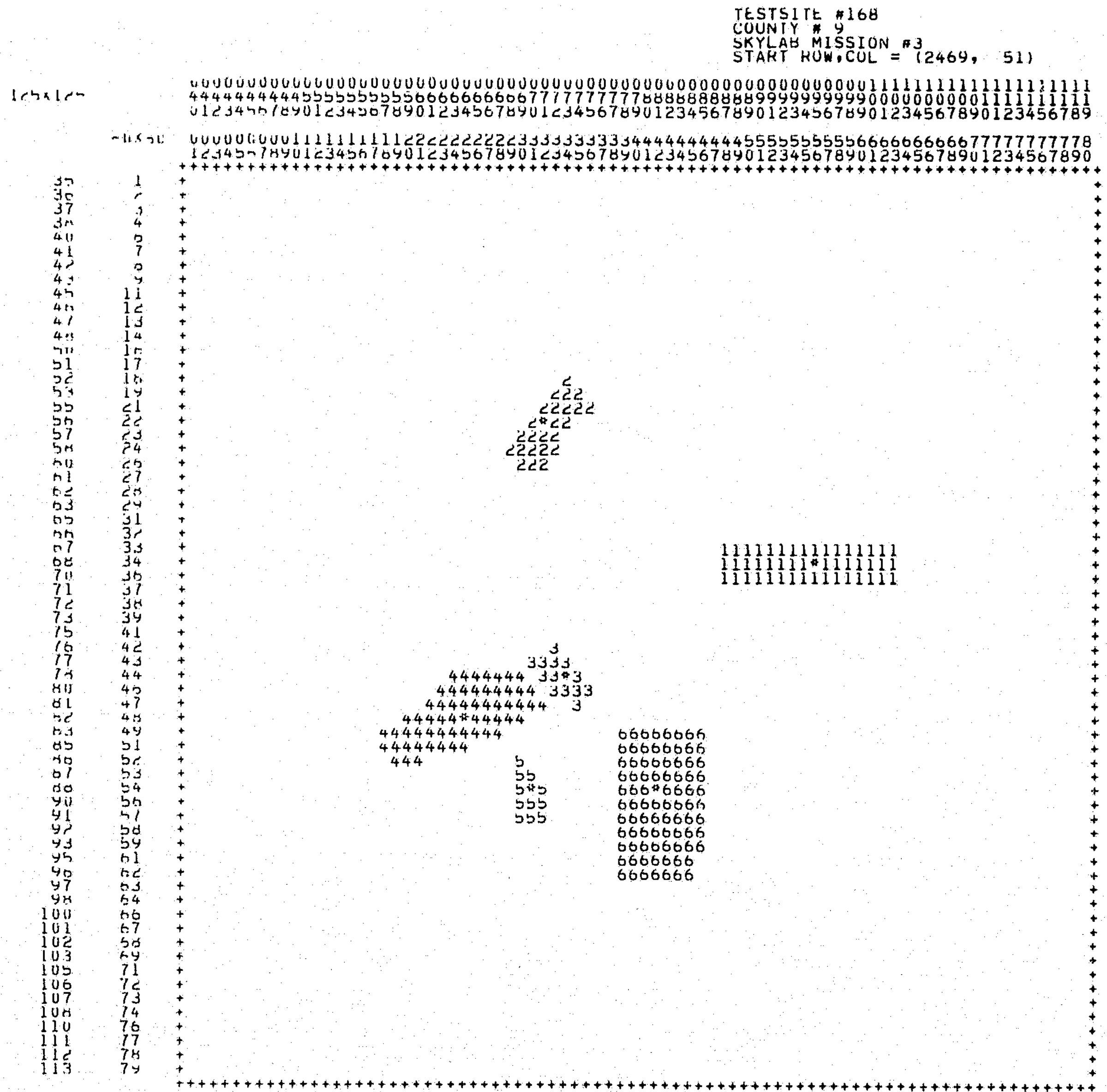
The remaining 1,200 bytes of each record were segmented into fifteen corresponding to a particular test site could be thought of as containing 15 windows each 80 rows (80 records) by 80 columns (80 bytes per block). The first 80 x 80 window, contained the field boundary map. Each data byte in the window corresponded to a pixel in the original data tape. If the value of the data byte was zero, then that particular pixel had not been assigned to a ground truth field. If the value of the data byte was non-zero then that pixel had been assigned to a field and the value of the data bite was the field number. The next 80 x 80 window contained a land use map. The land use map was identical to the field boundary map except that the non-zero entries specified the land use code assigned to that pixel by the ground truth collection teams. The remaining thirteen windows contained the spectral data for the thirteen spectral bands detected by the S-192 Multispectral Scanner. It should be noted that within each individual record in this data file, the user not only had access to the gray values recorded by all thirteen spectral channels of the S-192 Multispectral Scanner, he also knew which pixels were assigned to fields, what the field numbers were, and the land use categories the pixels belonged to. This data file was a very powerful tool in the analysis which followed.

1.5.2.2 <u>Generation of Field Boundary Maps</u>

The test site field boundary/video data file was constructed in two steps. The first step in the construction of this file was the generation of the field boundary and land use boundary maps. The 13 channel video data was then added to the file in the second step. The file was initially built on a random access, direct storage device. This access method simplified adding the video data to the file since the file could be updated in place. Once the file had been generated it was copied onto a computer tape for storage and future use.

The generation of the test site field boundary maps (see processing block D-15) required accessing the test site data base to determine the row, column coordinates of the upper lefthand corner of the 125 x 125 test site windows. Using the 80 x 80 subwindow offset coordinates the row, column coordinates of the subwindow were calculated. The field information cards were then used to reconstruct the field boundary maps in computer-compatible form. It is important to note that during the reconstruction process record was kept of the position of each line which had been deleted during the shade print geometric rectification process. As the field boundary maps were processed the missing lines on the shade prints were reinserted.

In order to verify the correctness of the field boundary maps a printer plot of each test site field boundary map was generated. The printer plots deleted the same lines which had been deleted from the shade prints. It was a simple procedure to photographically enlarge the printer field boundary plots by the same factor used to enlarge the original shade prints. By overlaying the enlarged printer plots on the shade



C.2

Figure 10. Field Boundary Map for Test Site 168

print enlargements the accuracy of the field boundary maps could be rapidly checked.

As a further check the test site field boundary map program calculated the actual size in pixels of each field as well as the number of pixels within each field visible on the shade prints. These two field sizes were, of course, different because of the lines deleted from the shade prints. By comparing the field size estimates made by the photointerpreters with the actual size of the fields as seen on the shade prints it was possible to quickly verify that the correct field numbers had been assigned to the fields in the field boundary maps. Figure 10 shows the field boundary map generated for test site 168 (see Figure 9 for shade print).

Once the field boundary maps had been generated, the field information cards were used to determine the land use categories assigned to each field in the map. The land use boundary maps were generated by copying the field boundary maps while replacing the field numbers by land use codes.

1.5.2.3 Video Data Extraction

The final step in the generation of the test site field boundary/video data file was the extraction of the 80 x 80 subwindows of test site video data from the reformatted S-192 digital data tapes.

The reformatted S-192 digital data tapes contained all 22 Scientific Data Output (SDO) channels. Since the test site field boundary/video data file had been designed to store only 13 subwindows of video data (see Section 1.5.2.1) it was necessary to select a set of 13 SDO's to be extracted from the reformatted digital data tape. The first step in the analysis of the SKYLAB data was designed to analytically determine the best set of 13 SDO's for use in the other phases of the analysis. This selection process required analyzing all 22 SDO's. This difficulty was overcome by generating three test site field boundary/video data files for some of the test sites. The first two files were generated for use in the SDO selection process. These two files did not contain a complete set of 13 spectral bands. After the SDO selection process was completed a third file, which did contain a complete set of 13 spectral bands as determined by the selection process, was generated for all of the test sites (see Section 2.1 for a complete description of the selection process).

1.5.3 Calculation of Test Site Field Signatures

The final task of the ground truth-digital data merger was the calculation of ground truth field signatures (see processing block D-18). The spectral field signatures included:

1) the number of pixels in a field

2) the mean grey values in each of the 13 spectral bands

3) the 13 x 13 covariance matrix

The field signatures calculated from video data provided a concise statistical representation of the spectral properties of the ground truth fields.

1.5.3.1 Signature Calculation

The test site field boundary/video data file was used to calculate the spectral field signatures. Each block of 80 records in the field boundary/video data file corresponding to a particular test site were read in and the field boundary maps were searched for pixels assigned to ground truth fields. As individual pixels assigned to a specific field were located, the corresponding gray values were included in the signature calculation for that field. After processing all of the records for a specific test site, the spectral field signatures for the fields within that test site were printed. In the spectral signature editing process the correlation matrix was actually printed rather than the covariance matrix. Table 3 shows the spectral field signatures calculated for two fields from test site 168.

1.5.3.2 Test Site Signature File

After calculating the spectral field signatures for all of the fields from a specific test site

TEST SITE: 168 FIELD: 1

COUNTY: 9 WINDOW ROW: 2469 LAND USE CODE: 5110 CENTER ROW: 2502.50 NUMBER OF BANDS: 13
SL MISSION: 3 COL: 51 QUALITY RATING: 5 COL: 105.50 NUMBER OF PIXELS: 64.

MEAN VALUES	STD. DEV.	SDO NUMBERS	CORRELATION MATRIX 22	18	1	3	5	7	9	19	20	17	11	14	21
112.9	6.1	22	1.000	-0.022	0.024	0.131	0.148	0.083	0.009	0.066	0.056	0.169	-0.047	0.070	0.050
114.4	25.6	18	-0.022	1.000	0.011	0.004	-0.071	0.069	0.182	0.021	0.040	0.204	0.068	-0.044	-0.083
70.6	4.3	1	0.024	0.011	1.000	0.674	0.460	0.673	0.567	0.489	0.424	0.449	0.619	0.409	-0.156
40.6	4.2	3	0.131	0.004	0.674	1.000	0.568	0.480	0.585	0.373	0.430	0.446	0.542	0.425	-0.210
38.6	6.5	5	0.148	-0.071	0.460	0.568	1.000	0.203	0.429	0.155	0.114	0.133	0.355	0.314	-0.045
46.6	5.7	7	0.083	0.069	0.673	0.480	0.203	1.000	0.530	0.555	0.543	0.469	0.542	0.324	-0.081
25.8	4.3	9	0.009	0.182	0.567	0.585	0.429	0.530	1.000	0.488	0.543	0.483	0.721	0.451	-0.259
21.7	6.4	19	0.066	0.021	0.489	0.373	0.155	0.555	0.488	1.000	0.703	0.675	0.621	0.321	-0.060
26.5	7.0	20	0.056	0.040	0.424	0.430	0.114	0.543	0.543	0.703	1.000	0.553	0.563	0.298	-0.090
18.5	6.4	17	0.169	0.204	0.449	0.446	0.133	0.469	0.483	0.675	0.553	1.000	0.498	0.342	-0.271
11.5	3.8	11	-0.047	0.068	0.619	0.542	0.355	0.542	0.721	0.621	0.563	0.498	1.000	0.528	-0.087
7.8	4.5	14	0.070	-0.044	0.409	0.425	0.314	0.324	0.451	0.321	0.298	0.342	0.528	1.000	-0.107
149.5	4.9	21	0.050	-0.083	-0.156	-0.210	-0.045	-0.081	-0.259	-0.060	-0.090	-0.271	-0.087	-0.107	1.000

TEST SITE: 168 FIELD: 2

COUNTY: 9 WINDOW ROW: 2469 LAND USE CODE: 4200 CENTER ROW: 2490.36 NUMBER OF BANDS: 13
SL MISSION: 3 COL: 51 QUALITY RATING: 5 COL: 82.42 NUMBER OF PIXELS: 33.

MEAN VALUES	STD. DEV.	SDO NUMBERS	CORRELATION MATRIX 22	18	1	3	5	7	9	19	20	17	11	14	21
118.8	6.3	22	1.000	0.046	-0.066	0.183	0.035	0.182	0.063	0.079	0.225	0.293	0.280	0.424	-0.078
111.3	14.6	18	0.046	1.000	0.249	0.157	0.183	-0.157	-0.296	0.127	0.003	0.162	-0.355	-0.063	0.267
68.2	2.9	1	-0.066	0.249	1.000	0.308	0.325	-0.159	-0.034	-0.127	-0.087	0.004	0.040	0.006	0.040
38.2	2.6	3	0.183	0.157	0.308	1.000	0.448	0.266	0.375	0.251	0.207	0.290	0.429	-0.020	0.012
37.0	5.2	5	0.035	0.183	0.325	0.448	1.000	0.219	0.348	0.258	0.381	0.401	0.320	0.250	0.125
75.6	4.0	7	0.182	-0.157	-0.159	0.266	0.219	1.000	0.479	0.606	0.572	0.485	0.506	0.307	0.418
81.3	4.1	9	0.063	-0.296	-0.034	0.375	0.348	0.479	1.000	0.270	0.339	0.229	0.595	0.138	0.064
92.9	6.2	19	0.079	0.127	-0.127	0.251	0.258	0.606	0.270	1.000	0.693	0.621	0.448	0.030	0.076
87.6	5.5	20	0.225	0.003	-0.087	0.207	0.381	0.572	0.339	0.693	1.000	0.640	0.675	0.212	-0.224
78.9	6.1	17	0.293	0.162	0.004	0.290	0.401	0.485	0.229	0.621	0.640	1.000	0.385	0.244	-0.016
35.6	3.0	11	0.280	-0.355	0.040	0.429	0.320	0.506	0.595	0.448	0.675	0.385	1.000	0.163	-0.253
16.8	4.2	14	0.424	-0.063	0.006	-0.020	0.250	0.307	0.138	0.030	0.212	0.244	0.163	1.000	0.138
151.3	5.4	21	-0.078	0.267	0.040	0.012	0.125	0.418	0.064	0.076	-0.224	-0.016	-0.253	0.138	1.000

Table 3. Spectral Signatures for two Fields from Test Site 168

the signatures were added to a spectral signature data file. As the spectral signature data file was generated the test site information cards were accessed to determine the field quality ratings assigned to each field. Upon completion the spectral signature file contained the following information about each field:

1. Test Site Number
2. County Number
3. SKYLAB Mission Number
4. Field Number
5. Land Use Code
6. Field Quality Rating
7. Number of Spectral Channels included in Signature
8. SDO Numbers of those Channels
9. Number of pixels in the field
10. Mean Gray Values for each Channel
11. Symmetric Form of the Covariance Matrix.

The test site data base was then updated to indicate which test sites had been processed to calculate spectral signatures (the processing block D-19). The generation of the test site signature file was the last step prior to beginning the analysis of the SKYLAB data.

2.0 DATA ANALYSIS

The data processing developed for this investigation and described in Section 1.0, provided the necessary data bases required to study the fundamental questions posed in the Executive Summary. In particular, the test site signature file (See Section 1.5.3) provided the information necessary to investigate questions related to:

1) the increased spectral coverage available with the S-192 Skylab Multispectral Scanner
2) identification of optical spectral bands for discriminating various land use categories
3) relationships between classes of data which are spectrally discriminable and land use categories of value to the land use planner.

The test site window data file (See Section 1.5.2) provided the necessary data to perform multispectral classification using the Skylab S-192 Multispectral Scanner data. In addition to addressing the questions related to the accuracy levels achievable when performing multispectral classification with the S-192 Multispectral Scanner data, it was also possible to comment on questions related to:

1) boundary delineation by multispectral classification
2) scan line misregistration between spectral bands
3) the effect of field quality on reported multispectral classification results.

The results of the analysis performed to investigate these questions along with the techniques and methodologies employed, will be discussed in this Section.

2.1 Discriminant Analysis

The step-wise discriminant analysis techniques discussed in this section were employed to address two distinctly different types of questions. The first question was totally operational in nature; the second question was more fundamentally related to the discipline area of land use planning and resource management.

As discussed earlier, the S-192 digital data included data from 13 separate spectral bands. The digital data tapes, however, contained output from 22 Scientific Data Output (SDO) channels. The first task in the analysis of the S-192 digital data was to select a single set of 13 SDO channels providing complete spectral coverage of all bands. These 13 channels would then be used exclusively in the other analysis tasks. The selection process was designed to identify the "least noisy" channel for those spectral bands sampled by more than one SDO channel. Since five of the spectral bands were sampled by only one SDO channel, there was no apparent reason to include those channels in the preliminary analysis.

After completing the SDO channel selection process, the second question could be addressed. Using the 13 SDO channels identified in the first analysis task, the step-wise discriminant analysis technique was employed to determine the best set of spectral bands for discriminating various land use categories. The differences between the spectral characteristics of various land use categories were analyzed to determine which spectral bands provided the greatest information to discriminate one category from another.

The step-wise discriminant analysis technique and the methods used to employ it in this investigation are discussed first. The results of the analysis performed are then presented and followed by a discussion of the implications of these results. Finally, the results obtained here are compared with other previously reported work.

2.1.1 Analysis Technique

Step-wise discriminant analysis procedures are widely reported in the literature (see T.W. Anderson, 1958, and C.R. Rao, 1965), and therefore will only be briefly summarized here. The analysis program used in this investigation was a modified form of the step-wise discriminant analysis program available as part of the Biomedical Computer Programs (BMD) package (Dixon, 1973).

In this application, the subject variables of the analysis were the 13 spectral bands. Data groups were defined as combinations of spectral signatures from various land use categories. The spectral signature combinations varied depending on the type of analysis being performed. Suppose, for example, that one wished to determine which single spectral band provided the greatest discrimination between urban and agricultural areas. In this case, two groups would be defined: a first group "U" would be defined to represent all pixels in the data base assigned to the class "urban;" and a second group "A" would be defined as all pixels assigned to the class "agriculture." The BMD program, which normally would expect unprocessed input data (actual gray values from 13 spectral channels), was modified to accept spectral signatures instead. Through these modifications, it was possible to specify, for each signature to be included in the analysis, its group number, test site number, and field number. The program would then access the test site signature file, find that particular field signature, and add it to the composite signature formed from that particular group. After specifying all of the field signatures to be used in a particular analysis, the composite group signatures were identical to those that would have been calculated directly from the corresponding pixel gray values. This modification greatly reduced the amount of processing necessary to perform the stepwise discriminant analysis.

The composite group signatures were used to calculate a within-group cross-product matrix, as well as a total cross-product matrix. The total cross-product matrix was directly proportional to the variance-covariance matrix for all of the data treated as a single data set. The within-group cross-product matrix was directly proportional to the within-group variance-covariance matrix, that is, the variance-covariance matrix obtained by performing a weighted sum of each <u>group</u> variance-covariance matrix. Note that a specific spectral band provides a good discriminant between groups if the total variance for all of the data (diagonal element of the total cross-product matrix) is much greater than the variance obtained by treating the data in groups (diagonal element of the within-group cross-product matrix). Let W be the within-group cross-product matrix with elements (w_{ij}), and T be the total cross-product matrix with elements (t_{ij}). Suppose that a total of "n" data points are used to define "g" distinct groups. Then, if it is assumed that the spectral gray values are normally distributed, the statistics w_{jj} and t_{jj} obey Chi-Squared distributions with n-g and n-1 degrees of freedom, respectively. Selection of the best spectral band to discriminate between all groups was made by calculating the likelihood ratio to test the equality over all g groups for each spectral band. Define the likelihood ratio statistic F_j for the j^{th} spectral band as:

$$F_j = \frac{(t_{jj}-w_{jj})/(g-1)}{(w_{jj})/(n-g)}$$

Note that as F_j increases, the probability of equality over all g groups measured by the j^{th} spectral band decreases. Large values of F_j indicate good discrimination between groups by the j^{th} spectral band. As defined above, the statistic F_j obeys an F- distribution with g-1 and n-g degrees of freedom.

The statistic F_j was calculated for each spectral band. The spectral band having the largest F- value was then selected as being the best single discriminant between all g groups.

After selecting the best spectral band to discriminate between the g groups, the next step was to select another band which, when in combination with the first band selected, would provide the best pair of spectral bands to discriminate between the groups. It should be noted that the program did <u>not</u>

perform an "all possible pairs" analysis. In order to perform this, and all subsequent steps in the analysis, the W and T matrices were partitioned into two disjoint sets; those variables which had been previously selected and those variables not yet selected. The sub-matrix corresponding to those variables not yet selected was modified to reflect the effect of data correlations between the previously selected variables and those not yet selected. The F_j statistic was therefore interpreted as the likelihood ratio test of equality over all g groups of the conditional distribution of the j^{th} spectral band, given that specific bands had already been selected. This step-wise procedure was repeated until all spectral bands had entered the calculation.

2.1.2 Methodologies

The methods used to employ the stepwise discriminant analysis were relatively straightforward. After deciding on a particular land use category to be included in an analysis, the test site field signature file was inspected and specific field signatures corresponding to that land use category were selected. The stepwise discriminant analysis procedure was intended to provide information about the "best" or optimal set of spectral bands available for discriminating between groups. It was believed that in this type of analysis, only the best or most representative field signatures should be included in the various signature groups. Throughout all of the stepwise discriminant analysis, only those fields with field quality ratings of three or higher were included (see Section 1.5.1.2). Fields with field quality ratings of three were only used when there were no quality rating four or five fields available.

As an additional aid to the analyst, the LARS transformed divergence measure (see Swain, Robertson, and Walker, 1971) was used to estimate the spectral separability of the input data classes at the end of each step of the discriminant analysis. The transformed divergence measure is a measure of spectral separability between two separate spectral signatures. While the transformed divergence measure is related to the probability of correct classification the relationship is extremely complex and in general, cannot be explicitly determined in advance of performing an actual multispectral classification. Swain and King (1973) have investigated the relationship between the transformed divergence measure and the probability of correct classification for the two-class case. The transformed divergence measure was found most useful during this analysis as a tool for measuring the relative improvement in spectral separability when additional spectral bands were added during each step of the discriminant analysis.

2.1.3 Results

Stepwise spectral discriminant analysis was initially employed to address the question of selecting an optimal set of 13 SDO channels providing complete spectral coverage. For this analysis, 350 fields were selected from the August 5th, 1973 SL/3 overpass of the Washington, D.C. - Baltimore, Maryland test site.

Spectral signatures were calculated using SDO channels 1 through 16 and 21. Table 4 shows the relationship between SDO number and spectral band number. Note that the thermal channel, spectral band number 13, was detected by three SDO channels. SDO channel numbers 17 through 20 and 22 were not included in these analyses because they were the only channels which sample those specific spectral bands. Since the stepwise discriminant analysis program had been modified to accept at most 13 spectral channels, the analysis was performed in two parts. First, spectral signatures for SDO channels 1 through 12 were calculated and analyzed. Finally, the spectral signatures for the same 350 fields were recalculated using SDO channels 13 through 16 and 21.

After selecting a complete set of spectral channels, the test site field signature file and the test site field boundary/video data file were recreated using the 13 SDO's identified in the first phase of the analysis. Thirteen spectral band signatures were calculated for all of the fields identified on the SL/3 overpass. Of the original 155 test sites planned to be covered by SL/3, only 87 test sites contained fields identified by the photointerpreters. From these 87 test sites, spectral signatures were calculated for 609 fields. Table 5 shows a summarized list for all of the land use categories identified in the 87 test sites. Figure 11 shows the geographic location of the 87 test sites superimposed on an image of spectral band number 8, of Skylab mission SL/3. These 609 fields formed the data base used in all of the remaining analyses.

Table 4. S-192 MULTISPECTRAL SCANNER CONFIGURATION

IFOV - 79.3 Meter Square Ground Coverage
Swath Width - 72.4 km.

SPECTRAL BAND NUMBER	DESCRIPTION	SPECTRAL RANGE	SDO CHANNEL(S)
1	Violet	0.41-0.46um	22
2	Violet-Blue	0.46-0.51um	18
3	Blue-Green	0.52-0.56um	1,2
4	Green-Yellow	0.56-0.61um	3,4
5	Orange-Red	0.62-0.67um	5,6
6	Red	0.68-0.76um	7,8
7	Near Infrared	0.78-0.88um	9,10
8	Near Infrared	0.98-1.08um	19
9	Near Infrared	1.09-1.19um	20
10	Mid Infrared	1.20-1.30um	17
11	Mid Infrared	1.55-1.75um	11,12
12	Mid Infrared	2.10-2.35um	13,14
13	Thermal Infrared	10.2-12.5um	15,16,21

Table 5. SUMMARY OF LAND USE CATEGORIES

IDENTIFIED ON DIGITAL S-192 IMAGERY

LAND USE CODE	NUMBER OF FIELDS	TOTAL NUMBER OF PIXELS	LAND USE CATEGORY
1110	24	1267	Single-Family Household Units
1111	1	33	Rural, Low Density, with Trees
1112	11	433	Rural, Low Density, without Trees
1113	27	2339	Urban, High Density, with Trees
1114	21	1536	Urban, High Density, without Trees
1120	12	432	Multi-Family Household Units
1121	2	218	Low Density, with Trees
1122	3	146	Low Density, without Trees
1123	2	106	High Density, with Trees
1124	23	560	High Density, without Trees
1150	3	53	Mobile Home Parks or Courts
1154	1	7	Mobile Home, High Density, Without Trees
1200	1	40	Commercial & Services
1210	6	248	Wholesale Trade Areas
1215	2	25	Farm Products (Raw Materials)

LAND USE CODE	NUMBER OF FIELDS	TOTAL NUMBER OF PIXELS	LAND USE CATEGORY
1220	19	474	Retail Trade Areas (Business Dist., Shopping Centers, Comm.)
1230	1	96	Business, Professional, & Personal Services
1243	2	69	Sports (Stadiums, Arenas, Racetracks, Other)
1246	3	46	Recreational (Tennis, Ice Skating, Stables, Play Areas)
1300	8	489	Industrial
1310	2	63	Mechanical Processing
1330	3	76	Chemical Processing
1340	1	18	Fabrication & Assembly
1410	1	127	Stone Quarries
1420	3	116	Sand & Gravel Pits
1511	2	102	Highways
1520	5	252	Railroads & Associated Facilities
1530	1	23	Airports & Associated Facilities
1531	1	68	Commercial
1532	1	46	General
1533	1	46	Military
1544	3	145	Port Facilities
1560	4	98	Electric, Gas, Water, Sewage Disposal, Solid Waste, Utilities

LAND USE CODE	NUMBER OF FIELDS	TOTAL NUMBER OF PIXELS	LAND USE CATEGORY
1610	5	47	Educational Facilities
1611	4	36	Primary
1612	6	76	Secondary
1614	2	27	College
1615	2	221	University
1616	2	34	Other
1621	4	89	Hospitals
1631	2	12	Churches
1640	1	44	Military Areas
1641	2	39	Housing
1643	3	412	Storage Areas
1644	1	14	Training Areas
1650	2	57	Correctional
1653	1	37	State
1660	5	181	Government & Administrative Offices
1911	10	777	Golf Courses
1912	12	669	Cemeteries
1913	3	141	Park
1914	4	152	Parking Lots

LAND USE CODE	NUMBER OF FIELDS	TOTAL NUMBER OF PIXELS	LAND USE CATEGORY
2112	7	237	Bare-Recently Plowed
2113	151	3876	Growing Crop Present
2114	1	12	Harvested
2120	24	799	Abandoned
2142	58	2130	Improved
2143	8	167	Unimproved
2250	5	171	Turf Farm
2320	1	8	Poultry and Egg Houses
4100	30	1118	Deciduous
4200	7	138	Evergreen (Coniferous & Other)
4300	9	364	Mixed
4400	3	76	Upland Shrubs
5110	6	366	Natural(Rivers & Creeks)
5120	1	55	Man-Made (Canals, Ditches, & Aquaducts)
5210	4	161	Natural Lakes & Ponds
5300	1	54	Reservoirs
5410	3	906	Bays

LAND USE CODE	NUMBER OF FIELDS	TOTAL NUMBER OF PIXELS	LAND USE CATEGORY
6000	4	149	Wetlands
6100	6	198	Vegetated Wetlands
6110	2	202	Brackish Marsh
6200	2	379	Forested Wetlands
7500	5	315	Disturbed Land

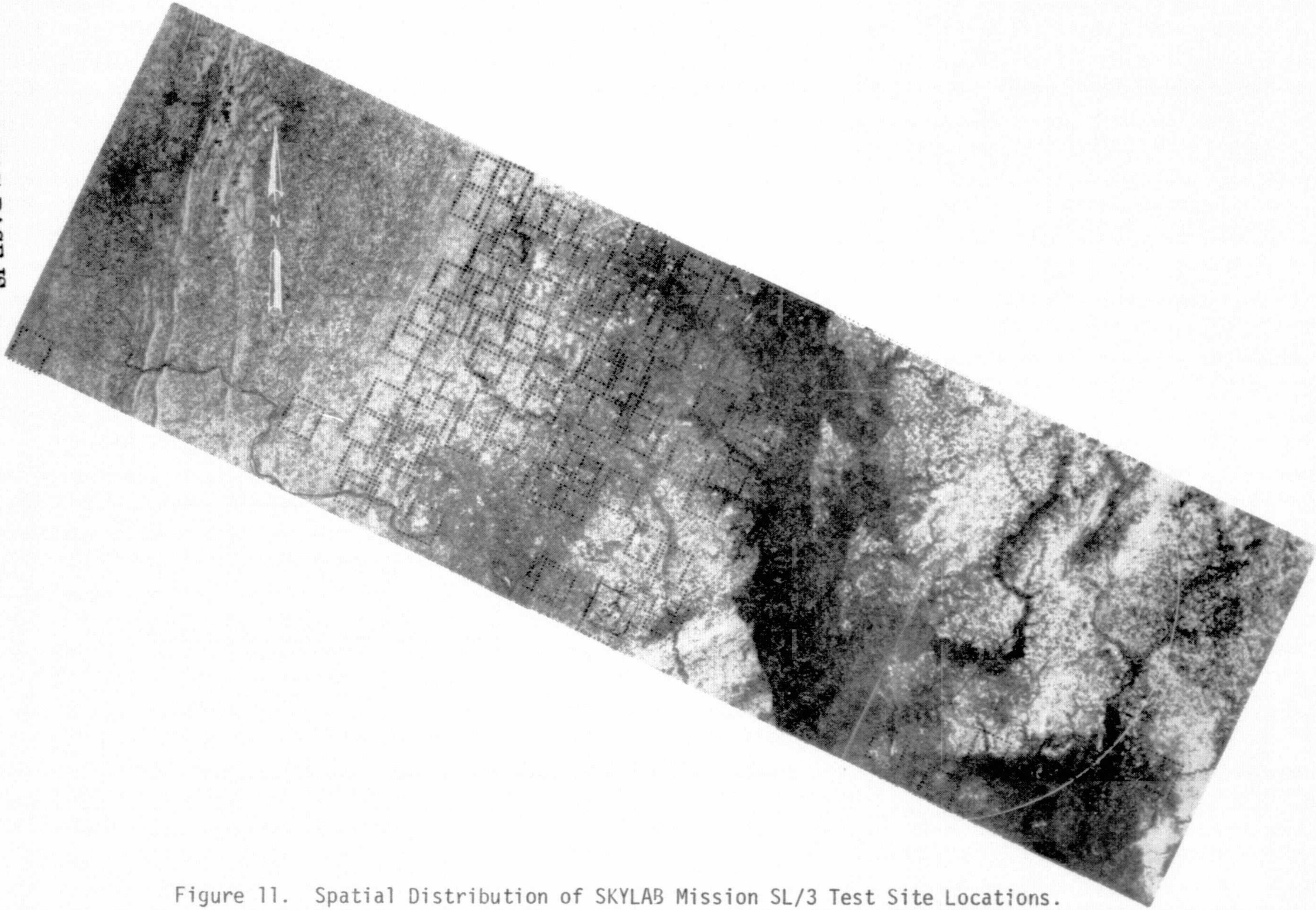


Figure 11. Spatial Distribution of SKYLAB Mission SL/3 Test Site Locations.

The test site location map has been superimposed on an image from spectral band number 8 of the SKYLAB Multispectral Scanner. The scale of the reproduction is approximately 1:1,130,000.

After completing the preliminary analysis just described, the stepwise discriminant analysis program was used to select the optimal spectral bands for discriminating between various land use categories. The first analysis performed was designed to identify those spectral bands most useful for discriminating between various level I categories. Five data groups corresponding to the level I categories; Urban, Agricultural, Forest, Water, and Wetlands, were defined. The stepwise discriminant analysis program selected the best spectral bands for discriminating between these categories.

Each level I category was then considered separately. The stepwise discriminant analysis program determined which spectral bands provided the greatest discrimination between various urban categories, agricultural categories, and so forth. The detailed results of these two analyses will now be discussed.

2.1.3.1 SDO Channel Selection

Spectral signatures were calculated for 350 fields selected on the SL/3 overpass of the Washington, D.C. - Baltimore, Maryland test site. Since only 13 channels of data could be analyzed at any one time and 17 SDO channels had to be included in this analysis, two sets of spectral signatures were computed for each of the 350 fields. The first signature in each set was a 12 channel signature calculated using SDO channel numbers 1 through 12. The second signature in each set was a five channel signature including SDO channel numbers 13 through 16 and

21. As is shown in Table 4 SDO channels 17 through 20 and 22 were the only channels used to sense those spectral bands, so those channels had to be included in the later analysis.

The 350 fields were partitioned into 19 groups of approximately equal size. Each group corresponded to field signatures from separate level III land use categories. The 19 groups were selected to provide a representative sample of all land use categories in the data base. The 19 groups were then analyzed by the stepwise discriminant analysis program. In the analysis using the 12 channel signatures, it was thought that high data correlations between the SDO channels sampling the same spectral bands would prevent two SDO channels with identical spectral coverage from being selected prior to selecting one channel from each available spectral band. This, in fact, was observed. All of the odd numbered SDO channels between 1 and 12 were selected before any even numbered SDO channel. In the analysis using the five channel signatures, the same effect was present. SDO channels 14 and 21 were selected first. Table 6 shows the final results of the SDO channel selection.

2.1.3.2 Spectral Band Selection

Thirteen band spectral signatures using the SDO channels shown in Table 6 were calculated for the 609 fields identified on the SL/3 overpass of the Washington,

Table 6. STEPWISE DISCRIMINANT ANALYSIS RESULTS FOR SDO CHANNEL SELECTION

SPECTRAL BAND NUMBER	SDO CHANNEL SELECTED	SPECTRAL COVERAGE
1	22	0.41-0.46 um
2	18	0.46-0.51 um
3	1	0.52-0.56 um
4	3	0.56-0.61 um
5	5	0.62-0.67 um
6	7	0.68-0.76 um
7	9	0.78-0.88 um
8	19	0.98-1.08 um
9	20	1.09-1.19 um
10	17	1.20-1.30 um
11	11	1.55-1.75 um
12	14	2.10-2.35 um
13	21	10.2-12.5 um

D.C. - Baltimore, Maryland test site. Field signatures were selected from land use levels III and IV categories and then aggregated up to five level I composite groups. The five composite groups included:

1) urban (all land use code 1000);
2) agriculture (all land use code 2000);
3) forest (all land use code 4000);
4) water (all land use code 5000);
5) wetlands (all land use code 6000).

Using this type of data grouping, the stepwise discriminant analysis would select the spectral bands of greatest utility for separating the various broad categories of land use. Results of this stepwise discriminant analysis are shown in Table 7. From the F-values listed, it should be noted that very little additional discriminability was added after five spectral channels had been selected. Table 8 shows the transformed divergence values calculated after five variables had entered the analysis. From the results of the divergence calculation, it was noted that a relatively high degree of separability existed between the five broad categories used. The least well discriminated categories were forest and wetlands. This is undoubtedly due to the presence of "forested wetlands."

After selection of the optimal bands for discriminating between level I categories, the stepwise discriminant technique was employed to select bands for discriminating within each general level I category separately.

Table 7. SPECTRAL BAND SELECTION FOR DISCRIMINATING BETWEEN ALL LEVEL 1 CATEGORIES

STEP NUMBER	SPECTRAL BAND NUMBER	SPECTRAL COVERAGE	F-VALUE
1	11	1.55-1.75 um	1,437,292
2	9	1.09-1.19 um	432,780
3	13	10.2-12.5 um	90,188
4	5	0.62-0.67 um	8,747
5	6	0.68-0.76 um	1,053
6	7	0.78-0.88 um	276
7	1	0.41-0.46 um	252
8	12	2.10-2.35 um	147
9	3	0.52-0.56 um	96
10	10	1.20-1.30 um	78
11	8	0.98-1.08 um	29
12	4	0.56-0.61 um	22
13	2	0.46-0.51 um	7

Table 8. Transformed Divergence Values Calculated After Five Spectral Bands had Entered the Between All Level I Analysis. Spectral Band Numbers Included: 5,6,9,11,13.

GROUP	1000	2000	4000	5000	6000
1000:	0				
2000:	1930	0			
4000:	1950	1682	0		
5000:	1975	1985	1965	0	
6000:	1831	1969	1589	1895	0

The spectral response of different land use categories varied greatly. It was anticipated that the optimal spectral bands for discriminating within one land use category may be completely different from the set of optimal bands for another category. Five analyses were performed; one for each of the general categories used in the "between level I" analysis.

The first of the five analyses included 19 level III and IV urban groups:

1) Single Family Household Units,
2) Multi-family Household Units,
3) Wholesale Trade Areas,
4) Retail Trade Areas,
5) Industrial,
6) Chemical Processing,
7) Fabrication and Assembly,
8) Stone Quarries,
9) Sand and Gravel Pits,
10) Highways, Auto Parking, Bus Terminals, etc.,
11) Railroads and associated facilities,
12) Airports and associated facilities,
13) Marine craft facilities,
14) Electric, Gas, Water facilities,
15) Educational facilities,
16) Medical and Health facilities,
17) Military Areas,

18) Government and Administrative Offices,

19) Improved Lands.

Table 9 shows the order of spectral band selection using only urban categories. The range of F-values shown here is considerably smaller than was observed with the between level I analyses. This indicates that (all) 19 urban groups exhibited considerably greater spectral overlap. Table 10 shows the transformed divergence values calculated after five variables had entered the anlaysis. The divergence calculation indicated that many of the groups were highly indistinguishable. This was understandable since the 19 groups defined included categories of such fine distinction as single family versus multi-family household units.

Only four groups were available for the within agricultural analysis:

1) Active Cropland,

2) Abandoned Cropland,

3) Pasture,

4) Turf Farms.

Table 11 shows the spectral band selection obtained for these groups and Table 12 shows the divergence values after five variables were entered. The divergence calculation indicated only moderate separability between these groups. As might be expected, the categories active cropland and turf farms were very poorly discriminated.

Table 9. SPECTRAL BAND SELECTION FOR DISCRIMINATING WITHIN URBAN CATEGORIES

STEP NUMBER	SPECTRAL BAND NUMBER	SPECTRAL COVERAGE	F-VALUE
1	9	1.09-1.19 um	114,731
2	3	0.52-0.56 um	36,275
3	11	1.55-1.75 um	19,801
4	6	0.68-0.76 um	9,805
5	13	10.2-12.5 um	7,810
6	1	0.41-0.46 um	3,768
7	7	0.78-0.88 um	3,039
8	5	0.62-0.67 um	2,302
9	12	2.10-2.35 um	1,447
10	4	0.56-0.61 um	1,140
11	2	0.46-0.51 um	988
12	8	0.98-1.08 um	449
13	10	1.20-1.30 um	506

Table 10. Transformed Divergence Values Calculated After Five Spectral Bands Had Entered the All Urban Analysis. Spectral Band Numbers Included: 3,6,9,11,13.

GROUP	1110	1120	1210	1220	1300	1330	1340	1410	1420	1510	1520	1530	1540	1560	1610	1620	1640	1660	1910
1110:	0																		
1120:	265	0																	
1210:	995	519	0																
1220:	1212	805	641	0															
1300:	1359	801	782	635	0														
1330:	1743	1419	1000	669	929	0													
1340:	1518	1128	1312	1655	1627	1578	0												
1410:	1690	1701	1622	1561	1708	1575	1841	0											
1420:	824	943	1289	1119	1734	1603	1705	1811	0										
1510:	1323	1329	1283	1606	1907	1759	1480	1896	1258	0									
1520:	1660	1202	1308	1382	555	1242	1715	1849	1893	1978	0								
1530:	1157	835	641	800	939	1440	1678	1442	1269	1410	1523	0							
1540:	1987	1903	1788	1852	1144	1730	1942	1832	1989	1991	681	1878	0						
1560:	1141	940	1191	1625	1095	1904	1895	1912	1672	1726	1350	1096	1900	0					
1610:	563	775	1341	1660	1710	1929	1745	1950	1211	1630	1852	1469	1990	1340	0				
1620:	1035	1320	1704	1910	1903	1988	1939	1970	1680	1551	1965	1600	1987	1347	602	0			
1640:	1436	1200	949	776	1101	1041	1619	1250	1212	1268	1575	855	1862	1628	1781	1883	0		
1660:	654	482	568	852	1239	1484	1582	1890	868	950	1670	784	1975	1072	1028	1354	1070	0	
1910:	760	834	999	1460	1333	1867	1740	1555	1274	1488	1784	1204	1987	1039	1153	1141	1356	1209	0

Table 11. SPECTRAL BAND SELECTION FOR DISCRIMINATING WITHIN AGRICULTURAL CATEGORIES

STEP NUMBER	SPECTRAL BAND NUMBER	SPECTRAL COVERAGE	F-VALUE
1	9	1.09-1.19 um	34,760
2	7	0.78-0.88 um	21,798
3	1	0.41-0.46 um	4,137
4	11	1.55-1.75 um	713
5	6	0.68-0.76 um	342
6	10	1.20-1.30 um	185
7	8	0.98-1.08 um	95
8	3	0.52-0.56 um	86
9	5	0.62-0.67 um	269
10	12	2.10-2.35 um	94
11	13	10.2-12.5 um	51
12	4	0.56-0.61 um	35
13	2	0.46-0.54 um	6

Table 12. Transformed Divergence Values Calculated After Five Spectral Bands had Entered the All Agricultural Analysis. Spectral Band Numbers Included: 1,6,7,9,11.

GROUP	2110	2120	2140	2250
2110:	0			
2120:	725	0		
2140:	989	1402	0	
2250:	483	936	1361	0

Somewhat better discrimination was observed in the remaining three analyses. In the within forest analysis four groups were included:

1) Deciduous,
2) Evergreens (conifers and other),
3) Mixed,
4) Upland Shrubs.

Table 13 shows the spectral band selection obtained using these four forest categories and Table 14 shows the divergence calculation for the forest analysis. Again, as expected, coniferous forests were well discriminated from all other forested areas except mixed coniferous and deciduous forests. The four water groups analyzed showed surprising separability. Those groups were:

1) Natural Rivers and Creeks,
2) Man-made Canals, Ditches, and Aqueducts,
3) Natural Lakes and Ponds,
4) Bays.

As shown in Table 15 and 16 the spectral bands in the visible part of the spectrum were selected as providing the best discrimination between water categories. Finally, the analysis of wetland categories is shown in Tables 17 and 18. Four wetland groups were included in the analysis:

1) General Wetlands,
2) Vegetated Wetlands,
3) Brackish Marsh,
4) Forested Wetlands.

Table 13. SPECTRAL BAND SELECTION FOR DISCRIMINATING WITHIN FOREST CATEGORIES

STEP NUMBER	SPECTRAL BAND NUMBER	SPECTRAL COVERAGE	F-VALUE
1	9	1.09-1.19 um	70,881
2	3	0.52-0.56 um	16,365
3	5	0.62-0.67 um	2,864
4	4	0.56-0.61 um	1,581
5	11	1.55-1.75 um	626
6	6	0.68-0.76 um	301
7	12	2.10-2.35 um	173
8	13	10.2-12.5 um	137
9	1	0.41-0.46 um	112
10	7	0.78-0.88 um	90
11	10	1.20-1.30 um	76
12	8	0.98-1.08 um	96
13	2	0.46-0.54 um	72

Table 14. Transformed Divergence Values Calculated After Five Spectral Bands had Entered the All Forest Analysis. Spectral Band Numbers Included: 3,4,5,9,11

GROUP	4100	4110	4200	4300	4400
4100:	0				
4110:	1400	0			
4200:	1432	1674	0		
4300:	1139	1356	995	0	
4400:	981	890	1941	1725	0

Table 15. SPECTRAL BAND SELECTION FOR DISCRIMINATING WITHIN WATER CATEGORIES

STEP NUMBER	SPECTRAL BAND NUMBER	SPECTRAL COVERAGE	F-VALUE
1	6	0.68-0.76 um	32,111
2	1	0.41-0.46 um	10,353
3	3	0.52-0.56 um	4,180
4	8	0.98-1.08 um	331
5	9	1.09-1.19 um	201
6	13	10.2-12.5 um	146
7	12	2.10-2.35 um	128
8	4	0.56-0.61 um	79
9	11	1.55-1.75 um	91
10	7	0.78-0.88 um	67
11	2	0.46-0.54 um	54
12	5	0.62-0.67 um	29
13	10	1.20-1.30 um	13

Table 16. Transformed Divergence Values Calculated After Five Spectral Bands had Entered the All Water Analysis.
Spectral Band Numbers Included: 1,3,6,8,9.

GROUP	5110	5120	5210	5410
5110:	0			
5120:	934	0		
5210:	980	1382	0	
5410:	1091	1020	1555	0

Table 17. SPECTRAL BAND SELECTION FOR DISCRIMINATING WITHIN WETLANDS CATEGORIES

STEP NUMBER	SPECTRAL BAND NUMBER	SPECTRAL COVERAGE	F-VALUE
1	10	1.20-1.30 um	318,695
2	3	0.52-0.56 um	98,144
3	1	0.41-0.46 um	5,096
4	8	0.98-1.08 um	136
5	6	0.68-0.76 um	102
6	13	10.2-12.5 um	65
7	4	0.56-0.61 um	64
8	11	1.55-1.75 um	34
9	7	0.78-0.88 um	27
10	9	1.09-1.19 um	13
11	12	2.10-2.35 um	9
12	5	0.62-0.67 um	1
13	2	0.46-0.54 um	1

Table 18. Transformed Divergence Values Calculated After Five Spectral Bands had Enter the All Wetlands Analysis.
Spectral Band Numbers Included: 1,3,6,8,10.

GROUP	6000	6100	6110	6200
6000:	0			
6100:	1040	0		
6110:	1831	1991	0	
6200:	1989	1975	1992	0

The forested wetlands as well as the brackish marsh areas were excellently discriminated from the other vegetated wetlands.

2.1.4 Implications of Reported Results

After completing the analysis just discussed, the results were studied to determine the effect these findings had on subsequent analysis tasks. The results obtained from the spectral band selection were used in the design of the multispectral classification algorithm employed in the final analysis task.

2.1.4.1 Implications from SDO Channel Selection Results

At first study, the SDO channel selection results appear to be completely consistent with what one would expect. In both the 12 channel analysis and the 5 channel analysis, SDO channels were selected from each available spectral band before any single spectral band was selected twice.

Careful study of the 12 channel spectral signatures, however, revealed a curious anomaly. As reported earlier, the SDO channel pairs which sampled data from the same spectral band were expected to show very high data correlations. This was observed for some of the field signatures analyzed, but in a great many other cases surprisingly low correlations were calculated for the spectral band pairs. It was observed that the degree of correlation between spectral band pairs was roughly proportional to the field size. This phenomenon was eventually explained in terms of the S-192 sensor design characteristics.

The S-192 digital data included data from 13 disjoint spectral regions. The digital data tapes, however, contained output from 22 Scientific Data Output channels. The 22 SDO channels were produced by sampling the 13 detectors at two different sampling rates. The detectors sensing spectral bands 1, 2, 8, 9, 10, and 13 were sampled at a low sampling rate which corresponded to an approximate 72.6 meter center to center spacing. The detectors sensing spectral bands 3, 4, 5, 6, 7, 11, 12, and 13 were sampled at twice the low rate which corresponds to an approximate 36.3 meter center to center spacing. The detectors which were sampled at the high rate, produced two SDO channels. All of the odd numbered samples from a high rate detector were combined to form one of the SDO channels while the even numbered samples were combined to form the other channel. It is important to note that this type of sampling design produces a one-half pixel misregistration between the two SDO channels produced from a single detector. In addition, spatial misregistration was introduced by the scan lines straightening algorithm as well as the scanner electronics and tape recording system. This additional misregistration effect has been reported to be as large as two pixels (Sattinger, 1975).

It was also learned that the sampling procedure used to sample the low rate detectors was not necessarily synchronized with the sampling of either of the corresponding SDO channels produced from the high rate sampling.

The anomalies observed in the 12 channel spectral signatures could be explained by the spectral band to band misregistration. The misregistration between bands only affected field boundary pixels. Large fields would generally have a smaller ratio of boundary pixels to internal pixels than small fields. The effect of the spacial misregistration should, therefore, be greatest for small fields. Table 19 shows spectral signatures calculated for two fields from test site 9. Note that field number 3 contained 91 pixels while field number 4 contained only 19 pixels. Consider the first four SDO channels shown on the correlation matrices. SDO channel pair 1 and 2 had a correlation coefficient of 0.636 for the 91 pixel field but only 0.293 for the 19 pixel field. SDO channel pair 3 and 4 showed a similar effect; correlation of 0.471 for the 91 pixel field and 0.259 for the 19 pixel field. Note, also, that the correlation coefficients between SDO numbers 1 and 3 are higher than the correlation coefficients between SDO numbers 1 and 2 for both fields. Since SDO channels 1 and 3 should be in registration, even though they sampled different spectral regions the correlations between them were higher than the correlations between two SDO's sampling the same spectral region but misregistered by one-half pixel (i.e. SDO channels 1 and 2). The same effect was observed for SDO channels 2 and 4 with spectral band pairs 3 and 4. Analysis of this type has lead to the following conclusions:

ORIGINAL PAGE IS OF POOR QUALITY

TEST SITE: 9 FIELD: 3

COUNTY: 5 WINDOW ROW: 1619 LAND USE CODE: 4100 CENTER ROW: 1661.22 NUMBER OF BANDS: 12
SL MISSION: 3 COL: 954 QUALITY RATING: 5 COL: 991.87 NUMBER OF PIXELS: 91.

MEAN VALUES	STD. DEV.	SDO NUMBERS	CORRELATION MATRIX 1	2	3	4	5	6	7	8	9	10	11	12	0
56.9	4.6	1	1.000	0.636	0.743	0.419	0.462	0.146	0.663	0.476	0.151	0.015	0.729	0.569	0.0
56.4	4.5	2	0.636	1.000	0.522	0.711	0.404	0.415	0.602	0.635	0.203	0.153	0.692	0.748	0.0
30.7	5.4	3	0.743	0.522	1.000	0.471	0.625	0.290	0.632	0.439	0.212	0.019	0.667	0.549	0.0
30.5	5.0	4	0.419	0.711	0.471	1.000	0.373	0.638	0.426	0.553	0.079	0.172	0.539	0.659	0.0
29.2	6.1	5	0.462	0.404	0.625	0.373	1.000	0.352	0.514	0.371	0.224	0.002	0.476	0.381	0.0
28.6	5.9	6	0.146	0.415	0.290	0.638	0.352	1.000	0.180	0.459	0.190	0.184	0.256	0.399	0.0
76.3	8.1	7	0.663	0.602	0.632	0.426	0.514	0.180	1.000	0.681	0.496	0.232	0.721	0.645	0.0
75.5	7.2	8	0.476	0.635	0.439	0.553	0.371	0.459	0.681	1.000	0.527	0.443	0.646	0.664	0.0
84.7	7.3	9	0.151	0.203	0.212	0.079	0.224	0.190	0.496	0.527	1.000	0.554	0.364	0.330	0.0
84.2	7.1	10	0.015	0.153	0.019	0.172	0.002	0.184	0.232	0.443	0.554	1.000	0.173	0.213	0.0
43.7	8.0	11	0.729	0.692	0.667	0.539	0.476	0.256	0.721	0.646	0.364	0.173	1.000	0.876	0.0
42.8	6.9	12	0.569	0.748	0.549	0.659	0.381	0.399	0.645	0.664	0.330	0.213	0.878	1.000	0.0
0.0	-0.0	0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

TEST SITE: 9 FIELD: 4

COUNTY: 5 WINDOW ROW: 1619 LAND USE CODE: 2120 CENTER ROW: 1663.74 NUMBER OF BANDS: 12
SL MISSION: 3 COL: 954 QUALITY RATING: 3 COL: 980.32 NUMBER OF PIXELS: 19.

MEAN VALUES	STD. DEV.	SDO NUMBERS	CORRELATION MATRIX 1	2	3	4	5	6	7	8	9	10	11	12	0
64.4	2.7	1	1.000	0.293	0.580	0.250	0.522	0.151	0.161	0.324	-0.461	-0.324	0.592	0.302	0.0
64.6	3.1	2	0.293	1.000	0.204	0.759	0.687	0.333	0.021	0.560	-0.030	-0.384	0.368	0.437	0.0
38.4	3.9	3	0.580	0.204	1.000	0.259	0.505	0.479	0.482	0.212	-0.265	-0.035	0.581	0.535	0.0
38.1	4.7	4	0.250	0.759	0.259	1.000	0.617	0.303	0.080	0.587	0.002	-0.307	0.473	0.530	0.0
37.9	6.5	5	0.522	0.687	0.505	0.617	1.000	0.398	0.057	0.517	-0.393	-0.343	0.624	0.421	0.0
38.8	5.6	6	0.151	0.333	0.479	0.303	0.398	1.000	-0.058	-0.126	-0.436	-0.480	0.415	0.615	0.0
84.3	4.9	7	0.161	0.021	0.482	0.080	0.057	-0.058	1.000	0.300	0.372	0.223	0.275	0.315	0.0
83.4	5.2	8	0.324	0.560	0.212	0.587	0.517	-0.126	0.300	1.000	0.165	0.290	0.286	0.116	0.0
79.8	7.1	9	-0.461	-0.030	-0.265	0.002	-0.393	-0.436	0.372	0.165	1.000	0.390	-0.519	-0.117	0.0
78.9	6.1	10	-0.324	-0.384	-0.035	-0.307	-0.343	-0.480	0.223	0.290	0.390	1.000	-0.400	-0.605	0.0
58.4	6.3	11	0.592	0.368	0.581	0.473	0.624	0.415	0.275	0.286	-0.519	-0.400	1.000	0.560	0.0
59.4	5.0	12	0.302	0.437	0.535	0.530	0.421	0.615	0.315	0.116	-0.117	-0.605	0.560	1.000	0.0
0.0	-0.0	0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 19. Spectral Signatures for Two Fields from Test Site 9

1) band to band spatial registration must be considered with high priority in the design of subsequent sensor systems. Spatial misregistration by as little as one-half pixel has been seen to introduce "noise" into the system which may seriously negate the advantages obtained through improved spectral resolution;

2) small spatial misregistration between spectral bands may seriously degrade any attempt at boundary delineation for area mensuration using multispectral data processing techniques. For that matter, any multispectral data analysis will be degraded.

Finally, it should be noted that the oversampling of the high rate detectors which introduced the inherent misregistration problem does not, for this scanner, improve the spatial resolution by a factor of two as might be expected. This is due to the large instantaneous field of view (IFOV) of the scanner detectors. The IFOV of the scanner is about 75 meters, as determined by a field stop at the entrance slit of the spectrometer. At the best system resolution for a low rate channel determined by sampling theory (the spatial wavelength equals twice the sample spacing), and assuming uniform illumination of the entrance slit, the modulation transfer function (MTF) of the detector is about $2/\pi$ for a fixed spatial wavelength, so long as the IFOV is unchanged. Oversampling does not change the MTF. Thus, the high channels use the same aperture as the low rate channels. This aperture is oversized for the new channel resolution based on the new sample spacing. In fact, the high rate channels have an MTF of near zero at their sampling theory resolution limit. This means that little information is actually perceived at these shorter spatial wavelengths. Consequently, little improvement in spatial resolution is produced by the high rate channels.

Using these conclusions as a basis, it is possible to explain results observed by the SDO channel selection analysis. The SDO channels selected were identified not only on the basis of which channels were least noisy but rather on the basis of which channels were spatially registered with one another.

2.1.4.2 Implications from the Spectral Band Selection

Table 20 shows a summary of all of the spectral band selection analyses. Listed in that table are the step numbers in which each band was selected during the various analyses. The average step number in which each spectral band entered the six analyses was calculated (see Overall Average). An overall rating was assigned each band on the basis of the average step numbers.

Implied in the results shown in Table 20 is the conclusion that the set of spectral channels which provided optimal discrimination between broad land use Level

Table 20. SPECTRAL BAND SELECTION ORDER FOR VARIOUS SIGNATURE GROUP DEFINITIONS

SPECTRAL BAND NUMBER	SPECTRAL COVERAGE	BETWEEN ALL LEVEL 1	WITHIN SINGLE LEVEL 1					OVERALL AVERAGE	OVERALL RATING
			1000	2000	4000	5000	6000		
1	0.41-0.46 um	7	6	3	9	2	3	5.0	4
2	0.46-0.51 um	13	11	13	13	11	13	12.3	13
3	0.52-0.56 um	9	2	8	2	3	2	4.3	2
4	0.56-0.61 um	12	10	12	4	8	7	8.8	11
5	0.62-0.67 um	4	8	9	3	12	12	8.0	8
6	0.68-0.76 um	5	4	5	6	1	5	4.3	2
7	0.78-0.88 um	6	7	2	10	10	9	7.3	7
8	0.98-1.08 um	11	12	7	12	4	4	8.3	9
9	1.09-1.19 um	2	1	1	1	5	10	3.3	1
10	1.20-1.30 um	10	13	6	11	13	1	9.0	12
11	1.55-1.75 um	1	3	4	5	9	8	5.0	4
12	2.10-2.35 um	8	9	10	7	7	11	8.7	10
13	10.2-12.5 um	3	5	11	8	6	6	6.5	6

I categories differed significantly from the optimal bands for discriminating within any single land use category. This information was used in the design of the multispectral classification algorithm used in the final analysis task.

Selection of a fixed set of spectral bands for use in classifying all land use categories would tend to degrade the classification accuracy since no single set of bands was optimal for all land use categories. Performing multispectral classification using all 13 spectral bands would greatly increase the amount of processing necessary to classify an area. It was therefore decided that a two-staged classification algorithm would provide the best overall technique for performing multispectral classification. In the first stage of the classification, a maximum likelihood classifier using the five best spectral bands identified for discriminating between all Level I land use categories was used to assign each pixel to a general Level I category. The first stage classification assignment of a particular pixel was used to determine what spectral bands would be used to further classify that pixel during the second stage. The spectral bands, as well as the training class signatures used in the second stage, were entirely determined by the results of the first stage analysis. In each stage two classification, the best five spectral bands selected for that particular Level I category were used. For example, suppose that a particular pixel was identified as "forest" during the first stage classification. That pixel would be assigned to the class "4000" (Level I land use category for forest). During the second stage classification, the best five spectral bands identified by the "within forest" discriminant analysis would be used, and the training set signatures would correspond to Level II forest categories. The second stage classifier may then assign that pixel to the Level II forest category "2," coniferous forest. The final classification assignment would therefore be 4200 -coniferous forest. Both stage classifiers used a maximum likelihood decision rule.

2.2 CLUSTERING ANALYSIS

The value of clustering in automatic data processing is that the clusters so defined are "natural" groups in the data and represent the way the the scanner perceives the environment in a multispectral sense. To the degree that these multispectral clusters are pure land use categories of a sort desired by the land use planner, the clusters will be regarded as valuable. To the degree that the clusters are "mixed" - again in a land use classification sense - they will be regarded as of little value.

In the present analysis a simple clustering algorithm was used to test SL/3 cluster groups. A mixed Level III and Level IV land use classification was used to show the overlap and discriminability between functional land use categories useful to land use planners.

2.2.1. Analysis Techniques

The distance measure employed in the clustering is the reduction of the divergence measure for the assumption of diagonal covariance matrices:

$$2J = \sum_{i=1}^{N} \frac{\{ \Sigma_{gi}^2 + \Sigma_{xi}^2 \}\{ \mu_{gi} - \mu_{xi} \}^2 + \{ \Sigma_{gi}^2 - \Sigma_{xi}^2 \}}{\Sigma_{gi}^2 \, \Sigma_{xi}^2}$$

Where	N	is the number of spectral bands
	μ_{gi}	is the cluster group mean for band i
	μ_{xi}	is the signature mean for band i
	Σ_{gi}^2	is the variance of the cluster group for band i
	Σ_{xi}^2	is the variance of the signature for band i
	J	is the divergence distance between group "g" and signature "x"

The clustering procedures used all 13 bands selected as described earlier, and were tear-down, rather than build-up, starting with a threshold of five, giving seven groups, and then sequentially to thresholds of four (13 groups) and three (23 groups), representing a rough doubling of groups at each step. No further breakdown was attempted because close inspection of the distances of individual items from the group means showed that fairly tight groups were established at a threshold of three. Only outliers would be broken away with further splitting, and no reasonable further breaks of major groups within clusters would take place. This topic is picked up again later on.

The distance measure employed is very closely related to measures in common use but has the advantage of reduced computer processing time - arising from the assumption of diagonal covariance matrices - with a parallel disadvantage of some loss in discrimination. It is not in a strict sense a metric, but nor are almost all distance measures in common use. It is our judgement, that the precision obtained is acceptable for the exploratory study reported here.

2.2.2. Selection of Sites and Thresholds for Clustering

The present study uses SL/3 ground truth data only in a clustering mode for comparison with the land use categories as given in Table 2.

The test site data actually used in the clustering was considerably smaller than the original six percent (over land) sample initially intended. Amoung the reasons in addition to those already mentioned were:

- All fields or areas less than ten pixels were excluded because of extensive boundary pixel problems, and field boundary identification problems on imagery.

- Many test sites were eliminated in toto because of haze, thin clouds in part of the northwest area of the strip, and because of deep haze, high cirrus, and altostratus clouds in the Eastern Shore area of Maryland (in the southern third of the strip).
- Many fields could not be identified on either band 4 or 8 used for production of shade prints and perforce had to be eliminated. (A considerable variety of data was available, including RC-10 photography, alphanumeric SKYLAB shade prints, and channel 4 and 8 black and white images, together with the polaroid enlargements of the area and ground truth. Despite this wide array of material, there were some sites in which it was absolutely infeasible to place boundaries around individual fields; these therefore had to be eliminated.)
- Some test sites as a whole were so lacking in constrast features as shown on the band 4 and 8 shade prints and images that they also were eliminated as a whole. If the window allocation were perfect, this would have been no problem. However, the digital window was off by several pixels and thus a blind overlay could not properly be used.
- Finally, only high-quality fields or units were used in order that clean signatures be used for clustering.

As a result of all of these processes, the total number of high-quality, large (greater than 10 pixels), cloud-free and shadow-free ground truth fields or land units available for clustering was 527 in SL/3. The total area involved was something less than a one percent sample of the land area, but was concentrated because of cloud cover problems, and therefore was no longer a strictly random sample, although it is still well distributed throughout the entire test site. Proportionately more data was available for the Baltimore, Washington and Virginia areas and less for the eastern shore, and Delmarva Peninsula areas of Maryland.

The various thresholds used in the study started with a threshold of five (5) - a coarse threshold leading to seven broad classes roughly comparable to Level I in the Anderson et. al. (1972) classification. This was followed by thresholding to distances of four (4) and three (3), using the distance measure noted above. The threshold of three (3) is reported on here, because it is close to Level II of the Anderson et. al. (1972) classification. The threshold five results are considered first.

2.2.3. Results of the Clustering

2.2.3.1. The Threshold Five (5) Results

Seven groups were obtained employing a threshold of five (5), as follows:

Group 1

This group is overwhelmingly urban (code 1,000), with 100 of 117 cases being urban.

The 17 aberrant cases included:

2112	Bare - recently ploughed (1)
2113	Growing Crop Present (9)
2120	Abandoned Agricultural Land (2)
2240	Nursery and Horticultural Area (1)
4300	Mixed Forest (1)
7500	Disturbed Land (2)

Nine (9) of these in turn were relative outliers in the group, lying at distances of greater than 2 from the group mean. Only 24 cases exceeded a distance of 2, so that the aberrant inclusions are proportionately a large component of the peripheral members of the cluster.

Group 2

This group is principally water (code 5,000), comprising 14 out of 20 cases. Of the remaining six cases all but two (deciduous forest) were probably very wet at the time of overflight - which occurred shortly after flooding rains - or naturally would have some water within the group in any case. The latter comment applies to:

1544	Port Facilities
6000	Wetlands
6110	Brackish Marsh (2)

Group 3

This group is a very small and peculiar mixed bag. The composition of the group is:

1330	Chemical Processing Plant
1641	Housing in Military Areas
2142	Improved Pasture Land
2143	Unimproved Pasture Land
7500	Disturbed Land (2 cases)

Group 4

This group is predominantly code 2,000 (Agricultural Land): 173 of 223 cases. Of the remaining 50 confusion cases, a significant number contain much growing vegetation (golf courses, cemetery, deciduous forest, park, vegetated wetlands, etc.) as listed below:

1110	Single Family Household Unit (6)
1112	Rural, Low Density Housing (1)
1113	Urban High Density with Trees (9)
1114	Urban High Density without Trees (6)
1124	High Density Multi-Family Residential Without Trees (1)
1220	Urban Retail Trade Areas (2)
1246	Recreational (Tennis, Stables, Play Areas, Etc.) (1)
1611	Primary Education Facility (1)
1911	Golf Course (8)
1912	Cemetary (2)
1913	Urban Parkland (2)
4100	Deciduous Forest (3)
6100	Vegetated Wetlands (2)
6200	Forested Wetlands (1)

The group may properly be thought of then as a growing crop-pasture- and grassed area group in which case the only evident "outsiders" in the cluster are:

1110	Single Family Household Unit (6)
1113	Urban High Density with Trees (9)
1114	Urban High Density without Trees (6)
1220	Urban Retail Trade Areas (2)
1246	Recreational (Tennis, Stables, Play Areas, etc.) (1)
1611	Primary Education Facility (1)
4100	Deciduous Forest (3)
6100	Vegetated Wetlands (2)
6200	Forested Wetlands (1)

Of these 31 cases in turn, a reasonable number are likely to contain much grass.

This cluster therefore appears to be a very rational group mostly of green growing vegetation without trees and with most "outsiders" being explicable.

Group 5

Group 5 is almost exclusively a category of growing crop present (code 2113), and discriminates just two test sites - 126 and 199 - in the Eastern Shore. Of the 13 members of the group, ten consist of growing crop present (principally corn), one is a golf course, one an abandoned farm, and one a mechanical processing plant surrounded by green lawns. In essence, this group also is green vegetation but with no trees. It is quite similar to Group 4 and appears to be mainly discriminating a single crop (corn, at the ear stage) type from the mixture of crops present in Group 4.

Group 6

Group 6 is overwhelmingly an urban group (42 of 44 cases), comprising mostly intensive commercial and industrial developments with little vegetation. The various categories of the group are as below at level III and IV (as appropriate):

1114	High Density Single Family Residential without Trees (7)
1220	Retail Trade Areas (1)
1300	Industrial (5)
1400	Extractive Industry (1)
1500	Transportation (9)
1600	Institutional (6)
1914	Parking Lots (3)
6000	Wetlands (2)

Group 7

The final group is very mixed comprising primarily urban categories and deciduous and other forest land. Most of the urban categories are those in which considerable tree cover is indicated in the code, or is reasonably likely to be present as in cemetaries, golf courses, educational facilities, and so on.

The various groups present at level III and level IV (as appropriate) are:

1110	Single Family Residential (3)
1111	Rural Residential, Low Density with Trees (1)
1112	Rural, Low Density Single Family Residential, without Trees (1)
1113	Urban, High Residential Single Family Density with Trees (14)
1121	Rural, Low Density Multi-Family Residential, with Trees (1)
1123	High Density, Multi-Family Residential, with Trees (> 50% tree cover) (1)
1124	High Density, Multi-Family Residential, without Trees (Less than 50% Tree Cover) (3)
1210	Wholesale Trade Area (1)
1246	Recreational (1)
1300	Industrial (2)
1610	Educational Facilities (1)
1612	Secondary Education Facility (1)
1615	University (2)
1621	Hospitals (1)
1653	State Institutional Facility (1)
1911	Golf Courses (1)
1912	Cemetaries (4)
1913	Urban Parkland (1)
	Urban Total: 40
2113	Growing Crop Present (10)
2120	Abandoned Farmland (4)
2142	Improved Pasture (3)
2250	Turf Farm (1)
	Agricultural Total 18
4100	Deciduous Forest (24)
4200	Evergreen Forest (5)
4300	Mixed Forest (7)
4400	Upland Shrub (2)
	Forest Total 38
5210	Natural Lakes and Ponds (1)
	Water Total 1
6000	Wetlands (2)
6100	Vegetated Wetlands (3)
6200	Forested Wetlands (1)
	Wetlands Total 6
	GRAND TOTAL 103

This group appears to be composite cluster in that all urban categories lie within a distance of 2.38 of the group mean, whereas 30 of 38 forest cases lie at distances of greater than 2.38 from the group mean. This group would, therefore, be expected to split to some degree in the later clustering (Table 25).

2.2.3.3 Clustering With A Threshold of Three and Four

The clusters formed by using thresholds respectively of four (13 groups) and three (23 groups) are shown in the following tables to which reference should be made.

Table 21 shows the group means and standard deviations for the 23 groups obtained with threshold 3. Table 22 gives the between group distance matrix for the 23 groups. Table 23 lists the 13 groups established with the threshold of 4, ordered within each group by land use code. Table 24 lists the 23 groups for threshold 3, ordered by distance from the group mean. Table 25 lists the 23 groups for threshold 3 ordered by land use code. It is thus readily feasible to make cross comparisons between Tables 23 and 25 and Tables 24 and 25 for different purposes.

The number of groups discriminated at threshold 3 (23) is close to the number of Level II categories in the land use classification (21), actually present in the SL/3 data set. On the other hand, the 13 groups separated at threshold 4 are intermediate between the 6 Level I categories and the 21 Level II categories. In this respect, it is worth noting that while both threshold 5 and 3 steps stabilized (less than 10 cases switching between all clusters with 3 cycles after establishment of the 7 and 23 groups respectively) this was not the case with threshold 4. Threshold 4 is evidently not a clean break point and switching from

Table 21. Group Means and Standard Deviations for Cluster Groups (23)

GROUP NUMBER	NUMBER OF PIXELS	NUMBER OF FIELDS	GROUP MEANS/STANDARD DEVIATIONS												
1	4402.0	93	108.5	117.5	72.7	45.0	48.9	93.4	85.4	98.2	99.1	98.2	63.8	39.0	164.6
			6.9	24.6	7.6	8.4	10.1	10.2	8.7	10.8	9.2	10.3	8.1	8.7	7.4
2	1137.0	7	101.1	107.4	57.0	29.3	30.0	34.3	20.7	19.1	22.2	16.1	10.6	9.0	152.3
			5.6	32.0	3.5	4.0	6.2	6.1	8.9	6.0	4.5	5.6	2.5	4.1	5.7
3	186.0	3	124.7	162.9	120.6	97.7	97.9	153.5	114.9	127.7	133.2	134.1	103.1	77.4	152.8
			20.5	40.4	39.2	45.2	37.3	38.4	23.3	27.0	28.1	29.1	23.1	24.0	12.9
4	1378.0	30	104.6	109.7	66.7	39.1	40.8	96.4	104.4	117.8	118.8	115.9	67.3	36.0	159.9
			6.3	30.6	4.7	5.2	7.3	8.9	17.6	16.6	13.8	14.8	7.5	7.2	6.6
5	234.0	9	130.5	133.7	87.1	55.0	58.3	100.7	89.9	100.0	102.7	95.4	54.7	36.0	122.3
			8.6	33.0	6.2	6.7	9.5	7.6	10.9	10.1	9.0	7.7	7.3	9.1	10.7
6	893.0	21	114.1	132.6	85.2	58.4	63.8	95.6	68.6	76.5	77.6	79.7	62.9	47.0	171.1
			6.7	17.3	11.6	12.8	14.1	17.4	14.0	16.2	14.5	17.8	12.8	12.7	9.4
7	928.0	13	102.3	102.5	59.6	32.8	33.7	73.5	75.9	82.3	84.7	79.6	45.2	25.0	156.4
			7.6	15.4	9.5	9.2	9.6	19.4	23.1	29.1	26.9	30.2	16.5	9.7	7.9
8	1051.0	31	108.1	117.6	72.3	44.3	48.5	82.3	65.9	75.7	76.4	77.1	55.9	38.7	169.9
			6.7	17.6	6.7	7.4	9.4	10.9	12.8	16.6	14.5	16.7	9.0	8.2	9.9
9	13.0	2	145.5	150.5	106.7	68.3	66.2	125.7	116.9	123.6	131.8	117.9	60.5	39.8	93.8
			3.5	35.3	3.6	6.6	7.4	8.5	7.7	7.1	3.8	7.8	4.0	7.4	8.4
10	2317.0	45	106.9	109.1	66.3	38.5	41.2	81.2	79.1	[illegible]	90.6	87.8	52.8	30.9	162.3
			7.4	25.9	6.1	6.5	8.1	8.2	8.5	11.4	9.5	11.4	7.4	7.0	7.4
11	5495.0	154	101.7	102.1	63.0	36.1	37.3	89.8	98.7	112.9	113.6	108.4	60.7	31.5	157.3
			6.4	15.8	5.0	5.2	7.1	8.3	12.8	11.3	9.9	10.4	8.5	7.4	6.7
12	298.0	7	116.3	120.6	71.7	41.5	44.4	73.5	68.1	76.9	75.8	70.9	37.9	21.3	143.2
			6.1	7.6	3.3	3.9	5.7	7.9	11.4	13.8	10.0	11.6	5.5	4.8	8.1
13	23.0	2	106.4	110.1	65.4	35.7	37.7	52.0	39.3	46.3	48.0	44.5	27.0	16.7	133.9
			6.7	16.0	3.4	3.4	4.9	5.4	3.0	4.9	4.8	4.9	2.1	4.8	3.9
14	845.0	14	110.6	120.3	71.0	42.8	48.6	69.7	45.6	52.6	52.9	54.5	44.4	34.1	176.7
			6.3	21.2	8.9	9.7	10.3	12.1	10.8	14.6	12.1	15.0	8.9	7.2	8.3
15	269.0	9	109.2	124.8	84.5	61.7	67.8	117.6	96.6	109.9	113.3	113.4	83.1	59.0	160.4
			10.1	25.8	16.8	20.8	23.6	24.9	10.0	12.0	11.3	16.0	21.3	21.6	8.0
16	20.0	1	112.6	106.5	55.1	28.6	35.6	98.4	119.2	138.7	138.5	131.5	66.8	30.6	158.8
			4.9	28.0	4.6	3.5	3.4	5.3	4.9	6.2	4.8	5.1	2.2	5.0	4.5
17	1844.0	57	100.5	98.7	56.0	29.3	30.6	77.3	90.3	103.8	103.2	97.3	47.1	22.3	153.0
			6.2	13.6	4.7	4.7	6.4	7.5	9.7	10.3	9.4	9.7	7.1	6.5	6.3
18	279.0	8	114.2	105.9	66.4	36.0	37.2	75.3	82.1	92.9	90.4	83.2	37.2	16.9	150.1
			6.7	27.1	3.2	3.4	5.2	4.3	4.9	7.8	5.6	7.8	3.9	5.0	7.9
19	11.0	2	155.1	181.5	115.5	79.0	79.3	113.5	79.1	88.6	92.1	87.5	62.1	45.1	111.2
			6.4	15.6	12.7	13.5	8.7	10.0	5.3	4.3	6.2	5.8	4.2	5.6	10.3
20	16.0	3	95.6	96.1	47.3	23.1	24.3	40.7	37.6	44.9	40.5	38.6	14.3	7.8	147.1
			5.7	7.0	2.4	3.2	3.7	6.5	4.4	23.0	11.7	20.7	2.6	4.3	6.0
21	76.0	4	101.3	98.9	55.7	28.7	31.7	61.9	65.8	75.4	74.7	70.3	36.9	19.0	154.0
			5.8	10.8	2.7	2.9	5.3	5.8	6.8	9.6	7.3	10.2	3.7	4.2	5.3
22	263.0	6	113.2	115.4	68.9	39.4	40.5	48.9	29.5	24.6	28.5	21.7	12.6	8.7	145.4
			7.3	20.3	3.7	4.1	5.8	5.6	7.6	5.6	7.5	6.2	3.6	4.5	6.3
23	83.0	6	107.5	111.6	68.1	38.0	37.7	96.6	116.8	126.5	121.6	113.1	53.6	24.4	152.8
			6.9	37.8	3.8	3.8	5.5	8.5	16.1	12.3	9.7	9.5	6.2	5.0	6.0

ORIGINAL PAGE IS
OF POOR QUALITY

Table 22. Between Group Distance Matrix for Cluster Groups (23)

GROUP NUMBER	1	2	3	4	5	6	7	8	9	10	11	12
1	0.0											
2	11.64	0.0										
3	5.39	19.51	0.0									
4	1.70	12.83	7.17	0.0								
5	2.60	12.18	5.85	3.51	0.0							
6	1.92	10.36	3.73	3.15	2.90	0.0						
7	2.40	8.37	5.10	2.38	3.84	2.47	0.0					
8	1.68	8.99	5.85	2.47	3.18	1.44	1.79	0.0				
9	6.96	18.91	7.26	7.28	4.35	7.41	8.85	8.28	0.0			
10	1.31	9.81	7.13	2.21	3.17	2.44	1.83	1.35	8.10	0.0		
11	1.74	12.16	7.65	0.98	3.63	3.54	2.34	2.66	7.76	1.87	0.0	
12	3.25	7.80	10.75	3.90	3.90	3.77	2.71	2.34	9.10	2.34	3.61	0.0
13	9.93	5.65	19.71	12.00	10.05	8.89	7.44	7.22	16.76	7.91	10.90	5.15
14	3.46	6.59	6.32	4.07	4.68	2.14	2.33	1.53	10.35	2.79	4.30	2.71
15	2.34	14.71	2.15	3.23	3.18	2.28	3.09	3.20	6.08	3.64	3.57	5.83
16	5.30	19.51	12.55	3.17	6.76	8.12	7.44	7.49	9.13	6.16	3.55	8.70
17	2.63	10.44	9.36	2.53	4.42	4.32	2.10	3.02	9.11	1.85	1.62	3.37
18	3.56	10.33	13.32	4.40	4.54	5.10	3.34	3.35	9.80	2.10	3.52	2.31
19	5.00	15.23	7.10	7.07	3.27	4.46	6.45	5.47	5.68	5.75	7.11	7.55
20	9.43	3.65	19.57	9.78	10.77	10.01	5.96	7.74	17.36	7.54	8.91	6.64
21	4.96	7.52	15.24	5.31	6.79	6.38	2.93	4.04	12.67	3.20	4.44	3.32
22	9.13	2.41	15.79	10.17	9.52	7.84	6.51	6.74	15.86	7.65	9.75	5.63
23	2.63	12.78	9.48	1.40	3.78	4.40	3.03	3.43	7.39	2.75	1.43	3.96

GROUP NUMBER	13	14	15	16	17	18	19	20	21	22	23
13	0.0										
14	5.18	0.0									
15	13.78	4.53	0.0								
16	17.20	10.32	7.14	0.0							
17	8.90	4.22	4.90	5.28	0.0						
18	7.58	4.53	7.24	8.03	2.39	0.0					
19	12.38	6.80	5.11	10.89	7.89	8.42	0.0				
20	4.71	6.43	13.03	15.29	7.01	7.73	15.44	0.0			
21	5.16	3.96	8.65	9.32	2.82	2.96	11.10	4.93	0.0		
22	3.95	4.84	11.56	17.12	8.57	8.31	12.44	4.60	6.29	0.0	
23	12.19	5.05	4.71	3.29	2.56	4.17	8.08	9.78	5.33	10.47	0.0

Table 23. 13 Groups Established Using a Threshold of 4.0

Table 23-1. Cluster Group #1

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
1	1	2	1110	3	35.0	1.02	SINGLE-FAMILY HOUSEHOLD UNITS
1	353	3	1110	4	10.0	1.31	SINGLE-FAMILY HOUSEHOLD UNITS
1	353	4	1110	4	20.0	1.64	SINGLE-FAMILY HOUSEHOLD UNITS
1	127	2	1110	4	57.0	0.91	SINGLE-FAMILY HOUSEHOLD UNITS
1	115	4	1110	5	88.0	1.44	SINGLE-FAMILY HOUSEHOLD UNITS
1	186	3	1110	5	201.0	1.39	SINGLE-FAMILY HOUSEHOLD UNITS
1	299	4	1110	5	16.0	2.49	SINGLE-FAMILY HOUSEHOLD UNITS
1	122	7	1110	5	24.0	1.70	SINGLE-FAMILY HOUSEHOLD UNITS
1	115	3	1110	5	33.0	1.35	SINGLE-FAMILY HOUSEHOLD UNITS
1	79	4	1110	5	38.0	1.06	SINGLE-FAMILY HOUSEHOLD UNITS
1	127	7	1110	5	42.0	1.38	SINGLE-FAMILY HOUSEHOLD UNITS
1	118	2	1110	5	52.0	1.12	SINGLE-FAMILY HOUSEHOLD UNITS
1	115	8	1110	5	67.0	0.91	SINGLE-FAMILY HOUSEHOLD UNITS
1	115	1	1112	5	198.0	1.13	RURAL, LOW DENSITY, WITHOUT TREES
1	64	6	1112	4	32.0	1.10	RURAL, LOW DENSITY, WITHOUT TREES
1	64	2	1112	5	17.0	1.56	RURAL, LOW DENSITY, WITHOUT TREES
1	65	2	1113	5	47.0	1.87	URBAN, HIGH DENSITY, WITH TREES
1	118	11	1113	4	32.0	1.38	URBAN, HIGH DENSITY, WITH TREES
1	355	2	1113	4	45.0	1.28	URBAN, HIGH DENSITY, WITH TREES
1	84	8	1113	5	38.0	1.50	URBAN, HIGH DENSITY, WITH TREES
1	97	3	1114	4	10.0	1.64	URBAN, HIGH DENSITY, WITHOUT TREES
1	110	2	1114	5	158.0	0.94	URBAN, HIGH DENSITY, WITHOUT TREES
1	122	3	1114	5	53.0	1.01	URBAN, HIGH DENSITY, WITHOUT TREES
1	86	3	1114	5	68.0	1.47	URBAN, HIGH DENSITY, WITHOUT TREES
1	86	5	1114	5	74.0	1.49	URBAN, HIGH DENSITY, WITHOUT TREES
1	65	3	1114	5	81.0	1.08	URBAN, HIGH DENSITY, WITHOUT TREES
1	110	4	1114	4	16.0	1.64	URBAN, HIGH DENSITY, WITHOUT TREES
1	127	3	1114	4	24.0	1.13	URBAN, HIGH DENSITY, WITHOUT TREES
1	124	1	1114	4	59.0	1.69	URBAN, HIGH DENSITY, WITHOUT TREES
1	127	1	1114	4	61.0	2.10	URBAN, HIGH DENSITY, WITHOUT TREES
1	65	5	1114	5	31.0	2.03	URBAN, HIGH DENSITY, WITHOUT TREES
1	200	1	1114	5	36.0	1.13	URBAN, HIGH DENSITY, WITHOUT TREES
1	118	1	1114	5	48.0	1.19	URBAN, HIGH DENSITY, WITHOUT TREES
1	86	1	1114	5	241.0	1.27	URBAN, HIGH DENSITY, WITHOUT TREES
1	353	7	1120	4	18.0	1.62	MULTI-FAMILY HOUSEHOLD UNITS
1	122	2	1120	5	16.0	1.64	MULTI-FAMILY HOUSEHOLD UNITS
1	68	14	1120	5	27.0	0.93	MULTI-FAMILY HOUSEHOLD UNITS
1	115	6	1120	5	42.0	1.32	MULTI-FAMILY HOUSEHOLD UNITS
1	84	12	1121	5	104.0	1.80	LOW DENSITY, WITH TREES
1	64	3	1122	5	45.0	1.13	LOW DENSITY, WITHOUT TREES
1	64	4	1122	5	62.0	0.75	LOW DENSITY, WITHOUT TREES
1	299	5	1123	5	21.0	1.97	HIGH DENSITY, WITH TREES
1	113	9	1124	5	41.0	0.76	HIGH DENSITY, WITHOUT TREES
1	115	5	1124	5	23.0	1.70	HIGH DENSITY, WITHOUT TREES
1	186	4	1124	5	18.0	1.30	HIGH DENSITY, WITHOUT TREES
1	179	3	1124	5	16.0	0.47	HIGH DENSITY, WITHOUT TREES

Table 23-1 (Continued)

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
1	113	8	1124	5	14.0	0.95	HIGH DENSITY, WITHOUT TREES
1	122	4	1124	5	11.0	2.08	HIGH DENSITY, WITHOUT TREES
1	97	4	1124	4	16.0	1.56	HIGH DENSITY, WITHOUT TREES
1	359	1	1150	5	16.0	1.02	MOBILE HOME PARKS OR COURTS
1	316	5	1150	4	21.0	0.92	MOBILE HOME PARKS OR COURTS
1	117	4	1150	3	16.0	1.44	MOBILE HOME PARKS OR COURTS
1	99	1	1210	4	24.0	2.19	WHOLESALE TRADE AREAS
1	1	7	1210	3	19.0	1.50	WHOLESALE TRADE AREAS
1	359	4	1220	4	25.0	0.99	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
1	54	3	1220	4	18.0	1.55	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
1	353	5	1230	4	96.0	0.95	BUSINESS, PROFESSIONAL, & PERSONAL SERVICES
1	186	6	1243	5	38.0	1.17	SPORTS (STADIUMS,ARENAS,RACETRACKS,OTHER)
1	122	12	1300	4	40.0	1.92	INDUSTRIAL
1	84	11	1420	5	58.0	0.86	SAND & GRAVEL PITS
1	41	4	1420	4	18.0	2.07	SAND & GRAVEL PITS
1	64	5	1511	4	19.0	1.62	HIGHWAYS
1	119	3	1511	5	83.0	1.20	HIGHWAYS
1	79	2	1532	4	46.0	1.00	GENERAL
1	64	12	1533	5	46.0	1.07	MILITARY
1	186	5	1560	5	12.0	1.68	ELECTRIC, GAS, WATER, SEWAGE DISPOSAL, SOLID WASTE, UTIL
1	86	4	1610	4	12.0	1.71	EDUCATIONAL FACILITIES
1	85	4	1612	4	15.0	1.48	SECONDARY
1	6	8	1614	4	19.0	1.46	COLLEGE
1	64	14	1616	5	25.0	0.96	OTHER
1	25	5	1621	5	46.0	1.18	HOSPITALS
1	64	11	1641	5	22.0	1.75	HOUSING
1	95	2	1643	4	348.0	0.74	STORAGE AREAS
1	359	7	1650	5	51.0	1.00	CORRECTIONAL
1	300	3	1660	5	18.0	2.22	GOVERNMENT & ADMINISTRATIVE OFFICES
1	85	1	1660	5	130.0	0.93	GOVERNMENT & ADMINISTRATIVE OFFICES
1	109	3	1912	4	29.0	0.85	CEMETERIES
1	8	6	1912	5	75.0	0.71	CEMETERIES
1	7	12	1912	5	73.0	0.91	CEMETERIES
1	108	1	1912	5	50.0	1.29	CEMETERIES
1	359	3	1912	5	57.0	1.28	CEMETERIES
1	109	4	1912	4	23.0	1.22	CEMETERIES
1	188	7	2112	4	11.0	3.28	BARE--RECENTLY PLOWED
1	113	14	2113	5	23.0	1.52	GROWING CROP PRESENT
1	199	6	2113	5	17.0	2.41	GROWING CROP PRESENT
1	113	13	2113	5	18.0	1.90	GROWING CROP PRESENT
1	199	7	2113	5	19.0	2.94	GROWING CROP PRESENT
1	122	9	2113	3	26.0	1.46	GROWING CROP PRESENT
1	192	9	2113	5	15.0	1.68	GROWING CROP PRESENT
1	307	8	2113	5	29.0	2.32	GROWING CROP PRESENT
1	191	2	2113	5	16.0	1.02	GROWING CROP PRESENT
1	188	1	2113	5	81.0	2.25	GROWING CROP PRESENT

Table 23-1 (Continued)

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
1	192	2	2113	5	69.0	1.65	GROWING CROP PRESENT
1	1	6	2120	3	17.0	1.78	ABANDONED
1	187	1	2120	5	136.0	1.29	ABANDONED
1	68	6	2120	4	16.0	2.02	ABANDONED
1	337	3	2120	5	10.0	2.67	ABANDONED
1	127	4	2142	4	15.0	1.58	IMPROVED
1	201	6	2142	5	145.0	1.20	IMPROVED
1	307	9	2240	5	15.0	2.24	NURSERIES AND FLORICULTURAL AREAS
1	307	7	2240	5	29.0	2.53	NURSERIES AND FLORICULTURAL AREAS
1	316	6	7500	5	34.0	1.39	DISTURBED LAND
1	316	4	7500	5	85.0	0.96	DISTURBED LAND

Table 23-2. Cluster Group #2

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
2	124	5	4100	5	16.0	3.51	DECIDUOUS
2	168	1	5110	5	64.0	1.69	NATURAL (RIVERS & CREEKS)
2	95	4	5110	5	69.0	2.57	NATURAL (RIVERS & CREEKS)
2	44	1	5110	5	147.0	2.18	NATURAL (RIVERS & CREEKS)
2	129	1	5110	5	16.0	2.08	NATURAL (RIVERS & CREEKS)
2	129	2	5110	5	16.0	2.98	NATURAL (RIVERS & CREEKS)
2	111	3	5110	5	54.0	3.09	NATURAL (RIVERS & CREEKS)
2	108	7	5120	5	55.0	1.43	MAN-MADE (CANALS, DITCHES, & AQUADUCTS)
2	135	2	5210	5	83.0	3.33	NATURAL LAKES & PONDS
2	98	6	5210	5	29.0	1.70	NATURAL LAKES & PONDS
2	117	12	5210	5	30.0	1.85	NATURAL LAKES & PONDS
2	57	8	5410	5	189.0	1.34	BAYS
2	45	1	5410	5	395.0	0.76	BAYS
2	98	1	5410	5	322.0	2.03	BAYS
2	102	1	6110	5	179.0	2.49	BRACKISH MARSH

Table 23-3. Cluster Group #3

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
3	35	4	1330	4	18.0	1.91	CHEMICAL PROCESSING
3	64	7	1641	4	17.0	2.73	HOUSING
3	308	12	2142	4	21.0	1.13	IMPROVED
3	121	10	2143	4	37.0	1.40	UNIMPROVED
3	359	6	7500	5	62.0	2.63	DISTURBED LAND
3	316	1	7500	5	128.0	0.76	DISTURBED LAND

Table 24. 23 Groups Established Using a Threshold of 3.0 and Ordered by Distant Measure

Table 24-1. Cluster Group #1

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
1	179	3	1124	5	16.0	0.52	HIGH DENSITY, WITHOUT TREES
1	8	6	1912	5	75.0	0.65	CEMETERIES
1	95	2	1643	4	348.0	0.69	STORAGE AREAS
1	64	4	1122	5	62.0	0.70	LOW DENSITY, WITHOUT TREES
1	113	9	1124	5	41.0	0.76	HIGH DENSITY, WITHOUT TREES
1	109	3	1912	4	29.0	0.81	CEMETERIES
1	84	11	1420	5	58.0	0.82	SAND & GRAVEL PITS
1	127	2	1110	4	57.0	0.84	SINGLE-FAMILY HOUSEHOLD UNITS
1	7	12	1912	5	73.0	0.88	CEMETERIES
1	68	14	1120	5	27.0	0.88	MULTI-FAMILY HOUSEHOLD UNITS
1	115	8	1110	5	67.0	0.89	SINGLE-FAMILY HOUSEHOLD UNITS
1	110	2	1114	5	158.0	0.89	URBAN, HIGH DENSITY, WITHOUT TREES
1	64	14	1616	5	25.0	0.89	OTHER
1	85	1	1660	5	130.0	0.91	GOVERNMENT & ADMINISTRATIVE OFFICES
1	316	4	7500	5	85.0	0.94	DISTURBED LAND
1	122	3	1114	5	53.0	0.95	URBAN, HIGH DENSITY, WITHOUT TREES
1	316	5	1150	4	21.0	0.96	MOBILE HOME PARKS OR COURTS
1	191	2	2113	5	16.0	0.96	GROWING CROP PRESENT
1	359	4	1220	4	25.0	0.98	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
1	113	8	1124	5	14.0	1.00	HIGH DENSITY, WITHOUT TREES
1	359	7	1650	5	51.0	1.00	CORRECTIONAL
1	353	5	1230	4	96.0	1.01	BUSINESS, PROFESSIONAL, & PERSONAL SERVICES
1	79	2	1532	4	46.0	1.01	GENERAL
1	79	4	1110	5	38.0	1.02	SINGLE-FAMILY HOUSEHOLD UNITS
1	64	6	1112	4	32.0	1.02	RURAL, LOW DENSITY, WITHOUT TREES
1	1	2	1110	3	35.0	1.02	SINGLE-FAMILY HOUSEHOLD UNITS
1	359	1	1150	5	16.0	1.03	MOBILE HOME PARKS OR COURTS
1	65	3	1114	5	81.0	1.03	URBAN, HIGH DENSITY, WITHOUT TREES
1	118	2	1110	5	52.0	1.04	SINGLE-FAMILY HOUSEHOLD UNITS
1	127	3	1114	4	24.0	1.05	URBAN, HIGH DENSITY, WITHOUT TREES
1	64	3	1122	5	45.0	1.07	LOW DENSITY, WITHOUT TREES
1	64	12	1533	5	46.0	1.12	MILITARY
1	118	1	1114	5	48.0	1.14	URBAN, HIGH DENSITY, WITHOUT TREES
1	25	5	1621	5	46.0	1.14	HOSPITALS
1	109	4	1912	4	23.0	1.14	CEMETERIES
1	115	1	1112	5	198.0	1.15	RURAL, LOW DENSITY, WITHOUT TREES
1	200	1	1114	5	38.0	1.15	URBAN, HIGH DENSITY, WITHOUT TREES
1	86	1	1114	5	241.0	1.18	URBAN, HIGH DENSITY, WITHOUT TREES
1	201	6	2142	5	145.0	1.19	IMPROVED
1	119	3	1511	5	83.0	1.20	HIGHWAYS
1	187	1	2120	5	136.0	1.22	ABANDONED
1	108	1	1912	5	50.0	1.23	CEMETERIES
1	359	3	1912	5	57.0	1.23	CEMETERIES
1	355	2	1113	4	45.0	1.24	URBAN, HIGH DENSITY, WITH TREES
1	353	3	1110	4	10.0	1.24	SINGLE-FAMILY HOUSEHOLD UNITS
1	115	3	1110	5	33.0	1.26	SINGLE-FAMILY HOUSEHOLD UNITS

Table 24-1 (Continued)

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
1	186	6	1243	5	38.0	1.26	SPORTS (STADIUMS,ARENAS,RACETRACKS,OTHER)
1	186	4	1124	5	18.0	1.27	HIGH DENSITY, WITHOUT TREES
1	186	3	1110	5	201.0	1.28	SINGLE-FAMILY HOUSEHOLD UNITS
1	115	6	1120	5	42.0	1.28	MULTI-FAMILY HOUSEHOLD UNITS
1	118	11	1113	4	32.0	1.28	URBAN, HIGH DENSITY, WITH TREES
1	127	7	1110	5	42.0	1.31	SINGLE-FAMILY HOUSEHOLD UNITS
1	115	4	1110	5	88.0	1.35	SINGLE-FAMILY HOUSEHOLD UNITS
1	86	3	1114	5	68.0	1.39	URBAN, HIGH DENSITY, WITHOUT TREES
1	316	6	7500	5	34.0	1.40	DISTURBED LAND
1	86	5	1114	5	74.0	1.41	URBAN, HIGH DENSITY, WITHOUT TREES
1	84	8	1113	5	38.0	1.43	URBAN, HIGH DENSITY, WITH TREES
1	113	14	2113	5	23.0	1.46	GROWING CROP PRESENT
1	117	4	1150	3	16.0	1.46	MOBILE HOME PARKS OR COURTS
1	64	2	1112	5	17.0	1.46	RURAL, LOW DENSITY, WITHOUT TREES
1	85	4	1612	4	15.0	1.49	SECONDARY
1	1	7	1210	3	19.0	1.50	WHOLESALE TRADE AREAS
1	127	4	2142	4	15.0	1.50	IMPROVED
1	122	9	2113	3	26.0	1.51	GROWING CROP PRESENT
1	6	8	1614	4	19.0	1.54	COLLEGE
1	64	5	1511	4	19.0	1.56	HIGHWAYS
1	97	3	1114	4	10.0	1.57	URBAN, HIGH DENSITY, WITHOUT TREES
1	353	4	1110	4	20.0	1.60	SINGLE-FAMILY HOUSEHOLD UNITS
1	353	7	1120	4	18.0	1.60	MULTI-FAMILY HOUSEHOLD UNITS
1	110	4	1114	4	16.0	1.61	URBAN, HIGH DENSITY, WITHOUT TREES
1	122	2	1120	5	16.0	1.62	MULTI-FAMILY HOUSEHOLD UNITS
1	86	4	1610	4	12.0	1.67	EDUCATIONAL FACILITIES
1	186	5	1560	5	12.0	1.67	ELECTRIC, GAS, WATER, SEWAGE DISPOSAL, SOLID WASTE, UTIL
1	122	7	1110	5	24.0	1.68	SINGLE-FAMILY HOUSEHOLD UNITS
1	115	5	1124	5	23.0	1.69	HIGH DENSITY, WITHOUT TREES
1	192	9	2113	5	15.0	1.72	GROWING CROP PRESENT
1	84	12	1121	5	104.0	1.78	LOW DENSITY, WITH TREES
1	65	2	1113	5	47.0	1.79	URBAN, HIGH DENSITY, WITH TREES
1	1	6	2120	3	17.0	1.79	ABANDONED
1	299	8	4300	4	15.0	1.81	MIXED
1	113	13	2113	5	16.0	1.84	GROWING CROP PRESENT
1	79	5	1110	3	19.0	1.86	SINGLE-FAMILY HOUSEHOLD UNITS
1	299	5	1123	5	21.0	1.87	HIGH DENSITY, WITH TREES
1	68	6	2120	4	16.0	1.95	ABANDONED
1	122	12	1300	4	40.0	1.98	INDUSTRIAL
1	65	5	1114	5	31.0	2.02	URBAN, HIGH DENSITY, WITHOUT TREES
1	127	1	1114	4	61.0	2.08	URBAN, HIGH DENSITY, WITHOUT TREES
1	122	4	1124	5	11.0	2.11	HIGH DENSITY, WITHOUT TREES
1	188	1	2113	5	81.0	2.26	GROWING CROP PRESENT
1	307	8	2113	5	29.0	2.33	GROWING CROP PRESENT
1	199	6	2113	5	17.0	2.35	GROWING CROP PRESENT
1	337	3	2120	5	10.0	2.70	ABANDONED
1	199	7	2113	5	19.0	2.85	GROWING CROP PRESENT

Table 24-2. Cluster Group #2

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
2	45	1	5410	5	395.0	0.69	BAYS
2	57	8	5410	5	189.0	1.18	BAYS
2	108	7	5120	5	55.0	1.23	MAN-MADE (CANALS, DITCHES, & AQUADUCTS)
2	96	1	5410	5	322.0	1.38	BAYS
2	98	6	5210	5	29.0	1.55	NATURAL LAKES & PONDS
2	44	1	5110	5	147.0	1.59	NATURAL (RIVERS & CREEKS)
2	95	4	5110	5	69.0	2.09	NATURAL (RIVERS & CREEKS)

Table 24-3 Cluster Group #3

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
3	316	1	7500	5	128.0	0.80	DISTURBED LAND
3	308	12	2142	4	21.0	1.16	IMPROVED
3	121	10	2143	4	37.0	1.34	UNIMPROVED

Table 24-4. Cluster Group #4

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
4	179	5	1913	5	35.0	0.56	PARK
4	202	5	2142	5	40.0	0.65	IMPROVED
4	84	4	2113	5	55.0	0.70	GROWING CROP PRESENT
4	57	7	1911	5	104.0	0.71	GOLF COURSES
4	64	10	1911	5	90.0	0.81	GOLF COURSES
4	308	11	2113	4	45.0	0.81	GROWING CROP PRESENT
4	355	1	2113	4	66.0	0.84	GROWING CROP PRESENT
4	190	9	2142	5	218.0	0.85	IMPROVED
4	7	10	1913	5	31.0	0.91	PARK
4	307	3	2142	5	23.0	0.92	IMPROVED
4	191	8	1220	5	57.0	0.96	RETAIL TRADE AREA (BUSINESS DIST.,SHOPPING CENT.,COMM.)
4	191	10	2142	5	76.0	1.08	IMPROVED
4	308	2	2142	5	95.0	1.08	IMPROVED
4	201	4	2142	5	17.0	1.09	IMPROVED
4	402	1	2120	5	57.0	1.12	ABANDONED
4	122	6	2120	5	20.0	1.16	ABANDONED
4	122	10	1911	4	58.0	1.16	GOLF COURSES
4	79	1	2113	5	58.0	1.27	GROWING CROP PRESENT
4	192	15	2113	5	39.0	1.29	GROWING CROP PRESENT
4	67	5	2113	4	20.0	1.29	GROWING CROP PRESENT
4	191	1	2113	5	25.0	1.30	GROWING CROP PRESENT
4	84	3	2113	5	59.0	1.39	GROWING CROP PRESENT
4	84	5	2113	5	28.0	1.41	GROWING CROP PRESENT
4	67	2	1911	4	18.0	1.41	GOLF COURSES
4	307	4	1112	5	28.0	1.47	RURAL, LOW DENSITY, WITHOUT TREES
4	201	1	2142	5	16.0	1.68	IMPROVED
4	199	4	2113	5	32.0	2.12	GROWING CROP PRESENT
4	121	9	2113	5	10.0	2.12	GROWING CROP PRESENT
4	265	4	2113	5	12.0	2.30	GROWING CROP PRESENT
4	307	1	2142	4	13.0	2.39	IMPROVED

Table 24-5. Cluster Group #5

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
5	126	5	2113	4	23.0	1.21	GROWING CROP PRESENT
5	126	16	1310	4	50.0	1.28	MECHANICAL PROCESSING
5	199	11	2113	5	21.0	1.30	GROWING CROP PRESENT
5	126	7	2113	4	22.0	1.32	GROWING CROP PRESENT
5	126	10	1911	5	81.0	1.49	GOLF COURSES
5	126	9	2113	5	37.0	1.59	GROWING CROP PRESENT
5	126	6	2113	4	30.0	2.43	GROWING CROP PRESENT
5	126	2	2113	4	13.0	2.73	GROWING CROP PRESENT
5	126	8	2113	5	30.0	2.87	GROWING CROP PRESENT

Table 24-6. Cluster Group #6

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
6	119	2	1914	5	31.0	0.68	PARKING LOTS
6	95	3	1640	5	44.0	0.68	MILITARY AREAS
6	8	2	1520	4	91.0	0.80	RAILROADS & ASSOCIATED FACILITIES
6	95	1	1210	5	85.0	0.89	WHOLESALE TRADE AREAS
6	107	1	1410	5	127.0	0.97	STONE QUARRIES
6	57	5	1644	4	14.0	1.04	TRAINING AREAS
6	54	3	1220	4	18.0	1.06	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
6	122	1	1220	5	42.0	1.07	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
6	97	2	1300	5	66.0	1.08	INDUSTRIAL
6	109	5	1330	5	29.0	1.11	CHEMICAL PROCESSING
6	109	6	1330	5	29.0	1.17	CHEMICAL PROCESSING
6	84	13	1210	5	43.0	1.35	WHOLESALE TRADE AREAS
6	99	3	1531	5	66.0	1.35	COMMERCIAL
6	115	2	1220	5	56.0	1.38	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
6	113	12	1124	4	12.0	1.39	HIGH DENSITY, WITHOUT TREES
6	57	6	1246	4	22.0	1.39	RECREATIONAL (TENNIS,ICE SKATING,STABLES,PLAY AREAS...)
6	64	8	1660	5	12.0	1.44	GOVERNMENT & ADMINISTRATIVE OFFICES
6	64	11	1641	5	22.0	1.47	HOUSING
6	97	6	1420	5	40.0	1.54	SAND & GRAVEL PITS
6	99	1	1210	4	24.0	1.74	WHOLESALE TRADE AREAS
6	300	3	1660	5	16.0	1.94	GOVERNMENT & ADMINISTRATIVE OFFICES

Table 24-7 Cluster Group #7

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
7	124	4	1113	5	303.0	0.54	URBAN, HIGH DENSITY, WITH TREES
7	1	9	1110	5	81.0	0.67	SINGLE-FAMILY HOUSEHOLD UNITS
7	179	12	1912	5	81.0	0.70	CEMETERIES
7	124	2	1113	4	200.0	0.70	URBAN, HIGH DENSITY, WITH TREES
7	179	1	1911	5	83.0	0.86	GOLF COURSES
7	308	3	2142	4	19.0	0.93	IMPROVED
7	57	4	1246	3	11.0	1.24	RECREATIONAL (TENNIS,ICE SKATING,STABLES,PLAY AREAS...)
7	1	8	2113	3	11.0	1.31	GROWING CROP PRESENT
7	355	3	2113	4	22.0	1.32	GROWING CROP PRESENT
7	9	5	2120	4	32.0	1.50	ABANDONED
7	57	10	1124	4	12.0	1.62	HIGH DENSITY, WITHOUT TREES
7	7	7	5300	5	54.0	1.67	RESERVOIRS
7	98	2	5210	4	19.0	1.75	NATURAL LAKES & PONDS

Table 24-8 Cluster Group #8

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
8	109	11	1643	4	54.0	0.62	STORAGE AREAS
8	8	1	1210	5	48.0	0.73	WHOLESALE TRADE AREAS
8	65	6	1560	5	70.0	0.84	ELECTRIC, GAS, WATER, SEWAGE DISPOSAL, SOLID WASTE, UTIL
8	179	10	1220	5	49.0	0.84	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
8	8	5	1220	5	58.0	0.87	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
8	6	5	1220	4	21.0	0.88	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
8	113	6	1653	5	37.0	0.90	STATE
8	8	4	1120	5	64.0	0.94	MULTI-FAMILY HOUSEHOLD UNITS
8	99	2	1300	5	48.0	0.99	INDUSTRIAL
8	95	5	1210	5	37.0	1.06	WHOLESALE TRADE AREAS
8	8	3	1120	5	38.0	1.11	MULTI-FAMILY HOUSEHOLD UNITS
8	108	3	1913	5	75.0	1.15	PARK
8	118	6	1220	3	30.0	1.18	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
8	57	9	1124	5	20.0	1.21	HIGH DENSITY, WITHOUT TREES
8	118	7	1643	4	10.0	1.23	STORAGE AREAS
8	54	2	1300	5	22.0	1.28	INDUSTRIAL
8	109	2	1621	4	17.0	1.29	HOSPITALS
8	64	9	1914	5	48.0	1.32	PARKING LOTS
8	119	5	1914	5	42.0	1.33	PARKING LOTS
8	68	16	1110	5	27.0	1.33	SINGLE-FAMILY HOUSEHOLD UNITS
8	118	9	1340	5	18.0	1.35	FABRICATION & ASSEMBLY
8	7	6	1611	4	10.0	1.40	PRIMARY
8	119	9	1220	4	34.0	1.42	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
8	7	1	1120	5	33.0	1.44	MULTI-FAMILY HOUSEHOLD UNITS
8	97	4	1124	4	16.0	1.44	HIGH DENSITY, WITHOUT TREES
8	157	3	2113	4	31.0	1.51	GROWING CROP PRESENT
8	119	4	1660	5	12.0	1.62	GOVERNMENT & ADMINISTRATIVE OFFICES
8	85	5	1124	4	15.0	1.69	HIGH DENSITY, WITHOUT TREES
8	7	5	1110	5	54.0	1.83	SINGLE-FAMILY HOUSEHOLD UNITS
8	157	2	1310	5	13.0	1.91	MECHANICAL PROCESSING
8	179	11	1124	5	31.0	2.10	HIGH DENSITY, WITHOUT TREES

Table 24-9. Cluster Group #9

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
9	126	3	2113	3	13.0	0.00	GROWING CROP PRESENT
9	126	4	2113	3	13.0	2.68	GROWING CROP PRESENT

Table 24-10. Cluster Group #10

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
10	6	1	1612	4	27.0	0.67	SECONDARY
10	67	1	1621	5	20.0	0.68	HOSPITALS
10	118	5	1110	4	73.0	0.75	SINGLE-FAMILY HOUSEHOLD UNITS
10	119	11	1123	5	85.0	0.77	HIGH DENSITY, WITH TREES
10	119	8	1113	5	137.0	0.79	URBAN, HIGH DENSITY, WITH TREES
10	84	10	1121	5	114.0	0.81	LOW DENSITY, WITH TREES
10	7	11	1912	5	55.0	0.85	CEMETERIES
10	85	2	1124	5	35.0	0.87	HIGH DENSITY, WITHOUT TREES
10	119	6	1912	5	53.0	1.01	CEMETERIES
10	85	7	1113	5	194.0	1.04	URBAN, HIGH DENSITY, WITH TREES
10	124	3	1110	4	88.0	1.06	SINGLE-FAMILY HOUSEHOLD UNITS
10	97	7	1113	5	156.0	1.09	URBAN, HIGH DENSITY, WITH TREES
10	186	2	1124	5	26.0	1.09	HIGH DENSITY, WITHOUT TREES
10	9	4	2120	3	19.0	1.11	ABANDONED
10	6	2	1124	5	154.0	1.11	HIGH DENSITY, WITHOUT TREES
10	68	10	1110	5	39.0	1.12	SINGLE-FAMILY HOUSEHOLD UNITS
10	98	4	1112	5	65.0	1.13	RURAL, LOW DENSITY, WITHOUT TREES
10	118	12	1912	5	101.0	1.14	CEMETERIES
10	118	4	1113	4	63.0	1.14	URBAN, HIGH DENSITY, WITH TREES
10	97	5	1113	5	100.0	1.14	URBAN, HIGH DENSITY, WITH TREES
10	45	3	2120	5	25.0	1.17	ABANDONED
10	117	5	4200	5	12.0	1.22	EVERGREEN (CONIFEROUS & OTHER)
10	84	9	1113	5	23.0	1.22	URBAN, HIGH DENSITY, WITH TREES
10	7	2	1615	5	50.0	1.25	UNIVERSITY
10	299	3	1113	5	17.0	1.28	URBAN, HIGH DENSITY, WITH TREES
10	64	1	1122	5	39.0	1.31	LOW DENSITY, WITHOUT TREES
10	65	1	1113	5	38.0	1.31	URBAN, HIGH DENSITY, WITH TREES
10	45	2	1113	5	178.0	1.32	URBAN, HIGH DENSITY, WITH TREES
10	179	9	1124	4	29.0	1.38	HIGH DENSITY, WITHOUT TREES
10	299	7	1124	3	13.0	1.40	HIGH DENSITY, WITHOUT TREES
10	85	3	1124	5	19.0	1.46	HIGH DENSITY, WITHOUT TREES
10	157	1	2113	4	26.0	1.49	GROWING CROP PRESENT
10	57	1	1124	5	10.0	1.50	HIGH DENSITY, WITHOUT TREES
10	299	1	1912	5	64.0	1.51	CEMETERIES
10	41	2	1210	4	11.0	1.52	WHOLESALE TRADE AREAS
10	85	6	1124	4	21.0	1.53	HIGH DENSITY, WITHOUT TREES
10	68	5	2113	4	36.0	1.59	GROWING CROP PRESENT
10	6	7	1610	4	12.0	1.62	EDUCATIONAL FACILITIES
10	84	6	2112	4	33.0	1.74	BARE--RECENTLY PLOWED
10	2	3	2250	4	15.0	1.77	TURF FARM
10	2	2	1300	4	20.0	1.78	INDUSTRIAL
10	68	1	1220	3	12.0	1.84	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
10	117	7	4300	4	10.0	1.92	MIXED
10	299	4	1110	5	16.0	2.33	SINGLE-FAMILY HOUSEHOLD UNITS
10	299	2	1113	5	30.0	2.34	URBAN, HIGH DENSITY, WITH TREES

Table 24-11. Cluster Group #11

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
11	402	3	2142	5	61.0	0.56	IMPROVED
11	344	3	2113	5	42.0	0.59	GROWING CROP PRESENT
11	402	4	2120	5	39.0	0.59	ABANDONED
11	188	3	2143	5	41.0	0.59	UNIMPROVED
11	14	1	1110	5	143.0	0.61	SINGLE-FAMILY HOUSEHOLD UNITS
11	365	2	2142	5	55.0	0.70	IMPROVED
11	308	6	2142	4	36.0	0.71	IMPROVED
11	12	1	2113	5	63.0	0.75	GROWING CROP PRESENT
11	192	5	2142	5	38.0	0.77	IMPROVED
11	355	4	1113	4	199.0	0.77	URBAN, HIGH DENSITY, WITH TREES
11	201	5	2142	5	100.0	0.77	IMPROVED
11	366	3	2113	5	48.0	0.78	GROWING CROP PRESENT
11	13	1	2120	5	96.0	0.79	ABANDONED
11	192	14	1911	5	126.0	0.80	GOLF COURSES
11	341	5	2142	5	15.0	0.80	IMPROVED
11	200	2	1114	5	185.0	0.80	URBAN, HIGH DENSITY, WITHOUT TREES
11	13	2	2113	3	22.0	0.81	GROWING CROP PRESENT
11	344	4	2120	4	25.0	0.87	ABANDONED
11	54	1	2113	4	28.0	0.89	GROWING CROP PRESENT
11	36	1	2113	3	43.0	0.89	GROWING CROP PRESENT
11	337	4	2240	5	93.0	0.89	NURSERIES AND FLORICULTURAL AREAS
11	344	2	1911	5	85.0	0.90	GOLF COURSES
11	352	6	2113	5	31.0	0.91	GROWING CROP PRESENT
11	107	4	2142	5	87.0	0.92	IMPROVED
11	201	2	2142	5	38.0	0.92	IMPROVED
11	13	5	2120	5	27.0	0.95	ABANDONED
11	340	11	2113	5	25.0	0.97	GROWING CROP PRESENT
11	1	10	2113	3	14.0	0.97	GROWING CROP PRESENT
11	36	2	1112	4	14.0	0.98	RURAL, LOW DENSITY, WITHOUT TREES
11	352	3	2142	5	64.0	0.98	IMPROVED
11	375	2	2142	5	21.0	0.98	IMPROVED
11	191	4	1113	5	130.0	0.99	URBAN, HIGH DENSITY, WITH TREES
11	86	6	2112	5	67.0	1.00	BARE--RECENTLY PLOWED
11	202	1	1114	5	136.0	1.00	URBAN, HIGH DENSITY, WITHOUT TREES
11	118	10	1911	5	55.0	1.01	GOLF COURSES
11	1	5	2113	5	46.0	1.01	GROWING CROP PRESENT
11	308	5	2142	4	30.0	1.03	IMPROVED
11	340	2	2113	5	19.0	1.04	GROWING CROP PRESENT
11	190	1	2142	4	30.0	1.05	IMPROVED
11	402	5	1114	5	172.0	1.07	URBAN, HIGH DENSITY, WITHOUT TREES
11	337	5	2113	5	24.0	1.08	GROWING CROP PRESENT
11	104	6	1113	4	13.0	1.08	URBAN, HIGH DENSITY, WITH TREES
11	340	10	2143	5	31.0	1.08	UNIMPROVED
11	104	7	2142	5	77.0	1.08	IMPROVED
11	6	6	1911	5	77.0	1.08	GOLF COURSES
11	202	3	2142	5	42.0	1.08	IMPROVED

Table 24-11 (Continued)

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
11	308	1	2113	5	35.0	1.09	GROWING CROP PRESENT
11	307	5	1112	4	30.0	1.12	RURAL, LOW DENSITY, WITHOUT TREES
11	366	2	2142	5	20.0	1.12	IMPROVED
11	402	8	2113	5	48.0	1.15	GROWING CROP PRESENT
11	352	9	2113	5	73.0	1.15	GROWING CROP PRESENT
11	308	8	2113	4	94.0	1.18	GROWING CROP PRESENT
11	352	5	2113	5	28.0	1.18	GROWING CROP PRESENT
11	65	4	2113	5	94.0	1.18	GROWING CROP PRESENT
11	366	5	2113	5	25.0	1.19	GROWING CROP PRESENT
11	192	4	2120	5	23.0	1.20	ABANDONED
11	352	7	2113	5	21.0	1.21	GROWING CROP PRESENT
11	13	9	2113	4	49.0	1.22	GROWING CROP PRESENT
11	191	3	1113	5	64.0	1.22	URBAN, HIGH DENSITY, WITH TREES
11	67	7	2113	3	15.0	1.27	GROWING CROP PRESENT
11	192	6	2113	5	21.0	1.27	GROWING CROP PRESENT
11	340	7	2142	5	46.0	1.28	IMPROVED
11	340	3	2113	5	27.0	1.30	GROWING CROP PRESENT
11	25	4	2113	4	27.0	1.30	GROWING CROP PRESENT
11	340	1	2120	5	26.0	1.31	ABANDONED
11	104	4	2113	5	43.0	1.32	GROWING CROP PRESENT
11	308	4	2142	4	33.0	1.32	IMPROVED
11	2	1	2113	4	26.0	1.35	GROWING CROP PRESENT
11	366	4	2113	5	74.0	1.35	GROWING CROP PRESENT
11	188	5	2113	4	25.0	1.35	GROWING CROP PRESENT
11	200	3	1113	5	54.0	1.36	URBAN, HIGH DENSITY, WITH TREES
11	45	4	2113	5	25.0	1.37	GROWING CROP PRESENT
11	366	14	2113	4	37.0	1.38	GROWING CROP PRESENT
11	12	2	2113	5	44.0	1.38	GROWING CROP PRESENT
11	190	11	1112	4	13.0	1.38	RURAL, LOW DENSITY, WITHOUT TREES
11	402	9	1113	5	21.0	1.39	URBAN, HIGH DENSITY, WITH TREES
11	35	2	2113	3	10.0	1.39	GROWING CROP PRESENT
11	366	15	2113	4	20.0	1.39	GROWING CROP PRESENT
11	9	1	2113	4	19.0	1.40	GROWING CROP PRESENT
11	337	8	2113	5	24.0	1.41	GROWING CROP PRESENT
11	121	1	2113	3	14.0	1.41	GROWING CROP PRESENT
11	352	10	2142	5	23.0	1.42	IMPROVED
11	187	2	2120	5	88.0	1.43	ABANDONED
11	13	4	2142	5	43.0	1.44	IMPROVED
11	353	6	1114	4	41.0	1.44	URBAN, HIGH DENSITY, WITHOUT TREES
11	113	11	1113	4	88.0	1.48	URBAN, HIGH DENSITY, WITH TREES
11	202	2	2113	5	69.0	1.48	GROWING CROP PRESENT
11	13	3	2113	5	39.0	1.49	GROWING CROP PRESENT
11	107	3	2142	4	27.0	1.51	IMPROVED
11	192	7	2142	5	48.0	1.51	IMPROVED
11	1	11	1110	3	14.0	1.51	SINGLE-FAMILY HOUSEHOLD UNITS
11	104	3	2113	5	30.0	1.51	GROWING CROP PRESENT

Table 24-11 (Continued)

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
11	190	6	2113	4	30.0	1.51	GROWING CROP PRESENT
11	188	4	2142	5	25.0	1.53	IMPROVED
11	265	1	2113	5	37.0	1.55	GROWING CROP PRESENT
11	341	2	2113	4	11.0	1.55	GROWING CROP PRESENT
11	341	4	2113	5	13.0	1.56	GROWING CROP PRESENT
11	36	5	2142	5	39.0	1.56	IMPROVED
11	375	5	2120	5	15.0	1.57	ABANDONED
11	340	8	2113	5	24.0	1.58	GROWING CROP PRESENT
11	364	3	2114	3	12.0	1.58	HARVESTED
11	266	3	4100	5	20.0	1.59	DECIDUOUS
11	366	11	2113	5	24.0	1.60	GROWING CROP PRESENT
11	192	11	2113	5	40.0	1.61	GROWING CROP PRESENT
11	341	3	2113	5	22.0	1.62	GROWING CROP PRESENT
11	113	3	2113	4	16.0	1.62	GROWING CROP PRESENT
11	375	8	2142	5	24.0	1.63	IMPROVED
11	201	7	2142	5	22.0	1.65	IMPROVED
11	104	2	2120	5	17.0	1.65	ABANDONED
11	127	10	2142	4	12.0	1.66	IMPROVED
11	352	11	2113	5	17.0	1.66	GROWING CROP PRESENT
11	129	4	6100	5	15.0	1.66	VEGETATED WETLANDS
11	191	5	1114	5	24.0	1.68	URBAN, HIGH DENSITY, WITHOUT TREES
11	364	2	2142	4	15.0	1.72	IMPROVED
11	366	9	2113	5	65.0	1.73	GROWING CROP PRESENT
11	13	8	2250	3	19.0	1.77	TURF FARM
11	192	10	1246	5	13.0	1.79	RECREATIONAL (TENNIS,ICE SKATING,STABLES,PLAY AREAS...)
11	366	6	2142	5	12.0	1.81	IMPROVED
11	41	1	1112	4	15.0	1.82	RURAL, LOW DENSITY, WITHOUT TREES
11	188	2	2142	5	22.0	1.82	IMPROVED
11	192	1	2113	5	69.0	1.83	GROWING CROP PRESENT
11	191	7	2143	5	12.0	1.83	UNIMPROVED
11	192	12	2113	5	17.0	1.83	GROWING CROP PRESENT
11	366	13	2142	4	20.0	1.84	IMPROVED
11	127	5	2120	4	18.0	1.86	ABANDONED
11	25	3	2113	4	14.0	1.87	GROWING CROP PRESENT
11	192	3	2113	5	74.0	1.89	GROWING CROP PRESENT
11	68	4	2113	3	13.0	1.91	GROWING CROP PRESENT
11	104	5	2142	4	17.0	1.92	IMPROVED
11	265	2	2113	5	10.0	1.92	GROWING CROP PRESENT
11	122	5	1611	3	12.0	2.05	PRIMARY
11	35	1	2113	3	18.0	2.05	GROWING CROP PRESENT
11	190	2	1112	4	14.0	2.07	RURAL, LOW DENSITY, WITHOUT TREES
11	366	12	2113	5	17.0	2.08	GROWING CROP PRESENT
11	67	6	2142	3	14.0	2.08	IMPROVED
11	307	11	2113	5	14.0	2.09	GROWING CROP PRESENT
11	375	1	2113	5	10.0	2.11	GROWING CROP PRESENT
11	375	9	2113	5	10.0	2.11	GROWING CROP PRESENT

Table 24-11 (Continued)

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
11	366	1	2120	5	32.0	2.12	ABANDONED
11	188	6	2143	5	23.0	2.13	UNIMPROVED
11	307	2	2142	5	20.0	2.17	IMPROVED
11	190	10	2142	5	80.0	2.19	IMPROVED
11	366	7	2142	5	10.0	2.21	IMPROVED
11	337	9	2113	5	10.0	2.25	GROWING CROP PRESENT
11	337	10	2113	5	12.0	2.28	GROWING CROP PRESENT
11	35	3	2142	4	14.0	2.30	IMPROVED
11	192	13	2113	5	24.0	2.34	GROWING CROP PRESENT
11	353	2	1110	4	12.0	2.35	SINGLE-FAMILY HOUSEHOLD UNITS
11	402	2	2113	5	12.0	2.39	GROWING CROP PRESENT
11	168	3	2113	3	18.0	2.45	GROWING CROP PRESENT
11	402	7	2113	5	12.0	2.48	GROWING CROP PRESENT
11	201	3	2142	5	11.0	2.51	IMPROVED
11	199	5	2113	5	16.0	2.59	GROWING CROP PRESENT
11	199	2	2113	5	13.0	2.61	GROWING CROP PRESENT

Table 24-12. Cluster Group #12

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
12	100	1	6000	5	74.0	0.90	WETLANDS
12	111	1	6100	5	48.0	1.26	VEGETATED WETLANDS
12	199	1	2113	5	43.0	1.41	GROWING CROP PRESENT
12	111	2	6100	5	48.0	1.65	VEGETATED WETLANDS
12	100	2	6000	5	30.0	1.72	WETLANDS
12	111	4	6100	5	55.0	1.89	VEGETATED WETLANDS
12	199	12	2113	5	16.0	2.53	GROWING CROP PRESENT

Table 24-13. Cluster Group #13

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
13	102	2	6110	4	23.0	0.00	BRACKISH MARSH
13	123	3	6000	5	20.0	2.22	WETLANDS

Table 24-14. Cluster Group #14

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
14	109	10	1300	5	184.0	0.56	INDUSTRIAL
14	109	9	1520	5	38.0	0.66	RAILROADS & ASSOCIATED FACILITIES
14	109	7	1520	5	49.0	0.68	RAILROADS & ASSOCIATED FACILITIES
14	108	2	1300	5	84.0	0.70	INDUSTRIAL
14	109	1	1520	4	37.0	0.86	RAILROADS & ASSOCIATED FACILITIES
14	108	8	1544	5	99.0	0.93	PART FACILITIES
14	109	8	1300	5	25.0	0.97	INDUSTRIAL
14	94	1	1520	5	37.0	0.98	RAILROADS & ASSOCIATED FACILITIES
14	7	3	1243	5	31.0	1.14	SPORTS (STADIUMS,ARENAS,RACETRACKS,OTHER)
14	109	9	1914	5	31.0	1.23	PARKING LOTS
14	7	9	1120	5	158.0	1.26	MULTI-FAMILY HOUSEHOLD UNITS
14	108	5	1544	4	32.0	1.61	PART FACILITIES
14	108	4	1200	5	40.0	1.71	COMMERCIAL & SERVICES
14	108	6	1544	4	14.0	2.19	PART FACILITIES

Table 24-15. Cluster Group #15

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
15	35	4	1330	4	18.0	0.77	CHEMICAL PROCESSING
15	41	4	1420	4	18.0	0.88	SAND & GRAVEL PITS
15	124	1	1114	4	59.0	1.02	URBAN, HIGH DENSITY, WITHOUT TREES
15	307	9	2240	5	15.0	1.04	NURSERIES AND FLORICULTURAL AREAS
15	64	7	1641	4	17.0	1.40	HOUSING
15	192	2	2113	5	69.0	1.46	GROWING CROP PRESENT
15	359	6	7500	5	62.0	1.53	DISTURBED LAND
15	188	7	2112	4	11.0	1.84	BARE--RECENTLY PLOWED
15	307	7	2240	5	29.0	2.22	NURSERIES AND FLORICULTURAL AREAS

Table 24-16. Cluster Group #16

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
16	191	9	2113	5	20.0	0.00	GROWING CROP PRESENT

Table 24-17. Cluster Group #17

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
17	202	7	2142	5	46.0	0.59	IMPROVED
17	63	1	1615	5	171.0	0.80	UNIVERSITY
17	9	3	4100	5	91.0	0.88	DECIDUOUS
17	190	8	1111	5	33.0	0.89	RURAL, LOW DENSITY, WITH TREES
17	365	1	4100	5	180.0	0.94	DECIDUOUS
17	316	3	1113	5	92.0	0.97	URBAN, HIGH DENSITY, WITH TREES
17	119	7	1113	5	37.0	0.99	URBAN, HIGH DENSITY, WITH TREES
17	402	6	1113	5	28.0	0.99	URBAN, HIGH DENSITY, WITH TREES
17	316	2	1113	5	12.0	1.00	URBAN, HIGH DENSITY, WITH TREES
17	340	6	4100	5	31.0	1.03	DECIDUOUS
17	122	13	4100	5	24.0	1.07	DECIDUOUS
17	202	8	2142	5	31.0	1.10	IMPROVED
17	340	12	4100	5	24.0	1.14	DECIDUOUS
17	359	5	[illegible]300	5	35.0	1.20	MIXED
17	179	8	4100	5	27.0	1.25	DECIDUOUS
17	110	6	2113	5	36.0	1.26	GROWING CROP PRESENT
17	110	5	2113	5	28.0	1.30	GROWING CROP PRESENT
17	200	5	4400	5	26.0	1.31	UPLAND SHRUBS
17	201	9	2142	5	12.0	1.33	IMPROVED
17	340	5	2113	5	17.0	1.35	GROWING CROP PRESENT
17	67	8	4100	5	84.0	1.35	DECIDUOUS
17	202	4	1220	5	16.0	1.36	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
17	107	2	4100	5	42.0	1.39	DECIDUOUS
17	121	3	2142	5	13.0	1.39	IMPROVED
17	117	8	4100	5	42.0	1.41	DECIDUOUS
17	337	6	4100	5	20.0	1.42	DECIDUOUS
17	366	6	4100	5	36.0	1.48	DECIDUOUS
17	364	1	4100	5	20.0	1.50	DECIDUOUS
17	202	6	4400	5	25.0	1.51	UPLAND SHRUBS
17	179	7	4100	5	20.0	1.53	DECIDUOUS
17	67	3	4100	5	49.0	1.54	DECIDUOUS
17	98	3	4300	5	85.0	1.56	MIXED
17	200	4	4400	5	25.0	1.56	UPLAND SHRUBS
17	352	2	2113	5	25.0	1.57	GROWING CROP PRESENT
17	98	5	4100	5	36.0	1.62	DECIDUOUS
17	352	4	2113	5	10.0	1.63	GROWING CROP PRESENT
17	352	8	4100	5	27.0	1.64	DECIDUOUS
17	308	10	4100	5	72.0	1.72	DECIDUOUS
17	337	7	4100	5	21.0	1.72	DECIDUOUS
17	121	8	4100	5	20.0	1.74	DECIDUOUS
17	9	6	2120	4	26.0	1.74	ABANDONED
17	386	1	4100	5	30.0	1.76	DECIDUOUS
17	84	7	4100	5	44.0	1.79	DECIDUOUS
17	308	9	4100	5	29.0	1.85	DECIDUOUS
17	84	2	2112	5	25.0	1.91	BARE--RECENTLY PLOWED
17	113	1	2113	5	11.0	1.93	GROWING CROP PRESENT

Table 24-17 (Continued)

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
17	113	2	2113	3	10.0	1.94	GROWING CROP PRESENT
17	129	3	4300	5	18.0	2.10	MIXED
17	299	9	4100	5	20.0	2.13	DECIDUOUS
17	190	7	4100	5	20.0	2.15	DECIDUOUS
17	186	1	6200	5	61.0	2.19	FORESTED WETLANDS
17	386	2	4100	5	30.0	2.20	DECIDUOUS
17	122	15	4100	5	12.0	2.24	DECIDUOUS
17	308	7	2113	4	21.0	2.36	GROWING CROP PRESENT
17	117	2	2113	4	22.0	2.45	GROWING CROP PRESENT
17	135	1	6200	5	318.0	2.47	FORESTED WETLANDS
17	192	8	2113	5	13.0	2.58	GROWING CROP PRESENT

Table 24-18. Cluster Group #18

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
18	168	2	4200	5	33.0	0.81	EVERGREEN (CONIFEROUS & OTHER)
18	168	6	4300	5	110.0	0.93	MIXED
18	123	2	4300	5	42.0	1.19	MIXED
18	126	13	2113	4	21.0	1.34	GROWING CROP PRESENT
18	126	12	2113	3	32.0	1.38	GROWING CROP PRESENT
18	123	1	4300	5	41.0	1.81	MIXED
18	126	11	2113	3	16.0	2.58	GROWING CROP PRESENT
18	126	15	2113	3	11.0	2.95	GROWING CROP PRESENT

Table 24-19. Cluster Group #19

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
19	199	9	2120	5	11.0	0.00	ABANDONED
19	199	10	2113	5	23.0	2.06	GROWING CROP PRESENT

Table 24-20. Cluster Group #20

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
20	124	5	4100	5	16.0	0.00	DECIDUOUS
20	121	7	4100	5	25.0	2.28	DECIDUOUS
20	102	1	6110	5	179.0	2.64	BRACKISH MARSH

Table 24-21. Cluster Group #21

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
21	75	1	4200	5	35.0	0.71	EVERGREEN (CONIFEROUS & OTHER)
21	75	4	4200	5	30.0	1.02	EVERGREEN (CONIFEROUS & OTHER)
21	75	2	4200	4	11.0	1.11	EVERGREEN (CONIFEROUS & OTHER)
21	123	4	6000	5	25.0	2.34	WETLANDS

Table 24-22. Cluster Group #22

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
22	168	1	5110	5	64.0	0.73	NATURAL (RIVERS & CREEKS)
22	129	1	5110	5	16.0	1.08	NATURAL (RIVERS & CREEKS)
22	111	3	5110	5	54.0	1.25	NATURAL (RIVERS & CREEKS)
22	117	12	5210	5	30.0	1.35	NATURAL LAKES & PONDS
22	129	2	5110	5	16.0	1.51	NATURAL (RIVERS & CREEKS)
22	135	2	5210	5	83.0	1.71	NATURAL LAKES & PONDS

Table 24-23. Cluster Group #23

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
23	45	5	2113	3	33.0	0.72	GROWING CROP PRESENT
23	117	11	6100	5	16.0	1.05	VEGETATED WETLANDS
23	199	3	2113	5	18.0	1.48	GROWING CROP PRESENT
23	117	10	6100	5	16.0	1.48	VEGETATED WETLANDS
23	168	4	2112	5	78.0	2.01	BARE--RECENTLY PLOWED
23	168	5	2112	4	15.0	2.11	BARE--RECENTLY PLOWED

Table 25. 23 Groups Established Using a Threshold of 3.0 and Ordered by Land Use Category

Table 25-1. Cluster Group #1

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
1	79	5	1110	3	19.0	1.86	SINGLE-FAMILY HOUSEHOLD UNITS
1	1	2	1110	3	35.0	1.02	SINGLE-FAMILY HOUSEHOLD UNITS
1	353	3	1110	4	10.0	1.24	SINGLE-FAMILY HOUSEHOLD UNITS
1	353	4	1110	4	20.0	1.60	SINGLE-FAMILY HOUSEHOLD UNITS
1	127	2	1110	4	57.0	0.84	SINGLE-FAMILY HOUSEHOLD UNITS
1	115	4	1110	5	88.0	1.35	SINGLE-FAMILY HOUSEHOLD UNITS
1	186	3	1110	5	201.0	1.28	SINGLE-FAMILY HOUSEHOLD UNITS
1	122	7	1110	5	24.0	1.68	SINGLE-FAMILY HOUSEHOLD UNITS
1	115	3	1110	5	33.0	1.26	SINGLE-FAMILY HOUSEHOLD UNITS
1	79	4	1110	5	38.0	1.02	SINGLE-FAMILY HOUSEHOLD UNITS
1	127	7	1110	5	42.0	1.31	SINGLE-FAMILY HOUSEHOLD UNITS
1	118	2	1110	5	52.0	1.04	SINGLE-FAMILY HOUSEHOLD UNITS
1	115	8	1110	5	67.0	0.89	SINGLE-FAMILY HOUSEHOLD UNITS
1	115	1	1112	5	198.0	1.15	RURAL, LOW DENSITY, WITHOUT TREES
1	64	6	1112	4	32.0	1.02	RURAL, LOW DENSITY, WITHOUT TREES
1	64	2	1112	5	17.0	1.46	RURAL, LOW DENSITY, WITHOUT TREES
1	65	2	1113	5	47.0	1.79	URBAN, HIGH DENSITY, WITH TREES
1	118	11	1113	4	32.0	1.28	URBAN, HIGH DENSITY, WITH TREES
1	355	2	1113	4	45.0	1.24	URBAN, HIGH DENSITY, WITH TREES
1	84	8	1113	5	38.0	1.43	URBAN, HIGH DENSITY, WITH TREES
1	97	3	1114	4	10.0	1.57	URBAN, HIGH DENSITY, WITHOUT TREES
1	110	2	1114	5	158.0	0.89	URBAN, HIGH DENSITY, WITHOUT TREES
1	118	1	1114	5	48.0	1.14	URBAN, HIGH DENSITY, WITHOUT TREES
1	122	3	1114	5	53.0	0.95	URBAN, HIGH DENSITY, WITHOUT TREES
1	86	3	1114	5	68.0	1.39	URBAN, HIGH DENSITY, WITHOUT TREES
1	86	5	1114	5	74.0	1.41	URBAN, HIGH DENSITY, WITHOUT TREES
1	65	3	1114	5	81.0	1.03	URBAN, HIGH DENSITY, WITHOUT TREES
1	110	4	1114	4	16.0	1.61	URBAN, HIGH DENSITY, WITHOUT TREES
1	127	3	1114	4	24.0	1.05	URBAN, HIGH DENSITY, WITHOUT TREES
1	127	1	1114	4	61.0	2.08	URBAN, HIGH DENSITY, WITHOUT TREES
1	65	5	1114	5	31.0	2.02	URBAN, HIGH DENSITY, WITHOUT TREES
1	200	1	1114	5	38.0	1.15	URBAN, HIGH DENSITY, WITHOUT TREES
1	66	1	1114	5	241.0	1.18	URBAN, HIGH DENSITY, WITHOUT TREES
1	353	7	1120	4	18.0	1.60	MULTI-FAMILY HOUSEHOLD UNITS
1	122	2	1120	5	16.0	1.62	MULTI-FAMILY HOUSEHOLD UNITS
1	68	14	1120	5	27.0	0.88	MULTI-FAMILY HOUSEHOLD UNITS
1	115	6	1120	5	42.0	1.28	MULTI-FAMILY HOUSEHOLD UNITS
1	84	12	1121	5	104.0	1.78	LOW DENSITY, WITH TREES
1	64	3	1122	5	45.0	1.07	LOW DENSITY, WITHOUT TREES
1	64	4	1122	5	62.0	0.70	LOW DENSITY, WITHOUT TREES
1	299	5	1123	5	21.0	1.87	HIGH DENSITY, WITH TREES
1	115	5	1124	5	23.0	1.69	HIGH DENSITY, WITHOUT TREES
1	113	9	1124	5	41.0	0.76	HIGH DENSITY, WITHOUT TREES
1	186	4	1124	5	18.0	1.27	HIGH DENSITY, WITHOUT TREES
1	179	3	1124	5	16.0	0.52	HIGH DENSITY, WITHOUT TREES
1	122	4	1124	5	11.0	2.11	HIGH DENSITY, WITHOUT TREES

Table 25-1 (Continued)

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
1	113	8	1124	5	14.0	1.00	HIGH DENSITY, WITHOUT TREES
1	359	1	1150	5	16.0	1.03	MOBILE HOME PARKS OR COURTS
1	316	5	1150	4	21.0	0.96	MOBILE HOME PARKS OR COURTS
1	117	4	1150	3	16.0	1.46	MOBILE HOME PARKS OR COURTS
1	1	7	1210	3	19.0	1.50	WHOLESALE TRADE AREAS
1	359	4	1220	4	25.0	0.98	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
1	353	5	1230	4	96.0	1.01	BUSINESS, PROFESSIONAL, & PERSONAL SERVICES
1	186	6	1243	5	38.0	1.26	SPORTS (STADIUMS,ARENAS,RACETRACKS,OTHER)
1	122	12	1300	4	40.0	1.98	INDUSTRIAL
1	84	11	1420	5	58.0	0.82	SAND & GRAVEL PITS
1	119	3	1511	5	83.0	1.20	HIGHWAYS
1	64	5	1511	4	19.0	1.56	HIGHWAYS
1	79	2	1532	4	46.0	1.01	GENERAL
1	64	12	1533	5	46.0	1.12	MILITARY
1	185	5	1560	5	12.0	1.67	ELECTRIC, GAS, WATER, SEWAGE DISPOSAL, SOLID WASTE, UTIL
1	86	4	1610	4	12.0	1.67	EDUCATIONAL FACILITIES
1	85	4	1612	4	15.0	1.49	SECONDARY
1	6	8	1614	4	19.0	1.54	COLLEGE
1	64	14	1616	5	25.0	0.89	OTHER
1	25	5	1621	5	46.0	1.14	HOSPITALS
1	95	2	1643	4	348.0	0.69	STORAGE AREAS
1	359	7	1650	5	51.0	1.00	CORRECTIONAL
1	85	1	1660	5	130.0	0.91	GOVERNMENT & ADMINISTRATIVE OFFICES
1	359	3	1912	5	57.0	1.23	CEMETERIES
1	109	4	1912	4	23.0	1.14	CEMETERIES
1	7	12	1912	5	73.0	0.88	CEMETERIES
1	106	1	1912	5	50.0	1.23	CEMETERIES
1	8	6	1912	5	75.0	0.65	CEMETERIES
1	109	3	1912	4	29.0	0.81	CEMETERIES
1	192	9	2113	5	15.0	1.72	GROWING CROP PRESENT
1	191	2	2113	5	16.0	0.96	GROWING CROP PRESENT
1	113	13	2113	5	18.0	1.84	GROWING CROP PRESENT
1	199	6	2113	5	17.0	2.35	GROWING CROP PRESENT
1	307	8	2113	5	29.0	2.33	GROWING CROP PRESENT
1	188	1	2113	5	81.0	2.26	GROWING CROP PRESENT
1	199	7	2113	5	19.0	2.65	GROWING CROP PRESENT
1	113	14	2113	5	23.0	1.46	GROWING CROP PRESENT
1	122	9	2113	3	26.0	1.51	GROWING CROP PRESENT
1	337	3	2120	5	10.0	2.70	ABANDONED
1	68	6	2120	4	16.0	1.95	ABANDONED
1	1	6	2120	3	17.0	1.79	ABANDONED
1	187	1	2120	5	136.0	1.22	ABANDONED
1	127	4	2142	4	15.0	1.50	IMPROVED
1	201	6	2142	5	145.0	1.19	IMPROVED
1	299	8	4300	4	15.0	1.81	MIXED
1	316	6	7500	5	34.0	1.40	DISTURBED LAND
1	316	4	7500	5	85.0	0.94	DISTURBED LAND

Table 25-2. Cluster Group #2

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
2	44	1	5110	5	147.0	1.59	NATURAL (RIVERS & CREEKS)
2	95	4	5110	5	69.0	2.09	NATURAL (RIVERS & CREEKS)
2	108	7	5120	5	55.0	1.23	MAN-MADE (CANALS, DITCHES, & AQUADUCTS)
2	98	6	5210	5	29.0	1.55	NATURAL LAKES & PONDS
2	45	1	5410	5	395.0	0.69	BAYS
2	98	1	5410	5	322.0	1.38	BAYS
2	57	8	5410	5	189.0	1.18	BAYS

Table 25-3. Cluster Group #3

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
3	308	12	2142	4	21.0	1.16	IMPROVED
3	121	10	2143	4	37.0	1.34	UNIMPROVED
3	316	1	7500	5	128.0	0.80	DISTURBED LAND

Table 25-4. Cluster Group #4

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
4	307	4	1112	5	28.0	1.47	RURAL, LOW DENSITY, WITHOUT TREES
4	191	8	1220	5	57.0	0.96	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
4	64	10	1911	5	90.0	0.81	GOLF COURSES
4	122	10	1911	4	58.0	1.16	GOLF COURSES
4	57	7	1911	5	104.0	0.71	GOLF COURSES
4	67	2	1911	4	18.0	1.41	GOLF COURSES
4	7	10	1913	5	31.0	0.91	PARK
4	179	5	1913	5	35.0	0.56	PARK
4	191	1	2113	5	25.0	1.30	GROWING CROP PRESENT
4	121	9	2113	5	10.0	2.12	GROWING CROP PRESENT
4	308	11	2113	4	45.0	0.81	GROWING CROP PRESENT
4	192	15	2113	5	39.0	1.29	GROWING CROP PRESENT
4	199	4	2113	5	32.0	2.12	GROWING CROP PRESENT
4	67	5	2113	4	20.0	1.29	GROWING CROP PRESENT
4	355	1	2113	4	66.0	0.84	GROWING CROP PRESENT
4	84	5	2113	5	28.0	1.41	GROWING CROP PRESENT
4	64	3	2113	5	59.0	1.39	GROWING CROP PRESENT
4	79	1	2113	5	58.0	1.27	GROWING CROP PRESENT
4	265	4	2113	5	12.0	2.30	GROWING CROP PRESENT
4	84	4	2113	5	55.0	0.70	GROWING CROP PRESENT
4	402	1	2120	5	57.0	1.12	ABANDONED
4	122	6	2120	5	20.0	1.16	ABANDONED
4	201	4	2142	5	17.0	1.09	IMPROVED
4	201	1	2142	5	16.0	1.68	IMPROVED
4	307	3	2142	5	23.0	0.92	IMPROVED
4	202	5	2142	5	40.0	0.65	IMPROVED
4	190	9	2142	5	218.0	0.85	IMPROVED
4	308	2	2142	5	95.0	1.08	IMPROVED
4	307	1	2142	4	13.0	2.39	IMPROVED
4	191	10	2142	5	76.0	1.08	IMPROVED

Table 25-5. Cluster Group #5,

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
5	126	16	1310	4	50.0	1.28	MECHANICAL PROCESSING
5	126	10	1911	5	81.0	1.49	GOLF COURSES
5	199	11	2113	5	21.0	1.30	GROWING CROP PRESENT
5	126	6	2113	4	30.0	2.43	GROWING CROP PRESENT
5	126	5	2113	4	23.0	1.21	GROWING CROP PRESENT
5	126	9	2113	5	37.0	1.59	GROWING CROP PRESENT
5	126	8	2113	5	30.0	2.87	GROWING CROP PRESENT
5	126	7	2113	4	22.0	1.32	GROWING CROP PRESENT
5	126	2	2113	4	13.0	2.73	GROWING CROP PRESENT

Table 25-6. Cluster Group #6,

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
6	113	12	1124	4	12.0	1.39	HIGH DENSITY, WITHOUT TREES
6	84	13	1210	5	43.0	1.35	WHOLESALE TRADE AREAS
6	95	1	1210	5	85.0	0.89	WHOLESALE TRADE AREAS
6	99	1	1210	4	24.0	1.74	WHOLESALE TRADE AREAS
6	54	3	1220	4	18.0	1.06	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
6	115	2	1220	5	56.0	1.38	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
6	122	1	1220	5	42.0	1.07	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
6	57	6	1246	4	22.0	1.39	RECREATIONAL (TENNIS,ICE SKATING,STABLES,PLAY AREAS...)
6	97	2	1300	5	66.0	1.08	INDUSTRIAL
6	109	6	1330	5	29.0	1.17	CHEMICAL PROCESSING
6	109	5	1330	5	29.0	1.11	CHEMICAL PROCESSING
6	107	1	1410	5	127.0	0.97	STONE QUARRIES
6	97	6	1420	5	40.0	1.54	SAND & GRAVEL PITS
6	8	2	1520	4	91.0	0.80	RAILROADS & ASSOCIATED FACILITIES
6	99	3	1531	5	68.0	1.35	COMMERCIAL
6	95	3	1640	5	44.0	0.68	MILITARY AREAS
6	64	11	1641	5	22.0	1.47	HOUSING
6	57	5	1644	4	14.0	1.04	TRAINING AREAS
6	64	8	1660	5	12.0	1.44	GOVERNMENT & ADMINISTRATIVE OFFICES
6	300	3	1660	5	18.0	1.94	GOVERNMENT & ADMINISTRATIVE OFFICES
6	119	2	1914	5	31.0	0.68	PARKING LOTS

Table 25-7. Cluster Group #7,

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
7	1	9	1110	5	81.0	0.67	SINGLE-FAMILY HOUSEHOLD UNITS
7	124	2	1113	4	200.0	0.70	URBAN, HIGH DENSITY, WITH TREES
7	124	4	1113	5	303.0	0.54	URBAN, HIGH DENSITY, WITH TREES
7	57	10	1124	4	12.0	1.62	HIGH DENSITY, WITHOUT TREES
7	57	4	1246	3	11.0	1.24	RECREATIONAL (TENNIS,ICE SKATING,STABLES,PLAY AREAS...)
7	179	1	1911	5	83.0	0.86	GOLF COURSES
7	179	12	1912	5	81.0	0.70	CEMETERIES
7	1	8	2113	3	11.0	1.31	GROWING CROP PRESENT
7	355	3	2113	4	22.0	1.32	GROWING CROP PRESENT
7	9	5	2120	4	32.0	1.50	ABANDONED
7	308	3	2142	4	19.0	0.93	IMPROVED
7	98	2	5210	4	19.0	1.75	NATURAL LAKES & PONDS
7	7	7	5300	5	54.0	1.67	RESERVOIRS

Table 25-8. Cluster Group #8

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
8	7	5	1110	5	54.0	1.83	SINGLE-FAMILY HOUSEHOLD UNITS
8	68	16	1110	5	27.0	1.33	SINGLE-FAMILY HOUSEHOLD UNITS
8	8	4	1120	5	64.0	0.94	MULTI-FAMILY HOUSEHOLD UNITS
8	8	3	1120	5	38.0	1.11	MULTI-FAMILY HOUSEHOLD UNITS
8	7	1	1120	5	33.0	1.44	MULTI-FAMILY HOUSEHOLD UNITS
8	85	5	1124	4	15.0	1.69	HIGH DENSITY, WITHOUT TREES
8	57	9	1124	5	20.0	1.21	HIGH DENSITY, WITHOUT TREES
8	97	4	1124	4	16.0	1.44	HIGH DENSITY, WITHOUT TREES
8	179	11	1124	5	31.0	2.10	HIGH DENSITY, WITHOUT TREES
8	95	5	1210	5	37.0	1.06	WHOLESALE TRADE AREAS
8	8	1	1210	5	48.0	0.73	WHOLESALE TRADE AREAS
8	6	5	1220	4	21.0	0.88	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
8	179	10	1220	5	49.0	0.84	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
8	118	6	1220	3	30.0	1.18	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
8	119	9	1220	4	34.0	1.42	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
8	8	5	1220	5	58.0	0.87	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
8	54	2	1300	5	22.0	1.28	INDUSTRIAL
8	99	2	1300	5	48.0	0.99	INDUSTRIAL
8	157	2	1310	5	13.0	1.91	MECHANICAL PROCESSING
8	118	9	1340	5	18.0	1.35	FABRICATION & ASSEMBLY
8	65	6	1560	5	70.0	0.84	ELECTRIC, GAS, WATER, SEWAGE DISPOSAL, SOLID WASTE, UTIL
8	7	6	1611	4	10.0	1.40	PRIMARY
8	109	2	1621	4	17.0	1.29	HOSPITALS
8	118	7	1643	4	10.0	1.23	STORAGE AREAS
8	109	11	1643	4	54.0	0.62	STORAGE AREAS
8	113	6	1653	5	37.0	0.90	STATE
8	119	4	1660	5	12.0	1.62	GOVERNMENT & ADMINISTRATIVE OFFICES
8	108	3	1913	5	75.0	1.15	PARK
8	64	9	1914	5	48.0	1.32	PARKING LOTS
8	119	5	1914	5	42.0	1.33	PARKING LOTS
8	157	3	2113	4	31.0	1.51	GROWING CROP PRESENT

Table 25-9. Cluster Group #9

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
9	126	3	2113	3	13.0	0.00	GROWING CROP PRESENT
9	126	4	2113	3	13.0	2.68	GROWING CROP PRESENT

Table 25-10. Cluster Group #10

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
10	124	3	1110	4	88.0	1.06	SINGLE-FAMILY HOUSEHOLD UNITS
10	299	4	1110	5	16.0	2.33	SINGLE-FAMILY HOUSEHOLD UNITS
10	68	10	1110	5	39.0	1.12	SINGLE-FAMILY HOUSEHOLD UNITS
10	118	5	1110	4	73.0	0.75	SINGLE-FAMILY HOUSEHOLD UNITS
10	98	4	1112	5	65.0	1.13	RURAL, LOW DENSITY, WITHOUT TREES
10	299	2	1113	5	30.0	2.34	URBAN, HIGH DENSITY, WITH TREES
10	65	1	1113	5	38.0	1.31	URBAN, HIGH DENSITY, WITH TREES
10	118	4	1113	4	63.0	1.14	URBAN, HIGH DENSITY, WITH TREES
10	299	3	1113	5	17.0	1.28	URBAN, HIGH DENSITY, WITH TREES
10	84	9	1113	5	23.0	1.22	URBAN, HIGH DENSITY, WITH TREES
10	45	2	1113	5	178.0	1.32	URBAN, HIGH DENSITY, WITH TREES
10	85	7	1113	5	194.0	1.04	URBAN, HIGH DENSITY, WITH TREES
10	97	5	1113	5	100.0	1.14	URBAN, HIGH DENSITY, WITH TREES
10	119	8	1113	5	137.0	0.79	URBAN, HIGH DENSITY, WITH TREES
10	97	7	1113	5	156.0	1.09	URBAN, HIGH DENSITY, WITH TREES
10	84	10	1121	5	114.0	0.81	LOW DENSITY, WITH TREES
10	64	1	1122	5	39.0	1.31	LOW DENSITY, WITHOUT TREES
10	119	11	1123	5	85.0	0.77	HIGH DENSITY, WITH TREES
10	6	2	1124	5	154.0	1.11	HIGH DENSITY, WITHOUT TREES
10	186	2	1124	5	26.0	1.09	HIGH DENSITY, WITHOUT TREES
10	57	1	1124	5	10.0	1.50	HIGH DENSITY, WITHOUT TREES
10	85	2	1124	5	35.0	0.87	HIGH DENSITY, WITHOUT TREES
10	299	7	1124	3	13.0	1.40	HIGH DENSITY, WITHOUT TREES
10	179	9	1124	4	29.0	1.38	HIGH DENSITY, WITHOUT TREES
10	85	6	1124	4	21.0	1.53	HIGH DENSITY, WITHOUT TREES
10	85	3	1124	5	19.0	1.46	HIGH DENSITY, WITHOUT TREES
10	41	2	1210	4	11.0	1.52	WHOLESALE TRADE AREAS
10	68	1	1220	3	12.0	1.84	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
10	2	2	1300	4	20.0	1.78	INDUSTRIAL
10	6	7	1610	4	12.0	1.62	EDUCATIONAL FACILITIES
10	6	1	1612	4	27.0	0.67	SECONDARY
10	7	2	1615	5	50.0	1.25	UNIVERSITY
10	67	1	1621	5	20.0	0.68	HOSPITALS
10	299	1	1912	5	64.0	1.51	CEMETERIES
10	7	11	1912	5	55.0	0.85	CEMETERIES
10	118	12	1912	5	101.0	1.14	CEMETERIES
10	119	6	1912	5	53.0	1.01	CEMETERIES
10	84	6	2112	4	33.0	1.74	BARE--RECENTLY PLOWED
10	68	5	2113	4	36.0	1.59	GROWING CROP PRESENT
10	157	1	2113	4	26.0	1.49	GROWING CROP PRESENT
10	45	3	2120	5	25.0	1.17	ABANDONED
10	9	4	2120	3	19.0	1.11	ABANDONED
10	2	3	2250	4	15.0	1.77	TURF FARM
10	117	5	4200	5	12.0	1.22	EVERGREEN (CONIFEROUS & OTHER)
10	117	7	4300	4	10.0	1.92	MIXED

Table 25-11. Cluster Group #11

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
11	14	1	1110	5	143.0	0.61	SINGLE-FAMILY HOUSEHOLD UNITS
11	353	2	1110	4	12.0	2.35	SINGLE-FAMILY HOUSEHOLD UNITS
11	1	11	1110	3	14.0	1.51	SINGLE-FAMILY HOUSEHOLD UNITS
11	307	5	1112	4	30.0	1.12	RURAL, LOW DENSITY, WITHOUT TREES
11	190	11	1112	4	13.0	1.38	RURAL, LOW DENSITY, WITHOUT TREES
11	190	2	1112	4	14.0	2.07	RURAL, LOW DENSITY, WITHOUT TREES
11	36	2	1112	4	14.0	0.98	RURAL, LOW DENSITY, WITHOUT TREES
11	41	1	1112	4	15.0	1.82	RURAL, LOW DENSITY, WITHOUT TREES
11	191	4	1113	5	130.0	0.99	URBAN, HIGH DENSITY, WITH TREES
11	200	3	1113	5	54.0	1.36	URBAN, HIGH DENSITY, WITH TREES
11	191	3	1113	5	64.0	1.22	URBAN, HIGH DENSITY, WITH TREES
11	104	6	1113	4	13.0	1.08	URBAN, HIGH DENSITY, WITH TREES
11	402	9	1113	5	21.0	1.39	URBAN, HIGH DENSITY, WITH TREES
11	113	11	1113	4	88.0	1.48	URBAN, HIGH DENSITY, WITH TREES
11	355	4	1113	4	199.0	0.77	URBAN, HIGH DENSITY, WITH TREES
11	202	1	1114	5	136.0	1.00	URBAN, HIGH DENSITY, WITHOUT TREES
11	402	5	1114	5	172.0	1.07	URBAN, HIGH DENSITY, WITHOUT TREES
11	200	2	1114	5	185.0	0.80	URBAN, HIGH DENSITY, WITHOUT TREES
11	191	5	1114	5	24.0	1.68	URBAN, HIGH DENSITY, WITHOUT TREES
11	353	6	1114	4	41.0	1.44	URBAN, HIGH DENSITY, WITHOUT TREES
11	192	10	1246	5	13.0	1.79	RECREATIONAL (TENNIS,ICE SKATING,STABLES,PLAY AREAS...)
11	122	5	1611	3	12.0	2.05	PRIMARY
11	118	10	1911	5	55.0	1.01	GOLF COURSES
11	192	14	1911	5	126.0	0.80	GOLF COURSES
11	6	6	1911	5	77.0	1.08	GOLF COURSES
11	344	2	1911	5	85.0	0.90	GOLF COURSES
11	86	6	2112	5	67.0	1.00	BARE--RECENTLY PLOWED
11	13	3	2113	5	39.0	1.49	GROWING CROP PRESENT
11	352	6	2113	5	31.0	0.91	GROWING CROP PRESENT
11	402	8	2113	5	48.0	1.15	GROWING CROP PRESENT
11	104	3	2113	5	30.0	1.51	GROWING CROP PRESENT
11	104	4	2113	5	43.0	1.32	GROWING CROP PRESENT
11	352	5	2113	5	28.0	1.18	GROWING CROP PRESENT
11	12	2	2113	5	44.0	1.38	GROWING CROP PRESENT
11	366	9	2113	5	65.0	1.73	GROWING CROP PRESENT
11	340	3	2113	5	27.0	1.30	GROWING CROP PRESENT
11	308	1	2113	5	35.0	1.09	GROWING CROP PRESENT
11	340	11	2113	5	25.0	0.97	GROWING CROP PRESENT
11	366	4	2113	5	74.0	1.35	GROWING CROP PRESENT
11	366	5	2113	5	25.0	1.19	GROWING CROP PRESENT
11	45	4	2113	5	25.0	1.37	GROWING CROP PRESENT
11	366	11	2113	5	24.0	1.60	GROWING CROP PRESENT
11	192	13	2113	5	24.0	2.34	GROWING CROP PRESENT
11	340	8	2113	5	24.0	1.58	GROWING CROP PRESENT
11	337	8	2113	5	24.0	1.41	GROWING CROP PRESENT
11	337	5	2113	5	24.0	1.08	GROWING CROP PRESENT

Table 25-11 (Continued)

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
11	344	3	2113	5	42.0	0.59	GROWING CROP PRESENT
11	1	5	2113	5	46.0	1.01	GROWING CROP PRESENT
11	341	3	2113	5	22.0	1.62	GROWING CROP PRESENT
11	352	7	2113	5	21.0	1.21	GROWING CROP PRESENT
11	192	6	2113	5	21.0	1.27	GROWING CROP PRESENT
11	12	1	2113	5	63.0	0.75	GROWING CROP PRESENT
11	192	3	2113	5	74.0	1.89	GROWING CROP PRESENT
11	265	1	2113	5	37.0	1.55	GROWING CROP PRESENT
11	340	2	2113	5	19.0	1.04	GROWING CROP PRESENT
11	65	4	2113	5	94.0	1.18	GROWING CROP PRESENT
11	366	3	2113	5	48.0	0.78	GROWING CROP PRESENT
11	192	12	2113	5	17.0	1.83	GROWING CROP PRESENT
11	352	9	2113	5	73.0	1.15	GROWING CROP PRESENT
11	352	11	2113	5	17.0	1.66	GROWING CROP PRESENT
11	192	1	2113	5	69.0	1.83	GROWING CROP PRESENT
11	366	12	2113	5	17.0	2.08	GROWING CROP PRESENT
11	192	11	2113	5	40.0	1.61	GROWING CROP PRESENT
11	199	5	2113	5	16.0	2.59	GROWING CROP PRESENT
11	202	2	2113	5	69.0	1.48	GROWING CROP PRESENT
11	188	5	2113	4	25.0	1.35	GROWING CROP PRESENT
11	190	6	2113	4	30.0	1.51	GROWING CROP PRESENT
11	308	8	2113	4	94.0	1.18	GROWING CROP PRESENT
11	2	1	2113	4	26.0	1.35	GROWING CROP PRESENT
11	13	2	2113	3	22.0	0.81	GROWING CROP PRESENT
11	25	3	2113	4	14.0	1.87	GROWING CROP PRESENT
11	68	4	2113	3	13.0	1.91	GROWING CROP PRESENT
11	265	2	2113	5	10.0	1.92	GROWING CROP PRESENT
11	168	3	2113	3	18.0	2.45	GROWING CROP PRESENT
11	337	10	2113	5	12.0	2.28	GROWING CROP PRESENT
11	307	11	2113	5	14.0	2.09	GROWING CROP PRESENT
11	54	1	2113	4	28.0	0.89	GROWING CROP PRESENT
11	13	9	2113	4	49.0	1.22	GROWING CROP PRESENT
11	402	7	2113	5	12.0	2.48	GROWING CROP PRESENT
11	121	1	2113	3	14.0	1.41	GROWING CROP PRESENT
11	1	10	2113	3	14.0	0.97	GROWING CROP PRESENT
11	67	7	2113	3	15.0	1.27	GROWING CROP PRESENT
11	341	2	2113	4	11.0	1.55	GROWING CROP PRESENT
11	25	4	2113	4	27.0	1.30	GROWING CROP PRESENT
11	375	9	2113	5	10.0	2.11	GROWING CROP PRESENT
11	375	1	2113	5	10.0	2.11	GROWING CROP PRESENT
11	402	2	2113	5	12.0	2.39	GROWING CROP PRESENT
11	36	1	2113	3	43.0	0.89	GROWING CROP PRESENT
11	366	15	2113	4	20.0	1.39	GROWING CROP PRESENT
11	341	4	2113	5	13.0	1.56	GROWING CROP PRESENT
11	35	1	2113	3	18.0	2.05	GROWING CROP PRESENT
11	113	3	2113	4	16.0	1.62	GROWING CROP PRESENT

Table 25-11 (Continued)

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
11	337	9	2113	5	10.0	2.25	GROWING CROP PRESENT
11	35	2	2113	3	10.0	1.39	GROWING CROP PRESENT
11	9	1	2113	4	19.0	1.40	GROWING CROP PRESENT
11	366	14	2113	4	37.0	1.38	GROWING CROP PRESENT
11	199	2	2113	5	13.0	2.61	GROWING CROP PRESENT
11	364	3	2114	3	12.0	1.58	HARVESTED
11	366	1	2120	5	32.0	2.12	ABANDONED
11	402	4	2120	5	39.0	0.59	ABANDONED
11	13	5	2120	5	27.0	0.95	ABANDONED
11	13	1	2120	5	96.0	0.79	ABANDONED
11	187	2	2120	5	88.0	1.43	ABANDONED
11	127	5	2120	4	18.0	1.86	ABANDONED
11	344	4	2120	4	25.0	0.87	ABANDONED
11	340	1	2120	5	26.0	1.31	ABANDONED
11	375	5	2120	5	15.0	1.57	ABANDONED
11	104	2	2120	5	17.0	1.65	ABANDONED
11	192	4	2120	5	23.0	1.20	ABANDONED
11	307	2	2142	5	20.0	2.17	IMPROVED
11	366	2	2142	5	20.0	1.12	IMPROVED
11	375	2	2142	5	21.0	0.98	IMPROVED
11	341	5	2142	5	15.0	0.80	IMPROVED
11	201	7	2142	5	22.0	1.65	IMPROVED
11	188	2	2142	5	22.0	1.82	IMPROVED
11	366	8	2142	5	12.0	1.81	IMPROVED
11	201	3	2142	5	11.0	2.51	IMPROVED
11	366	7	2142	5	10.0	2.21	IMPROVED
11	352	10	2142	5	23.0	1.42	IMPROVED
11	375	8	2142	5	24.0	1.63	IMPROVED
11	188	4	2142	5	25.0	1.53	IMPROVED
11	308	6	2142	4	36.0	0.71	IMPROVED
11	308	4	2142	4	33.0	1.32	IMPROVED
11	308	5	2142	4	30.0	1.03	IMPROVED
11	190	1	2142	4	30.0	1.05	IMPROVED
11	107	3	2142	4	27.0	1.51	IMPROVED
11	192	5	2142	5	38.0	0.77	IMPROVED
11	366	13	2142	4	20.0	1.84	IMPROVED
11	201	5	2142	5	100.0	0.77	IMPROVED
11	104	5	2142	4	17.0	1.92	IMPROVED
11	364	2	2142	4	15.0	1.72	IMPROVED
11	201	2	2142	5	38.0	0.92	IMPROVED
11	35	3	2142	4	14.0	2.30	IMPROVED
11	104	7	2142	5	77.0	1.08	IMPROVED
11	127	10	2142	4	12.0	1.66	IMPROVED
11	190	10	2142	5	80.0	2.19	IMPROVED
11	67	6	2142	3	14.0	2.08	IMPROVED
11	107	4	2142	5	87.0	0.92	IMPROVED

Table 25-11 (Continued)

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
11	36	5	2142	5	39.0	1.56	IMPROVED
11	202	3	2142	5	42.0	1.08	IMPROVED
11	13	4	2142	5	43.0	1.44	IMPROVED
11	340	7	2142	5	46.0	1.28	IMPROVED
11	192	7	2142	5	48.0	1.51	IMPROVED
11	365	2	2142	5	55.0	0.70	IMPROVED
11	402	3	2142	5	61.0	0.56	IMPROVED
11	352	3	2142	5	64.0	0.98	IMPROVED
11	191	7	2143	5	12.0	1.83	UNIMPROVED
11	168	6	2143	5	23.0	2.13	UNIMPROVED
11	340	10	2143	5	31.0	1.08	UNIMPROVED
11	188	3	2143	5	41.0	0.59	UNIMPROVED
11	337	4	2240	5	93.0	0.89	NURSERIES AND FLORICULTURAL AREAS
11	13	8	2250	3	19.0	1.77	TURF FARM
11	266	3	4100	5	20.0	1.59	DECIDUOUS
11	129	4	6100	5	15.0	1.66	VEGETATED WETLANDS

Table 25-12. Cluster Group #12

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
12	199	12	2113	5	16.0	2.53	GROWING CROP PRESENT
12	199	1	2113	5	43.0	1.41	GROWING CROP PRESENT
12	100	2	6000	5	30.0	1.72	WETLANDS
12	100	1	6000	5	74.0	0.90	WETLANDS
12	111	1	6100	5	48.0	1.26	VEGETATED WETLANDS
12	111	2	6100	5	48.0	1.65	VEGETATED WETLANDS
12	111	4	6100	5	55.0	1.89	VEGETATED WETLANDS

Table 25-13. Cluster Group #13

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
13	123	3	6000	5	20.0	2.22	WETLANDS
13	102	2	6110	4	23.0	0.00	BRACKISH MARSH

Table 25-14. Cluster Group #14

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
14	7	9	1120	5	158.0	1.26	MULTI-FAMILY HOUSEHOLD UNITS
14	108	4	1200	5	40.0	1.71	COMMERCIAL & SERVICES
14	7	3	1243	5	31.0	1.14	SPORTS (STADIUMS,ARENAS,RACETRACKS,OTHER)
14	108	2	1300	5	84.0	0.70	INDUSTRIAL
14	109	10	1300	5	184.0	0.56	INDUSTRIAL
14	109	8	1300	5	25.0	0.97	INDUSTRIAL
14	109	1	1520	4	37.0	0.86	RAILROADS & ASSOCIATED FACILITIES
14	94	1	1520	5	37.0	0.98	RAILROADS & ASSOCIATED FACILITIES
14	108	9	1520	5	38.0	0.66	RAILROADS & ASSOCIATED FACILITIES
14	109	7	1520	5	49.0	0.68	RAILROADS & ASSOCIATED FACILITIES
14	108	6	1544	5	99.0	0.93	PART FACILITIES
14	108	6	1544	4	14.0	2.19	PART FACILITIES
14	108	5	1544	4	32.0	1.61	PART FACILITIES
14	109	9	1914	5	31.0	1.23	PARKING LOTS

Table 25-15. Cluster Group #15

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
15	124	1	1114	4	59.0	1.02	URBAN, HIGH DENSITY, WITHOUT TREES
15	35	4	1330	4	18.0	0.77	CHEMICAL PROCESSING
15	41	4	1420	4	18.0	0.88	SAND & GRAVEL PITS
15	64	7	1641	4	17.0	1.40	HOUSING
15	188	7	2112	4	11.0	1.84	BARE--RECENTLY PLOWED
15	192	2	2113	5	69.0	1.46	GROWING CROP PRESENT
15	307	9	2240	5	15.0	1.04	NURSERIES AND FLORICULTURAL AREAS
15	307	7	2240	5	29.0	2.22	NURSERIES AND FLORICULTURAL AREAS
15	359	6	7500	5	62.0	1.53	DISTURBED LAND

Table 25-16. Cluster Group #16

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
16	191	9	2113	5	20.0	0.00	GROWING CROP PRESENT

Table 25-17. Cluster Group #17

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
17	190	8	1111	5	33.0	0.89	RURAL, LOW DENSITY, WITH TREES
17	316	2	1113	5	12.0	1.00	URBAN, HIGH DENSITY, WITH TREES
17	119	7	1113	5	37.0	0.99	URBAN, HIGH DENSITY, WITH TREES
17	402	6	1113	5	28.0	0.99	URBAN, HIGH DENSITY, WITH TREES
17	316	3	1113	5	92.0	0.97	URBAN, HIGH DENSITY, WITH TREES
17	202	4	1220	5	16.0	1.36	RETAIL TRADE AREAS (BUSINESS DIST.,SHOPPING CENT.,COMM.)
17	63	1	1615	5	171.0	0.80	UNIVERSITY
17	84	2	2112	5	25.0	1.91	BARE--RECENTLY PLOWED
17	340	5	2113	5	17.0	1.35	GROWING CROP PRESENT
17	352	4	2113	5	10.0	1.63	GROWING CROP PRESENT
17	110	6	2113	5	36.0	1.26	GROWING CROP PRESENT
17	192	8	2113	5	13.0	2.58	GROWING CROP PRESENT
17	352	2	2113	5	25.0	1.57	GROWING CROP PRESENT
17	110	5	2113	5	28.0	1.30	GROWING CROP PRESENT
17	113	1	2113	5	11.0	1.93	GROWING CROP PRESENT
17	113	2	2113	3	10.0	1.94	GROWING CROP PRESENT
17	308	7	2113	4	21.0	2.36	GROWING CROP PRESENT
17	117	2	2113	4	22.0	2.45	GROWING CROP PRESENT
17	9	6	2120	4	26.0	1.74	ABANDONED
17	201	9	2142	5	12.0	1.33	IMPROVED
17	202	7	2142	5	46.0	0.59	IMPROVED
17	202	8	2142	5	31.0	1.10	IMPROVED
17	121	3	2142	5	13.0	1.39	IMPROVED
17	386	1	4100	5	30.0	1.76	DECIDUOUS
17	386	2	4100	5	30.0	2.20	DECIDUOUS
17	308	9	4100	5	29.0	1.85	DECIDUOUS
17	352	8	4100	5	27.0	1.64	DECIDUOUS
17	179	8	4100	5	27.0	1.25	DECIDUOUS
17	340	6	4100	5	31.0	1.03	DECIDUOUS
17	122	13	4100	5	24.0	1.07	DECIDUOUS
17	340	12	4100	5	24.0	1.14	DECIDUOUS
17	337	7	4100	5	21.0	1.72	DECIDUOUS
17	337	6	4100	5	20.0	1.42	DECIDUOUS
17	364	1	4100	5	20.0	1.50	DECIDUOUS
17	179	7	4100	5	20.0	1.53	DECIDUOUS
17	190	7	4100	5	20.0	2.15	DECIDUOUS
17	299	9	4100	5	20.0	2.13	DECIDUOUS
17	365	1	4100	5	180.0	0.94	DECIDUOUS
17	121	8	4100	5	20.0	1.74	DECIDUOUS
17	366	6	4100	5	36.0	1.48	DECIDUOUS
17	122	15	4100	5	12.0	2.24	DECIDUOUS
17	308	10	4100	5	72.0	1.72	DECIDUOUS
17	67	8	4100	5	84.0	1.35	DECIDUOUS
17	9	3	4100	5	91.0	0.88	DECIDUOUS
17	98	5	4100	5	36.0	1.62	DECIDUOUS
17	107	2	4100	5	42.0	1.39	DECIDUOUS

Table 25-17 (Continued)

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
17	117	8	4100	5	42.0	1.41	DECIDUOUS
17	67	3	4100	5	49.0	1.54	DECIDUOUS
17	84	7	4100	5	44.0	1.79	DECIDUOUS
17	129	3	4300	5	18.0	2.10	MIXED
17	359	5	4300	5	35.0	1.20	MIXED
17	98	3	4300	5	85.0	1.56	MIXED
17	202	6	4400	5	25.0	1.51	UPLAND SHRUBS
17	200	4	4400	5	25.0	1.56	UPLAND SHRUBS
17	200	5	4400	5	26.0	1.31	UPLAND SHRUBS
17	135	1	6200	5	318.0	2.47	FORESTED WETLANDS
17	166	1	6200	5	61.0	2.19	FORESTED WETLANDS

Table 25-18 Cluster Group #18

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
18	126	13	2113	4	21.0	1.34	GROWING CROP PRESENT
18	126	15	2113	3	11.0	2.95	GROWING CROP PRESENT
18	126	11	2113	3	16.0	2.58	GROWING CROP PRESENT
18	126	12	2113	3	32.0	1.38	GROWING CROP PRESENT
18	168	2	4200	5	33.0	0.81	EVERGREEN (CONIFEROUS & OTHER)
18	168	6	4300	5	110.0	0.93	MIXED
18	123	1	4300	5	41.0	1.81	MIXED
18	123	2	4300	5	42.0	1.19	MIXED

Table 25-19. Cluster Group #19

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
19	199	10	2113	5	23.0	2.06	GROWING CROP PRESENT
19	199	9	2120	5	11.0	0.00	ABANDONED

Table 25-20. Cluster Group #20

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
20	121	7	4100	5	25.0	2.28	DECIDUOUS
20	124	5	4100	5	16.0	0.00	DECIDUOUS
20	102	1	6110	5	179.0	2.64	BRACKISH MARSH

Table 25-21. Cluster Group #21

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
21	75	1	4200	5	35.0	0.71	EVERGREEN (CONIFEROUS & OTHER)
21	75	2	4200	4	11.0	1.11	EVERGREEN (CONIFEROUS & OTHER)
21	75	4	4200	5	30.0	1.02	EVERGREEN (CONIFEROUS & OTHER)
21	123	4	6000	5	25.0	2.34	WETLANDS

Table 25-22. Cluster Group #22

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
22	111	3	5110	5	54.0	1.25	NATURAL (RIVERS & CREEKS)
22	168	1	5110	5	64.0	0.73	NATURAL (RIVERS & CREEKS)
22	129	1	5110	5	16.0	1.08	NATURAL (RIVERS & CREEKS)
22	129	2	5110	5	16.0	1.51	NATURAL (RIVERS & CREEKS)
22	117	12	5210	5	30.0	1.35	NATURAL LAKES & PONDS
22	135	2	5210	5	83.0	1.71	NATURAL LAKES & PONDS

Table 25-23. Cluster Group #23

GROUP NUMBER	TESTSITE NUMBER	FIELD NUMBER	LAND USE CODE	FIELD QUALITY	FIELD SIZE	DISTANCE	DESCRIPTION OF LAND USE CODE
23	168	4	2112	5	78.0	2.01	BARE--RECENTLY PLOWED
23	168	5	2112	4	15.0	2.11	BARE--RECENTLY PLOWED
23	199	3	2113	5	18.0	1.48	GROWING CROP PRESENT
23	45	5	2113	3	33.0	0.72	GROWING CROP PRESENT
23	117	10	6100	5	16.0	1.48	VEGETATED WETLANDS
23	117	11	6100	5	16.0	1.05	VEGETATED WETLANDS

25 to 50 cases took place constantly through 17 cycles after establishment of the 13 groups. Because of the uncertainty about threshold 4, only a portion of the table is reproduced here and discussion is confined to threshold 3.

Inspection of Table 21 of group means and standard deviations for threshold 3 and Tables 22, 24, and 25 enables the following observations to be made:

Group 1

The largest single category in this group is urban residential, but it is still a very mixed group (50 urban residential of a total of 93 cases), indicating real difficulties in discriminating between urban residential without trees, but with grass, from other groups within which grass - or growing crops - are present. The 43 confusion categories (non-urban residential) are as follows:

1200	Commercial and Services	(4)
1300	Industrial	(1)
1400	Extractive	(1)
1500	Transportation	(5)
1600	Institutional	(8)
1900	Open and Other Urban	(6)
2100	Cropland and Pasture	(15)
4300	Mixed Forest	(1)
7500	Barren land	(2)

The other urban classes to greater or lesser extent will include grass and trees, particularly Institutional (large buildings usually surrounded by open space, such as hospitals, schools and similar facilities) and Open, and Other Urban. Multi-date imagery should enable improved separation of these sub-elements within the cluster. The balance of within pixel mixture of paved surfaces, buildings and roof tops, and green vegetation would shift throughout the year and should lead to better separation.

Group 2

This small group (7 cases) consists exclusively of water, both salt and fresh.

Group 3

This very small group (3 cases) is an outlier of cropland (2) and barren land (1).

Group 4

This group is dominated by crop land (22 cases) and open (grassy) urban land (6 cases). Two odd cases are single and multi-family residential (1) cases each with substantial grass lawns.

Group 5

This group of 9 is dominantly cropland (7 cases) with one case each of industrial and open and other urban, the latter also with grassy areas.

Group 6

This group of 20 cases is entirely urban, all but 2 cases (one residential and one open and other urban) being of a type likely to be low to very low in green vegetation. The remaining 18 cases cover a wide spectrum of land uses (retail trade areas, extractive, chemical processing, wholesale, government and administrative offices, etc.) but share the characteristic of low vegetation density. In this respect, it is interesting to note the very high thermal IR group mean, band 13. It is the second highest shown in Table 21, after group 14, itself an intensively urban low-vegetation group.

Group 7

Group 7 is very mixed comprising the following cases:

1100	Single Family Residential	(4)
1200	Multi-Family Residential	(1)
1900	Open and Other Urban	(2)
2100	Cropland and Pasture	(4)
5200	Lakes	(1)
5300	Reservoirs	(1)

It is very difficult to attribute meaning to this cluster. It is a measure of the degree of spectral overlap between land use classes as seen by the sensor, in comparison to the way the land use planner sees these classes, that such a complex group could be erected as a cluster.

Group 8

Group 8 is entirely urban except for one cropland outlier (30 urban, 1 cropland). However, the within-urban Level II categories range widely as shown below:

1100	Single Family Residential	(9)
1200	Multi-Family Residential	(7)
1300	Industrial	(4)
1500	Transportation	(1)
1600	Institutional	(6)
1900	Open and Other Urban	(3)
2100	Cropland and Pasture	(1)

All are, however, groups likely to be low in tree cover as shown by the detailed Level III and IV classes in Table 25-8.

Group 9

This group of 2 consists of a cropland and pasture outlier.

Group 10

Group 10 is composed of 46 cases, all but 8 of urban classes dominantly residential (see Table 25-10), both with and without trees.

Group 11

This group is primarily cropland and pasture having 124 cases of growing crop present or improved pastures out of a total of 154 cases. The only major confusion group is urban residential, both with and without trees (and grasses), numbering 20 cases. Multi-date imagery should enable separation of these outliers, through changes in reflectance in comparison to the cropland (differential bare ground and paved surface reflectance, etc.).

Group 12

This is a wetland group (5 cases) without 2 outliers of growing crop present. Since the SL/3 pass took place shortly after very heavy rainfalls some of the latter fields may have been unusually damp at the time of the overpass.

Group 13

This group of 2 is also wetlands.

Group 14

This group of 14 cases is entirely urban, with the major components being railroads and other transportation (7 cases), industrial (3 cases), and other land use types of very low vegetation cover. The high value group mean for the thermal IR band clearly is important in discriminating this group.

Group 15

This is a very mixed group of 9 cases very difficult to interpret. (See Table 25-15).

Group 16

This is an outlier of one growing crop present.

Group 17

This group of 57 cases is dominated by land use types with deciduous trees as shown below:

1111	Rural, Low Density Single Family Residential, with Trees	(1)
1113	Urban, High Density Single Family Residential, with Trees	(4)
2113	Growing crop present	(10)
2140	Pasture land	(5)
4100	Deciduous Forest	(26)
4300	Mixed Deciduous and Evergreen Forest	(3)
4400	Upland Shrubs	(3)
6200	Forested Wetlands	(2)

While principally forest land, the inclusions are explicable in that the urban areas contain trees, and both the areas of growing crop and pasture land may contain edge - pixels with treed land. Multi-date imagery may well enable these inclusions to be separated from the true forest land.

Group 18

This group is mainly mixed forest (3 cases) and growing crops (4 cases), all site 126, where there are significant forest edges to many fields, and one evergreen forest.

Group 19

This is another group of 2 cropland pasture lands from a single test site, 199, evidently a special outlier since it lies on a small island in Chesapeake Bay.

Group 20

An outlier of 2 deciduous forest and 1 brackish marsh.

Group 21

This group is composed of evergreen forest (3 cases) with a mixed forest (evergreen and deciduous wetland) (1 case).

Group 22

This group is composed entirely of fresh water rivers and lakes.

Group 23

The final group is a mixture of bare ground (2 cases), growing crop present (2), and vegetated wetlands (2). In each case there is rather more bare ground present in the vegetative members of the group than is common in other areas.

The relationships between the various groups and their closest affinity members (as shown by the between-group difference matrix, Table 22; and by the group means and standard deviations for 23 cluster groups, Table 21) are summarized in Table 26. It is seen in this table that the various Level II groups are ordered into a number of master Level I categories. These categories have a relatively high degree of consistency and classificatory rationality. They indicated that improvements in classification may be achievable with multi-date imagery. From this analysis of the clustering groups, it is clear that when all 13 spectral bands are available, considerable within level II discrimination should be feasible with multi-date imagery. Even though no explicit investigation of multi-date material was carried out, the character of the "outliers" included in each cluster should enable separation to be achieved. Despite the very difficult noise problems with the Skylab S-

Table 26. Relationship Between Groups and Their Closest Affinity Groups

Cluster Number	Number of Fields	Dominant Category	Number	Closest Affinities to Groups/Dominant Categories
		WATER		
2	7	Water	7	22, then 20 (Deciduous forest and Brackish Marsh)
22	6	Water	6	2 then 13 (Wetlands)
		CROPLAND (and OPEN URBAN RESIDENTIAL)		
4	30	Cropland Open Urban Land	22	11 (Cropland and Pasture) 23 (Bare Ground & Growing Crop), 10 (Urban Residential 7 (Very Mixed), 17 (Deciduous Trees)
11	154	Cropland and Pasture Urban Residential	124 20	4, 23, 17, 10
23	6	Mixed: Bare Ground Growing Crops Vegetated Wetlands	2 2 2	4, 11
		INTENSELY URBAN (LOW or NO TREE COVER)		
6	21	Urban: With Low Tree Cover	21	8, 14, 15, 10
7	13	Mixed, Principally Urban	6	8, 10, 17, 4
8	31	Urban: Principally Residential	30	10, 6, 14, 7
10	46	Urban: Residential With and Without Trees	38	8, 7, 17, 11
14	14	Urban: Transportation Industrial	10	8, 6, 7, 12
		URBAN (MIXED)		
1	93	Urban: Mixed Residential Other Urban Cropland	 50 25 15	10, 8, 4, 11

Cluster Number	Number of Fields	Dominant Category	Number	Closest Affinities to Groups/Dominant Categories
		DECIDUOUS FOREST LANDS		
17	57	Deciduous Forest and Cropland	29	11, 10, 7, 18
		MIXED FOREST		
18	8	Mixed Forest	4	10, 12, 17
		EVERGREEN FOREST		
21	7	Evergreen Forest	3	17
		WETLANDS		
12	7	Wetlands	5	18, 8, 10, 7
13	2	Wetland (outlier)	2	NONE
		OUTLIERS, PRIMARILY CROPLAND		
3	3	Cropland Barren	2	15
5	9	Cropland	7	6
9	2	Cropland Pasture	2	NONE
15	9	Mixed		3, 6, Difficult to interpret
16	1	Cropland	1	NONE
19	2	Cropland	2	NONE
20	3	Cropland	2	NONE

192 scanner, only partially removed by filtering and other pre-processing, this result is an encouraging indication of what could be achieved with both higher spatial and spectral resolution and low-noise multi-channel data. Since it is already known from many LANDSAT I and II experiments that multi-date information significantly improves identification accuracies, the clustering results seen here are therefore quite encouraging for greater discrimination as better multispectral scanner systems and multi-date information become available.

2.2.4 Implications of the Analysis Procedures on the Results

The results obtained with the clustering are influenced by the following factors:

1) The distance measure employed,
2) The time of year at which the data was obtained,
3) The degree of randomness or otherwise in the data collection,
4) The nature of the local between-class confusions, a product of the local land use mixture and spectral overlaps.

All of these factors in one manner or another make some indeterminate contributions to the results and consequently, the conclusions should be viewed with some caution.

- The distance measure suffers from the process of collapsing all off-diagonal elements out of the analysis to reduce computer time by assuming only diagonal covariance matrices. Because of the considerable spectral separation of many of the categories, these effects should be relatively small. There will of course be differences also arising from the thresholding adopted, as well as the general distance measure itself compared to other measures. While not inconsequential, these differences were not investigated and are part of the noise arising from different analytical techniques between different investigators.

- Early August - when the S-192 coverage was obtained - is a time of year when all trees are still in leaf and much cropland is still covered by active growing vegetation. It is also a time of year (and time of day near 11:00 am) when there will be substantial temperature differences between urban, unvegetated areas, and areas of higher vegetation cover. These two factors - degree of vegetation cover and temperature - have played a significant role in helping to differentiate between the various clusters (compare Tables 21 and 22).

2.2.5 Comparison of Results with Other Reported Work

Of the roughly one dozen papers dealing with analyses of Skylab S-192 data for Land Use Mapping, which have either been published, or were presented at the Skylab Regional Planning and Development Conference, Purdue University, September 9-10, 1975, the present study is the only one which examines in some detail the within and between cluster group membership and affinities. Silva and Biehl (1975) followed the standard LARSYS procedure of using a clustering routine to confirm the number of discriminable land use classes before selecting training sets for determination of the most important channels in their analysis in the Wabash River Basin. Comparable procedures were used by Hoffer and Fleming (1975) in defining the best lands for forest cover mapping in mountainous areas of the southwest USA, and had earlier been used by Biehl and Silva (1975) in Land Use Studies in Lawrence County, Indiana and vicinity. These papers used clustering simply as a precursor to channel selection, primarily to increase the number of separable categories where discrimination within a land use class is feasible and proper.

The significance of the scrutiny given to the clustering in the present paper is that it indicates where the problem confusion areas lie and suggest that approaches through use of multi-date imagery would inprove the separation of land use categories. For example, there is a high level of consistency and classificatory rationality in the threshold 3 groups which were established. Knowing the composition of each group it is feasible to make extrapolations to other times and circumstances in the study are when better segregation of within-group confusion may be possible. For example, multi-date imagery would almost certainly lead to the elimination of the "growing crop present" category from cluster I, threshold 3 (Table 24-1) which is overwhelmingly urban (75 of 93 cases). It also may enable separation between the urban residential and non-residential areas, based on the multi-date day and night thermal imagery, and on leaf-out, flush growth, and senescence of the deciduous urban tree cover. At the same time, the trend urban residential areas included in what is primarily natural deciduous communities (cluster group 17: Table 24-1) should also be separated with multi-date coverage.

2.3 Multispectral Classification

The final analysis task using the S-192 Multispectral Scanner data involved performing a multispectral classification on all of the field pixels identified by the photointerpreters. In order to make maximum use of the broad spectral coverage available with the S-192 Multispectral Scanner data, a two-stage multispectral classification algorithm was designed. The first stage of the classification algorithm assigned each pixel to a general level I land use category. Based on the results of the Stage 1 classification, the second stage classification assigned a land use level II or level III to that pixel. In each stage only five spectral bands were used to make the assignment. The spectral bands used in the stage 1 classification were the best five spectral bands selected during the discriminant analysis between general level I categories. Different sets of spectral bands were used during the stage 2 classification depending on the particular level I land use category assignment made by the stage 1 classification. In total, 24,634 pixels were analyzed using this technique.

The results of each stage of the multispectral classification were stored in a data file along with information about the ground truth assignment of each pixel as well as the boundary structure of each pixel surrounding each classified pixel. It was then possible to produce various "confusion matrices" from the multispectral data file to investigate the original questions posed in the Executive Summary.

2.3.1 Analysis Techniques

Many types of multispectral classification algorithms have been implemented and tested in the past. The accuracy levels obtainable through various multispectral classification algorithms depends largely on the spectral characteristics of the data classes which are to be discriminated. Cost effectiveness is frequently an important consideration when deciding which multispectral classification algorithm to use for a particular application. While there are many factors to be considered, it is generally true that the more sophisticated classification algorithms, such as the maximum likelihood technique, provide higher levels of accuracy when the data classes display spectral overlap than do the less expensive algorithms such as the Euclidian distance algorithm or the Boolean classification technique. Because of the large variety of data classes considered in this investigation and hence the anticipation of large spectral overlap, it was decided that the maximum likelihood algorithm would be implemented in this work. It was learned during the step-wise discriminant analysis (see section 2.1) that the optimal spectral bands selected to discriminate between the various level II and III land use categories varied according to the different level I category being considered. In order to take advantage of the broad spectral coverage available with the S-192 Multispectral Scanner data while at the same time not incurring the cost of using all 13 spectral bands in the classification of each pixel, a two-staged classifier was designed. The first stage of the classification assigned an unknown vector of gray values (corresponding to a single pixel) to one of five general level I land use categories. The stage 1 classification used the five spectral bands identified through the step-wise discriminant analysis as providing the best discrimination between general level I categories. No thresholding was performed during the stage 1 classification. The results of the stage 1 classification were stored along with the vector of gray values.

After assigning each pixel to a general level I category, the second stage of the classification algorithm used the results of the first stage to further refine the classification assignment. A separate set of training signatures and spectral bands were defined for each general level I category. The best five spectral bands selected to discriminate within a specific level I category were used during the second stage. Suppose for example that a particular pixel was assigned to the level I category Urban (land use code 1000) by the stage 1 classifier. The second stage classification algorithm would then use a set of all urban land use level II and III training set signatures to assign a specific level II or III code to that pixel. The stage 2 classification algorithm would use the five best spectral bands identifed for discriminating within the category Urban.

This technique took advantage of the broad spectral coverage of the S-192 multispectral scanner data, reducing the amount of computation that would have been required had all 13 spectral bands been used in the assignment of each pixel. It has been estimated that a one-pass maximum likelihood classification using the same set of training signatures and all 13 spectral bands would have required over 13 times as much computation as performed by the two-stage classification algorithm.

A maximum likelihood classification algorithm was used during each of the two stages of the classification process. Consequently, each class was assumed to be normally distributed and completely described by a mean vector and variance matrix. The a priori probabilities of each class were assumed to be equal; that is, no one class had a higher probability of occurrence than another. Thresholding was performed during the second stage classification. A threshold of 2.0 standard deviations was applied to each class separately. After a pixel had been assigned to a category by the stage 2 classifier if the discriminant function had a value greater than 2.0 standard deviations for that particular class that pixel was assigned to the category "not classified."

2.3.2 Methodology

The two-stage classification algorithm required gray values for all 13 spectral bands for each pixel analyzed even though for any given pixel only a subset of the 13 bands were used. A special data file was generated for the multispectral classification analysis. From the test site field boundary/video data file only those pixels assigned to land use classes by the photo interpreters were transferred to the multispectral classification data file. The multispectral classification data file contained one record per pixel. Each record contained the following ground truth information:

1. Test Site Number,
2. Field Number,
3. Pixel Number,
4. Ground Truth Land Use Code Assignment,
5. Number of nearest neighbor boundary points,
6. Nearest neighbor boundary structure,
7. Gray values from each of the 13 spectral bands.

In addition, information about the results of the multispectral classification were entered into the file at each stage of the analysis. A total of 24,634 pixels were entered into this file.

The boundary structure and number of nearest neighbor boundary points was defined by the four perpendicular nearest neighbors (that is, the pixels to the left, right, above, and below on a 3 x 3 matrix centered on that particular pixel). Following that sequence, and assigning a binary value to each of the four neighbors, (0 if that particular neighbor was assigned to the same field as the center pixel, 1 if that neighbor was not assigned to the same field) a four-bit number uniquely defined both the number of neighbors to a particular pixel and the position of those neighbors with respect to that pixel.

All of the fields for which field signatures had been calculated were stratified into two groups; 1) training fields, and 2) test fields. The training fields were used to define the class signatures used in the multispectral classification analysis, and the test fields were used to test the accuracy of the classification. Several criteria were used in stratifying the fields between the training set and the test sets. First, an attempt was made to obtain an approximately equal (spatial) stratification between the training fields and the test fields. Second, the results of the clustering analysis (see section 2.2) was used to obtain an approximately equal stratification between all cluster groups. These two criteria were applied to the selection of training fields and test fields for both stages of the classification process. Finally, an attempt was made to equally divide the fields between the training fields and test fields. In total, 10 classes were defined for use by the stage 1 classification analysis and 42 classes were defined for use by the stage 2 classification. Table 27 shows the distribution of training fields and test fields obtained for the 42 classes used by the stage 2 classifier. Each pixel in the multispectral classification data file was flagged to indicate where it was belonged to a training field or a test field.

The sequence of processing was initiated by reading the group data which defined the number of classes within that group and the spectral band selected for use for that group. For example, the stage 1 classifier would have one group and 10 classes. Each class represented a specific level I land use code. Correspondingly the stage 2 classifier would have five groups and five channel selections based upon the different land use categories. Each class signature was obtained by mathematically combining the field signature assigned for training for that class. After each class in the group was completely defined those values used in the discriminant function which would not depend upon the unknown vector (such as the inverse of the variance

TABLE 27. Distribution of Training and Test Fields Used During the Stage 2 Classification

Class Number	Land Use Code	Number of Pixels	Percentage Training	Percentage Test
1	1110	5,741	48.7	51.3
2	1120	1,485	45.3	54.7
3	1210	302	36.1	63.9
4	1220	474	47.7	52.3
5	1240	115	57.4	42.6
6	1300	672	40.8	59.2
7	1400	271	36.3	63.7
8	1500	798	45.7	54.3
9	1600	1,385	32.0	68.0
10	1900	1,750	46.4	53.6
Subtotal All Urban		12,993	45.1	54.9
11	2112	237	48.5	51.5
12	2113	2,293	50.0	50.0
13	2113	1,649	43.7	56.3
14	2120	799	44.9	55.1
15	2142	1,325	42.9	57.1
16	2142	810	30.1	69.9
17	2143	167	65.3	34.7
18	2250	171	80.1	19.9
Subtotal All Agriculture		7,451	45.6	54.4
19	4100	253	54.9	45.1
20	4100	325	45.2	54.8
21	4100	412	42.0	58.0
22	4100	76	52.6	47.4
23	4100	52	48.1	51.9
24	4200	72	48.6	51.4
25	4200	66	50.0	50.0
26	4300	291	29.2	70.8
27	4300	90	46.7	53.3
28	4400	76	32.9	67.1
Subtotal All Forest		1,713	43.4	56.6

Class Number	Land Use Code	Number of Pixels	Percentage Training	Percentage Test
29	5110	184	37.5	62.5
30	5110	149	36.2	63.8
31	5110	40	40.0	60.0
32	5120	55	100.0	0.0
33	5210	86	34.9	65.1
34	5210	75	38.7	61.3
35	5300	54	100.0	0.0
36	5410	906	35.5	64.5
Subtotal All Water		1,549	40.6	59.4
37	6000	72	41.7	58.3
38	6000	77	26.0	74.0
39	6100	123	39.0	61.0
40	6100	36	44.4	55.6
41	6100	241	74.3	25.7
42	6200	379	16.1	83.9
Subtotal All Wetlands		928	38.1	61.9
GRAND TOTAL		24,634	44.6	55.4

matrix, the determinant of the variance matrix and the threshold) were calculated. Each unknown vector was read and the appropriate classification performed. The threshold value was tested and the appropriate land use value assignment was made. The updated data record was output to store the results of the process and the sequence was repeated for each of the pixels.

After all classifications were performed, results were evaluated by forming "confusion matrices" based upon various combinations of pixel characteristics. Each of the pixel file entries listed above could be used to exclude that particular pixel from being used in the calculation of any given confusion matrix. For example, it was important to differentiate between low classification percentages caused by the contribution of edge or boundary effects from the basic inability to distinguish between classes. For this investigation only those pixels listed as internal points could be included in the calculation of the confusion matrix. That confusion matrix could be compared against confusion matrices calculated using all pixels with one nearest neighbor, two nearest neighbors, and so forth.

The results of the multispectral classification and the various confusion matrices generated will be described in the next section.

2.3.3 Results

The results of the multispectral classification are presented in this section in the form of confusion matrices. First, the results of the stage 1 classification are discussed.

The results of the stage 2 classification are then presented both at land use level II and then again aggregated up to land use level I. The results of these analyses are shown for:

1) all training fields,
2) all test fields, and
3) aggregate of all fields (training plus test).

Several questions are then investigated by preparing confusion matrices using different pixel selection criteria. First, the accuracy of boundary delineation was investigated by comparing the following confusion matrices;

1) considering only those pixels which were internal field points,
2) considering only those pixels which had one nearest neighbor boundary, and
3) considering only those pixels which had two nearest neighbor boundaries.

Second, questions concerning the effect of the along scan line misregistration between spectral bands was investigated. Confusion matrices for all pixels with boundaries above and below the scanline direction were compared with similar confusion matrices for all pixels with boundaries in the along scan line direction.

2.3.3.1 Stage 1 Classification Results

Tables 28, 29 and 30 show the confusion matrices generated for the stage 1 classification results for the training fields, test fields, and total fields (training plus test) respectively. Note that no thresholding was performed during stage 1.

Table 28. MULTISPECTRAL CLASSIFICATION RESULTS: All Level I Training Pixels at Stage I
No Thresholding Used at Stage I

GROUND TRUTH \ CLASSIFICATION RESULTS (%)	NUMBER OF PIXELS	URBAN 1000	AGRICULTURAL 2000	FOREST 4000	WATER 5000	WETLANDS 6000
1000: URBAN	5859	68.4	20.4	9.0	0.8	1.4
2000: AGRICULTURAL	3400	12.6	74.5	11.1	0.3	1.5
4000: FOREST	744	2.7	3.2	86.8	0.9	6.4
5000: WATER	629	1.9	0.3	0.6	95.1	2.1
6000: WETLANDS	354	1.4	5.4	13.8	17.2	62.2
TOTAL PIXELS	10,986	4473	3773	1603	722	415

Overall accuracy in individual pixels:
8005 out of 10,986 or 72.9%

Table 29. MULTISPECTRAL CLASSIFICATION RESULTS: All Level I Test Pixels at Stage I
No Thresholding Used at Stage I

GROUND TRUTH \ CLASSIFICATION RESULTS (%)	NUMBER OF PIXELS	URBAN 1000	AGRICULTURAL 2000	FOREST 4000	WATER 5000	WETLANDS 6000
1000: URBAN	7134	73.8	18.3	6.2	0.3	1.4
2000: AGRICULTURAL	4051	21.1	65.2	11.4	0.0	2.3
4000: FOREST	969	5.1	6.5	82.0	1.2	5.2
5000: WATER	920	0.3	0.2	0.1	97.1	2.3
6000: WETLANDS	574	6.3	11.7	45.8	0.0	36.2
TOTAL PIXELS	13,648	6205	4079	1962	924	478

Overall accuracy in individual pixels:
9798 out of 13,648 or 71.8%

Table 30. MULTISPECTRAL CLASSIFICATION RESULTS: All Level I Pixels (Training + Test) at Stage I
No Thresholding Used at Stage I

GROUND TRUTH \ CLASSIFICATION RESULTS (%)	NUMBER OF PIXELS	URBAN 1000	AGRICULTURAL 2000	FOREST 4000	WATER 5000	WETLANDS 6000
1000: URBAN	12,993	71.3	19.3	7.4	0.5	1.5
2000: AGRICULTURAL	7,451	17.3	69.4	11.3	0.1	1.9
4000: FOREST	1,713	4.0	5.1	84.1	1.1	5.7
5000: WATER	1,549	1.0	0.3	0.3	96.3	2.1
6000: WETLANDS	928	4.4	9.3	33.6	6.6	46.1
TOTAL PIXELS	24,634	10,678	7,852	3,565	1,646	893

Overall accuracy in individual pixels:
17,803 out of 24,634 or 72.3%

Several comments should be made about these results. The apparent confusion between urban categories and agricultural categories is probably attributable to the spectral similarity between residential grass areas, open urban areas such as parks and golf courses, and agricultural pasture areas. The urban - agricultural confusion is observed for both the training set and the test set. Additionally, note that the confusion is two-way in that the misclassification of agricultural pixels as urban areas was approximately equal in magnitude to the misclassification of urban pixels as agricultural areas. The confusion observed between urban and forest areas is attributable to the presence of treed residential areas surrounding both Washington, D.C. and Baltimore, Maryland. The 11% misclassification of agricultural areas as forestland observed for both the training set and test set should be noted. More will be said about this misclassification when the Stage 2 results are discussed. The correct classification percentages for all forest areas (84.1%) and all water areas (96.3%) indicate good separability for these categories. The poor classification accuracy of the wetlands category is largely due to the presence of a category "forested wetlands". This category was poorly descriminated from the forest category in all the analyses performed during this investigation. The rather large difference in accuracy of correctly classified wetlands areas between the training set (62.2% correct) is explainable in terms of the distribution of the forested wetlands pixels. The forested wetlands category included only two fields totaling 379 pixels. One of the fields (62 pixels) was used for training and the other (318 pixels) for test purposes. Of all the wetlands pixels assigned to the training set only 17% were forested wetlands while 55% of the wetlands pixels assigned to the test set were forested wetlands. It is, therefore, not surprising that 45.8% of the wetlands test set pixels were misclassified as forest.

The accuracy levels observed for the Stage I classification are disappointingly low. The major areas of misclassification are, however, rationally explainable and, to a high degree, were predicted by the cluster analysis discussed in Section 2.2. The major land use category "mixes" observed within the cluster groups (see Table 24 and 25) are seen as major areas of misclassification in Table 30. Multidate repetitive coverage may be necessary to remove the major areas of misclassification seen in Table 30.

Finally, it should be noted that the overall accuracy of the training set (72.9%) was very close to the overall accuracy of the test set (71.8%). This indicates that the stratification between training and test fields provided an adequate distribution of all spectral data classes between the two data sets. Except as noted for the wetlands category there appears to be no significant difference between the training set results and the test set results. All additional tables will, therefore, be given for the aggregate of the training set plus the test set pixels only.

2.3.3.2 Stage 2 Classification Results

The results of the second stage classifier are shown in Tables 31 through 34. Tables 31 and 32 shows the results of the multispectral classification includes all pixels. Note that the only difference between the Stage 1 results shown in Table 30 and Level I (Table 32) is the effect of the threshold applied at Stage 2. Since thresholding cannot possibly improve the accuracy of the pixels correctly classified, the only effect was to reduce errors of misclassification. In some categories, pixels which had been properly classified were "threshold out" thus reducing the overall accuracy of correctly classified pixels. The most significant effect of the thresholding operation was for the land use category "wetlands." At Stage 1, many of the pixels assigned to the category "forested wetlands" were assigned to the category forest. The thresheld at Stage 2 reduced the percentage of pixels misclassified as forest from 33.6% to 18.7%.

The results of the second stage classification shown in Table 31 indicate generally low spectral separability between the land use Level II categories. The land use

Table 31. MULTISPECTRAL CLASSIFICATION RESULTS: All Pixels Classified By Stage 2. Results Shown for all Level II Categories. All Classes Thresheld at 2.0 Standard Deviations.

GROUND TRUTH	CLASSIFICATION RESULTS (PERCENTAGES)	NO. OF PIXELS	URBAN							AGRICULTURAL		FOREST				WATER				WETLANDS			NOT CLASSIFIED
			1100	1200	1300	1400	1500	1600	1900	2100	2200	4100	4200	4300	4400	5100	5200	5300	5400	6000	6100	6200	
URBAN	1100 - RESIDENTIAL	7226	40.6	10.7	2.0	6.5	2.8	1.3	6.9	15.4	1.2	5.0	0.2	2.4	1.0	2.0	0.1	0.3	0.0	0.1	0.1	0.4	3.1
	1200 - COMMERCIAL	891	13.1	44.4	5.8	7.2	7.3	2.5	5.1	7.2	1.3	1.5	0.2	0.2	0.1	0.0	0.0	0.2	0.0	0.0	0.0	0.0	3.8
	1300 - INDUSTRIAL	672	5.9	24.1	28.1	6.7	17.0	5.2	3.0	4.6	0.1	0.1	0.0	0.9	0.3	0.0	0.0	0.1	0.0	0.1	0.0	0.1	3.4
	1400 - EXTRACTIVE	271	8.1	10.7	31.0	24.4	0.7	3.3	2.6	4.8	2.6	2.6	0.0	1.1	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.4	6.6
	1500 - TRANSPORTATION	798	9.8	27.2	14.4	7.7	19.4	1.5	4.7	10.9	1.0	1.1	0.0	0.1	0.3	0.1	0.9	0.0	0.0	0.3	0.0	0.1	2.0
	1600 - INSTITUTIONAL	1385	22.3	15.1	1.7	6.6	5.6	3.9	17.5	15.2	0.8	4.3	0.0	2.8	1.4	0.0	0.0	0.0	0.0	0.1	0.0	0.6	2.0
	1900 - OPEN	1750	12.6	7.4	2.6	1.9	2.8	1.4	16.2	43.4	3.0	3.1	0.1	1.5	0.6	0.0	0.1	0.0	0.0	0.1	0.2	0.3	2.7
AGRICULTURAL	2100 - CROPLAND	7280	7.5	1.4	1.2	1.0	0.5	0.4	4.7	64.9	2.7	7.3	0.1	0.5	0.7	0.1	0.1	0.1	0.1	0.2	0.3	0.2	6.2
	2200 - ORCHARDS, TURF FARMS, ETC.	171	3.5	1.2	0.6	5.8	0.6	1.1	7.0	53.8	8.8	6.4	0.0	0.0	1.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9.4
FOREST	4100 - DECIDUOUS	1118	2.1	0.0	0.0	0.0	0.0	0.4	0.4	6.3	0.0	57.0	0.2	8.6	19.0	0.0	0.2	1.1	0.0	0.0	0.2	0.6	4.3
	4200 - CONIFEROUS	138	5.8	0.7	0.0	0.7	0.0	0.0	0.0	0.0	0.0	10.1	23.9	18.1	0.0	0.0	0.0	0.0	0.0	5.3	0.0	8.0	26.8
	4300 - MIXED	381	3.9	1.3	0.0	0.5	0.0	0.5	0.5	3.1	0.0	16.8	13.4	46.2	2.1	0.0	0.0	0.0	0.0	0.5	0.8	1.3	7.9
	4400 - UPLAND SHRUBS	76	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0	43.4	0.0	5.3	40.8	0.0	0.0	0.0	0.0	0.0	0.0	1.3	6.6
WATER	5100 - STREAMS & WATERWAYS	428	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	75.9	8.9	0.9	10.0	0.7	0.0	0.0	3.5
	5200 - LAKES & PONDS	161	0.0	0.6	0.0	0.0	0.6	0.0	0.0	1.2	0.0	0.6	0.0	0.6	0.0	49.7	31.7	0.0	4.3	0.6	0.0	1.2	8.7
	5300 - RESERVOIRS	54	3.7	9.3	3.7	0.0	0.0	0.0	0.0	1.8	1.8	1.8	0.0	0.0	0.0	24.1	7.4	18.5	5.6	0.0	1.8	0.0	20.4
	5400 - BAYS	906	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	46.7	13.5	0.7	36.1	0.0	0.0	0.0	3.1
WETLANDS	6000 - GENERAL	149	0.7	6.0	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	3.4	0.0	40.3	40.3	0.0	8.1
	6100 - VEGETATED	400	0.2	1.5	2.5	0.2	0.0	0.0	0.2	9.7	0.0	0.7	0.5	0.2	0.0	1.0	11.7	0.7	0.0	9.5	53.7	0.0	7.2
	6200 - FORESTED	379	1.8	0.0	0.0	0.5	0.0	0.5	0.0	10.8	0.3	34.0	0.0	4.5	5.8	0.0	0.0	0.0	0.0	0.0	0.0	1.6	40.1
TOTAL PIXELS		24,634	4332	2042	755	916	699	290	1495	7266	392	1929	112	607	430	850	279	64	382	144	3211	92	1237

PERCENTAGE ACCURACY OVERALL: 10,689 OUT OF 24,634 (43.4PERCENT

Table 32. MULTISPECTRAL CLASSIFICATION RESULTS: All Pixels Classified By Stage 2. Results Shown for Level I Categories. All Classes Thresheld at 2.0 Standard Deviations.

CLASSIFICATION RESULTS (%) / GROUND TRUTH	NUMBER OF PIXELS	URBAN 1000	AGRICULTURAL 2000	FOREST 4000	WATER 5000	WETLANDS 6000	NOT CLASSIFIED
1000: URBAN	12,993	70.5	18.9	6.7	0.3	0.6	3.0
2000: AGRICULTURAL	7,451	16.8	67.6	8.6	0.1	0.7	6.3
4000: FOREST	1,713	3.9	4.9	80.9	0.8	2.5	7.0
5000: WATER	1,549	0.7	0.3	0.2	94.0	0.5	4.4
6000: WETLANDS	928	4.4	8.7	18.7	6.5	40.8	20.8
TOTAL PIXELS	24,634	10,529	7,658	3,078	1,575	557	1,237

Overall Percentage Accuracy: 17,417 out of 24,634 (70.7 percent)

Table 33. MULTISPECTRAL CLASSIFICATION RESULTS: All Field Pixels With Field Quality Ratings of 4 or 5. Results of Stage 2 Classification Shown for all Level II Categories. All Classes Thresheld at 210 Standard Deviations.

GROUND TRUTH	CLASSIFICATION RESULTS (PERCENTAGES)	NO. OF PIXELS	URBAN							AGRICULTURAL		FOREST				WATER				WETLANDS			NOT CLASSIFIED
			1100	1200	1300	1400	1500	1600	1900	2100	2200	4100	4200	4300	4400	5100	5200	5300	5400	6000	6100	6200	
URBAN	1100 - RESIDENTIAL	6905	40.8	10.7	2.1	6.6	2.7	1.2	6.9	15.0	1.0	5.1	0.2	2.5	1.0	0.1	0.1	0.3	0.0	0.1	0.1	0.4	3.1
	1200 - COMMERCIAL	775	13.3	44.1	6.2	7.4	5.8	2.8	5.8	7.5	1.2	1.7	0.0	0.1	0.1	0.0	0.0	0.3	0.0	0.0	0.0	0.0	3.7
	1300 - INDUSTRIAL	646	6.2	25.1	25.9	7.0	17.6	5.4	2.8	4.5	0.2	0.2	0.0	0.9	0.3	0.0	0.0	0.2	0.0	0.2	0.0	0.2	3.6
	1400 - EXTRACTIVE	243	9.1	11.9	34.6	15.6	0.8	3.7	2.9	5.3	2.9	0.0	0.0	1.2	0.4	0.0	0.0	0.0	0.0	0.8	0.0	0.4	7.4
	1500 - TRANSPORTATION	757	9.8	28.4	15.1	8.1	17.8	1.6	4.5	10.3	1.1	0.4	0.0	0.0	0.1	0.1	0.4	0.0	0.0	0.4	0.0	0.0	2.1
	1600 - INSTITUTIONAL	1287	21.8	15.4	1.9	6.4	5.4	3.2	18.0	15.5	0.7	4.3	0.0	3.0	1.5	0.0	0.0	0.0	0.0	0.1	0.0	0.7	2.2
	1900 - OPEN	1739	12.7	7.4	2.6	1.9	2.8	1.4	15.8	93.6	3.0	3.1	0.1	1.5	0.6	0.0	0.1	0.0	0.0	0.1	0.2	0.3	2.8
AGRICULTURAL	2100 - CROPLAND	6658	7.4	1.2	0.8	1.0	0.5	0.4	4.4	66.6	2.5	7.5	0.1	0.5	0.7	0.1	0.1	0.1	0.1	0.1	0.1	0.1	5.9
	2200 - ORCHARDS, TURF FARMS, ETC.	152	2.0	1.3	0.7	6.6	0.7	0.7	6.6	52.6	9.9	7.2	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9.9
FOREST	4100 - DECIDUOUS	1118	2.1	0.0	0.0	0.0	0.0	0.4	0.4	6.3	0.0	57.0	0.2	8.6	19.0	0.0	0.2	1.1	0.0	0.0	0.2	0.6	4.3
	4200 - CONIFEROUS	129	6.2	0.8	0.0	0.8	0.0	0.0	0.0	0.0	0.0	8.5	24.8	17.1	0.0	0.0	0.0	0.0	0.0	6.2	0.0	7.8	27.9
	4300 - MIXED	364	4.1	1.3	0.0	0.5	0.0	0.5	0.5	1.1	0.0	17.6	14.0	48.4	2.2	0.0	0.0	0.0	0.0	0.5	0.0	1.4	7.7
	4400 - UPLAND SHRUBS	76	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0	43.4	0.0	5.3	40.8	0.0	0.0	0.0	0.0	0.0	0.0	1.3	6.6
WATER	5100 - STREAMS & WATERWAYS	421	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	77.2	9.0	1.0	10.2	0.0	0.0	0.0	2.6
	5200 - LAKES & PONDS	161	0.0	0.6	0.0	0.0	0.6	0.0	0.0	1.2	0.0	0.6	0.0	0.6	0.0	49.7	31.7	0.0	4.3	0.6	0.0	1.2	8.7
	5300 - RESERVOIRS	54	3.7	0.0	0.0	0.0	0.0	9.3	3.7	1.8	1.8	1.8	0.0	0.0	0.0	24.1	7.4	18.5	5.6	0.0	1.8	0.0	20.4
	5400 - BAYS	906	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	46.7	13.5	0.7	36.1	0.0	0.0	0.0	3.1
WETLANDS	6000 - GENERAL	149	0.7	6.0	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	3.4	0.0	40.3	40.3	0.0	8.1
	6100 - VEGETATED	400	0.2	1.5	2.5	0.2	0.0	0.0	0.2	9.7	0.0	0.7	0.5	0.2	0.0	1.0	11.7	0.7	0.0	9.5	53.7	0.0	7.2
	6200 - FORESTED	379	1.8	0.0	0.0	0.5	0.0	0.5	0.0	10.8	0.3	34.0	0.0	4.5	5.8	0.0	0.0	0.0	0.0	0.0	0.0	1.6	40.1
	TOTAL PIXELS	23,319	4114	1920	699	851	634	264	1399	6847	337	1876	107	594	422	850	279	64	382	135	295	87	1163

PERCENTAGE ACCURACY OVERALL: 10,157 OUT OF 23,319 (43 PERCENT)

Table 34. MULTISPECTRAL CLASSIFICATION RESULTS: All Field Pixels With Field Quality Ratings of 4 or 5. Results of Stage 2 Classification Shown for all Level I Categories. All Classes Thresheld at 2.0 Standard Deviations.

CLASSIFICATION RESULTS (%) / GROUND TRUTH	NUMBER OF PIXELS	URBAN 1000	AGRICULTURAL 2000	FOREST 4000	WATER 5000	WETLANDS 6000	NOT CLASSIFIED
1000: URBAN	12,352	70.4	18.8	6.8	0.3	0.6	3.1
2000: AGRICULTURAL	6,810	15.7	69.0	8.8	0.1	0.4	6.0
4000: FOREST	1,687	4.0	4.5	81.7	0.8	2.1	6.9
5000: WATER	1,542	0.7	0.3	0.2	94.4	0.3	4.1
6000: WETLANDS	928	4.4	8.7	18.7	6.5	40.8	20.8
TOTAL PIXELS	23,319	9,881	7,184	2,999	1,575	517	1,163

Overall Percentage Accuracy: 16602 out of 23,319 (71.1 percent)

The Level II agricultural land use classification showed an interesting result in that the misclassification of agricultural pixels as forest observed earlier in the Stage 1 results is seen here to be a confusion between agriculture and deciduous forest. Repetitive coverage should resolve this difficulty. The Level II forest and water analysis results moderate spectral separation between Level II categories with good separation at Level I.

In addition to the confusion matrices generated using all of the field pixels, matrices were generated using only those pixels assigned to fields with field quality ratings of 4 or 5. Tables 33 and 34 show the results of considering only those fields with high field quality ratings. By comparing Tables 31 and 32 (all pixels) with Tables 33 and 34 (only fields with quality ratings of 4 or 5) it is seen that at Level I all categories have an approximately equal or slightly higher correctly classified percentage accuracy for the quality 4 and 5 fields than for the results using all of the fields. The differences, however, are not as large as might have been expected. The significant result here is that 94.6% of all of the pixels identified by the photointerpreters were categories as 4 or 5 quality fields.

2.3.3.3 Boundary Delineation Accuracy

An important question frequently posed by land use planners is whether or not multispectral classification can be used to determine land use boundaries. While this is a very difficult question to address some insight into the problem can be gained by using the boundary structure information stored in the multispectral classification data file. Confusion matrices were generated considering only those pixels which were internal field pixels (no nearest neighbor boundary points). Tables 35 and 36 show the results of this analysis for land use Levels II and I respectively. These results were compared with confusion matrices generated for all pixels with one nearest neighbor boundary (see Tables 37 and 38) and for all pixels with two nearest neighbor boundaries (see Tables 39 and 40). The results of all of these analyses are summarized in Table 41. Table 41 shows that for the Level I categories agriculture, forest, and water, there was a difference of between 8.0 and 13.3 percent between the pixels with zero boundaries and the pixels with two boundaries.

An interesting anomaly occurred in the urban category. It appears as though the classification accuracy went up between the zero boundary the two boundary analysis. After close inspection it appears as though the category 1900 - Open Urban - is responsible for this effect. Apparently, the internal field points in the Open Urban category were frequently misclassified as agriculture. This is understandable since the Open Urban category includes parks, golf

ORIGINAL PAGE IS OF POOR QUALITY

Table 35. MULTISPECTRAL CLASSIFICATION RESULTS: All Internal Field Pixels With Field Quality Ratings of 4 or 5. Results of Stage 2 Classification Shown for Level II Categories. All Classes Thresheld at 2.0 Standard Deviations.

GROUND TRUTH	CLASSIFICATION RESULTS (PERCENTAGES)	NO. OF PIXELS	URBAN 1100	1200	1300	1400	1500	1600	1900	AGRICULTURAL 2100	2200	FOREST 4100	4200	4300	4400	WATER 5100	5200	5300	5400	WETLANDS 6000	6100	6200	NOT CLASSIFIED
URBAN	1100 - RESIDENTIAL	3712	42.9	10.8	2.6	6.0	2.6	1.1	6.3	13.9	0.8	5.0	0.2	2.6	0.9	0.0	0.1	0.4	0.0	0.1	0.2	0.4	3.4
	1200 - COMMERCIAL	331	8.8	52.9	6.6	6.6	7.6	3.0	3.9	4.5	1.2	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.9
	1300 - INDUSTRIAL	316	4.7	21.8	32.0	4.1	20.6	7.3	2.2	2.5	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.3	3.5
	1400 - EXTRACTIVE	130	10.0	5.4	38.5	23.8	0.0	4.6	3.1	3.8	0.0	1.5	0.0	0.8	0.0	0.0	0.0	0.0	0.0	0.8	0.0	0.8	6.9
	1500 - TRANSPORTATION	303	6.9	31.4	15.5	6.9	19.5	2.6	2.6	11.6	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
	1600 - INSTITUTIONAL	638	24.0	13.5	0.8	7.1	4.9	3.6	22.7	13.5	0.6	2.7	0.0	3.4	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.9	1.9
	1900 - OPEN	917	9.4	4.0	1.2	1.3	1.1	1.5	16.5	52.9	4.8	2.9	0.0	1.4	0.4	0.0	0.0	0.0	0.0	0.0	0.1	0.0	2.4
AGRICULTURAL	2100 - CROPLAND	2829	7.7	0.9	0.5	0.7	0.5	0.2	4.1	69.7	2.6	6.3	0.0	0.1	0.3	0.0	0.1	0.1	0.0	0.1	0.1	0.1	6.1
	2200 - ORCHARDS, TURF FARMS, ETC.	46	0.0	0.0	0.0	4.3	2.2	2.2	4.3	65.2	8.7	4.3	0.0	0.0	2.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.5
FOREST	4100 - DECIDUOUS	513	1.0	0.0	0.0	0.0	0.0	0.4	0.0	3.9	0.0	58.8	0.0	9.3	21.8	0.0	0.4	0.6	0.0	0.0	0.2	0.4	3.5
	4200 - CONIFEROUS	31	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.2	41.9	12.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	41.9
	4300 - MIXED	183	3.2	1.1	0.0	0.0	0.0	0.5	0.0	0.0	0.0	13.7	12.6	59.0	2.2	0.0	0.0	0.0	0.0	0.0	0.0	0.5	7.1
	4400 - UPLAND SHRUBS	27	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	29.6	0.0	11.1	51.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.4
WATER	5100 - STREAMS & WATERWAYS	202	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	74.3	7.4	0.5	15.3	0.0	0.0	0.0	2.5
	5200 - LAKES & PONDS	78	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	64.1	29.5	0.0	5.1	0.0	0.0	0.0	1.3
	5300 - RESERVOIRS	28	0.0	7.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	39.3	10.7	14.3	10.7	0.0	0.0	0.0	17.9
	5400 - BAYS	711	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	48.4	13.2	0.7	35.1	0.0	0.0	0.0	2.5
WETLANDS	6000 - GENERAL	63	0.0	6.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.2	0.0	38.1	50.8	0.0	1.6
	6100 - VEGETATED	210	0.5	0.5	2.4	0.0	0.0	0.0	0.0	5.3	0.0	0.5	0.5	0.0	0.0	1.9	18.6	0.9	0.0	7.1	55.7	0.0	6.2
	6200 - FORESTED	286	2.1	0.0	0.0	0.3	0.0	0.7	0.0	12.6	0.0	32.2	0.0	3.1	5.6	0.0	0.0	0.0	0.0	0.0	0.0	1.0	42.3
	TOTAL PIXELS	11,554	2143	905	350	391	302	136	681	3218	166	841	43	308	195	559	182	32	288	44	158	30	582

PERCENTAGE ACCURACY OVERALL: 5041 OUT OF 11,554 (44 PERCENT)

Table 36 . MULTISPECTRAL CLASSIFICATION RESULTS: All Internal Field Pixels With Field Quality Ratings of 4 or 5. Results of Stage 2 Classification Shown for Level I Categories. All Classes Thresheld at 2.0 Standard Deviations.

CLASSIFICATION RESULTS (%) / GROUND TRUTH	NUMBER OF PIXELS	URBAN 1000	AGRICULTURAL 2000	FOREST 4000	WATER 5000	WETLANDS 6000	NOT CLASSIFIED
1000: URBAN	6347	70.1	19.5	6.5	0.3	0.5	3.1
2000: AGRICULTURAL	2875	14.6	72.3	6.7	0.1	0.1	6.1
4000: FOREST	754	2.1	2.6	87.9	0.7	0.5	6.1
5000: WATER	1019	0.2	0.0	0.0	97.0	0.0	2.8
6000: WETLANDS	559	3.6	8.4	21.3	8.4	34.2	24.2
TOTAL PIXELS	11,554	4908	3384	1387	1061	232	582

Percentage Accuracy Overall: 8373 out of 11,554 (72.5 percent)

Table 37. MULTISPECTRAL CLASSIFICATION RESULTS: All Field Pixels With Field Quality Ratings of 4 or 5 and 1 Nearest Neighbor Boundary Pixel. Results of Stage 2 Classification Shown for Level II Categories. All Classes Thresheld at 2.0 Standard Deviations.

GROUND TRUTH \ CLASSIFICATION RESULTS (PERCENTAGES)		NO. OF PIXELS	URBAN 1100	URBAN 1200	URBAN 1300	URBAN 1400	URBAN 1500	URBAN 1600	URBAN 1900	AGRICULTURAL 2100	AGRICULTURAL 2200	FOREST 4100	FOREST 4200	FOREST 4300	FOREST 4400	WATER 5100	WATER 5200	WATER 5300	WATER 5400	WETLANDS 6000	WETLANDS 6100	WETLANDS 6200	NOT CLASSIFIED
URBAN	1100 - RESIDENTIAL	1776	38.4	9.7	1.6	7.5	3.0	1.4	7.9	16.2	1.1	5.6	0.1	2.5	1.2	0.0	0.1	0.2	0.0	0.2	0.2	0.5	2.7
	1200 - COMMERCIAL	235	14.9	42.6	4.7	8.5	4.3	3.0	4.7	11.1	0.8	2.1	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0	3.0
	1300 - INDUSTRIAL	144	4.9	22.9	19.4	9.7	18.8	3.5	9.0	0.0	0.0	0.7	0.7	1.4	0.0	0.0	0.0	0.0	0.0	0.7	0.0	0.0	4.9
	1400 - EXTRACTIVE	74	1.4	17.6	33.8	14.9	0.0	4.1	0.0	8.1	8.1	1.4	0.0	2.7	1.4	0.0	0.0	0.0	0.0	1.4	0.0	0.0	5.4
	1500 - TRANSPORTATION	184	13.0	21.7	14.1	10.9	17.9	0.5	5.4	12.0	1.1	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	1.6
	1600 - INSTITUTIONAL	305	20.7	15.1	3.0	6.6	5.2	2.6	12.8	17.7	1.3	5.9	0.0	2.0	3.6	0.0	0.0	0.0	0.0	0.0	0.0	1.0	2.6
	1900 - OPEN	436	12.8	9.2	1.6	2.1	4.4	2.3	17.0	38.5	1.6	3.4	0.5	1.4	0.7	0.0	0.2	0.0	0.0	0.0	0.5	0.7	3.2
AGRICULTURAL	2100 - CROPLAND	2145	6.7	1.2	1.0	1.0	0.6	0.7	4.4	65.7	2.8	8.1	0.2	0.5	0.7	0.1	0.0	0.1	0.1	0.2	0.1	0.2	5.6
	2200 - ORCHARDS, TURF FARMS, ETC.	70	1.4	0.0	1.4	8.6	0.0	0.0	5.7	50.0	11.4	8.6	0.0	0.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	11.4
FOREST	4100 - DECIDUOUS	410	2.7	0.0	0.0	0.0	0.0	0.0	1.0	5.3	0.0	59.0	0.0	7.5	15.9	0.0	0.0	2.0	0.0	0.0	0.0	0.7	4.9
	4200 - CONIFEROUS	56	10.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.1	8.9	19.6	0.0	0.0	0.0	0.0	0.0	10.7	0.0	14.3	28.6
	4300 - MIXED	141	2.8	1.4	0.0	1.4	0.0	0.7	0.7	1.4	0.0	19.9	14.2	41.1	2.8	0.0	0.0	0.0	0.0	1.4	0.0	2.8	9.2
	4400 - UPLAND SHRUBS	36	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.8	0.0	44.4	0.0	2.8	44.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.6
WATER	5100 - STREAMS & WATERWAYS	176	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	79.5	9.6	1.1	6.8	0.0	0.0	0.0	2.8
	5200 - LAKES & PONDS	56	0.0	1.8	0.0	0.0	0.0	0.0	0.0	1.8	0.0	1.8	0.0	0.0	0.0	35.7	33.9	0.0	3.6	1.8	0.0	1.8	17.8
	5300 - RESERVOIRS	16	6.3	12.5	6.3	0.0	0.0	0.0	0.6	6.3	0.0	0.0	0.0	0.0	0.0	6.3	0.0	37.5	0.0	0.0	6.3	0.6	18.7
	5400 - BAYS	155	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	38.7	12.3	0.6	43.2	0.0	0.0	0.0	5.2
WETLANDS	6000 - GENERAL	64	1.6	6.3	1.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.7	0.0	40.6	35.9	0.0	9.4
	6100 - VEGETATED	154	0.0	2.6	2.6	0.6	0.0	0.0	0.6	12.3	0.0	0.0	0.6	0.6	0.0	0.0	5.2	0.0	0.0	12.3	55.2	0.0	7.1
	6200 - FORESTED	75	0.0	0.0	0.0	1.3	0.0	0.0	0.0	4.0	1.3	38.7	0.0	8.0	6.7	0.0	0.0	0.0	0.0	0.0	0.0	4.0	36.0
	TOTAL PIXELS	6708	1035	482	164	260	171	75	385	2074	110	640	34	177	145	223	66	25	82	64	116	39	341

PERCENTAGE ACCURACY OVERALL: 3021 OUT OF 6708 (45 PERCENT)

Table 38 . MULTISPECTRAL CLASSIFICATION RESULTS: All Field Pixels With Field Quality Ratings of 4 or 5 and 1 Nearest Neighbor Boundary Pixel. Results of Stage 2 Classification Shown for Level I Categories. All Classes Thresheld at 2.0 Standard Deviaitons.

CLASSIFICATION RESULTS (%) / GROUND TRUTH	NUMBER OF PIXELS	URBAN 1000	AGRICULTURAL 2000	FOREST 4000	WATER 5000	WETLANDS 6000	NOT CLASSIFIED
1000: URBAN	3154	68.8	19.6	7.7	0.2	0.8	2.9
2000: AGRICULTURAL	2215	15.7	68.3	9.5	0.2	3.5	5.8
4000: FOREST	643	4.8	4.5	77.9	1.2	3.6	7.9
5000: WATER	403	1.2	0.5	0.2	90.8	0.7	6.5
6000: WETLANDS	293	5.8	7.8	14.3	3.8	53.2	15.0
TOTAL PIXELS	6708	2572	2184	996	396	219	341

Percentage Accuracy Overall: 4707 out of 6708 (70.2 percent)

Table 39. MULTISPECTRAL CLASSIFICATION RESULTS: All Field Pixels With Field Quality Ratings of 4 or 5 and 2 Nearest Neighbor Boundary Pixels. Results of Stage 2 Classification Shown for Level II Categories. All Classes Thresheld at 2.0 Standard Deviations.

GROUND TRUTH \ CLASSIFICATION RESULTS (PERCENTAGES)		NO. OF PIXELS	URBAN							AGRICULTURAL		FOREST				WATER				WETLANDS			NOT CLASSIFIED
			1100	1200	1300	1400	1500	1600	1900	2100	2200	4100	4200	4300	4400	5100	5200	5300	5400	6000	6100	6200	
URBAN	1100 - RESIDENTIAL	1283	39.6	11.8	1.6	7.0	2.2	1.4	7.2	16.1	1.1	4.5	0.2	2.3	0.8	0.2	0.1	0.3	0.0	0.2	0.0	0.4	3.0
	1200 - COMMERCIAL	197	18.8	38.6	6.6	7.6	2.5	2.0	7.1	7.6	1.5	2.0	0.0	0.5	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	4.6
	1300 - INDUSTRIAL	172	8.7	30.8	25.0	9.3	11.0	2.9	2.9	4.1	0.0	0.6	0.0	1.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.9
	1400 - EXTRACTIVE	51	11.8	17.6	13.7	27.5	3.9	0.0	3.9	3.9	2.0	7.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.8
	1500 - TRANSPORTATION	236	9.7	29.7	11.4	7.6	19.1	1.3	5.9	8.0	0.0	0.4	0.0	0.0	0.4	0.4	1.3	0.0	0.0	0.4	0.0	0.0	4.2
	1600 - INSTITUTIONAL	301	16.9	19.9	3.0	4.7	7.3	4.7	13.0	16.6	0.3	6.3	0.0	3.3	1.7	0.0	0.0	0.0	0.0	0.3	0.0	0.0	2.0
	1900 - OPEN	344	19.2	13.1	6.7	2.9	4.4	0.3	13.1	30.2	0.3	3.2	0.0	1.7	0.9	0.0	0.3	0.0	0.0	0.3	0.0	0.3	3.2
AGRICULTURAL	2100 - CROPLAND	1592	8.0	1.6	1.1	1.3	0.5	0.3	4.8	62.3	2.2	8.6	0.2	1.0	1.2	0.0	0.0	0.0	0.1	0.0	0.0	0.6	6.1
	2200 - ORCHARDS, TURF FARMS, ETC.	30	3.3	3.3	0.0	6.7	0.0	0.0	10.0	46.7	6.7	10.0	0.0	0.0	3.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	10.0
FOREST	4100 - DECIDUOUS	189	3.7	0.0	0.0	0.0	0.0	1.1	0.0	12.2	0.0	49.2	1.1	9.0	16.9	0.0	0.0	0.0	0.0	0.0	0.5	1.0	5.3
	4200 - CONIFEROUS	38	5.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	15.8	36.8	13.2	0.0	0.0	0.0	0.0	0.0	5.3	0.0	5.3	18.4
	4300 - MIXED	39	10.3	2.6	0.0	0.0	0.0	0.0	2.6	5.1	0.0	28.2	20.5	25.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.1
	4400 - UPLAND SHRUBS	13	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.7	0.0	69.2	0.0	0.0	7.7	0.0	0.0	0.0	0.0	0.0	0.0	7.7	7.7
WATER	5100 - STREAMS & WATERWAYS	41	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	80.5	14.6	2.4	0.0	0.0	0.0	0.0	2.4
	5200 - LAKES & PONDS	27	0.0	0.0	0.0	0.0	3.7	0.0	0.0	3.7	0.0	0.0	0.0	3.7	0.0	37.0	33.3	0.0	3.7	0.0	0.0	3.7	11.1
	5300 - RESERVOIRS	10	10.0	10.0	10.0	0.0	0.0	0.0	0.0	0.0	10.0	10.0	0.0	0.0	0.0	10.0	10.0	0.0	0.0	0.0	0.0	0.0	30.0
	5400 - BAYS	39	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	48.7	20.5	0.0	25.6	0.0	0.0	0.0	5.1
WETLANDS	6000 - GENERAL	22	0.0	4.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.5	0.0	0.0	45.5	22.7	0.0	22.7
	6100 - VEGETATED	36	0.0	2.8	2.3	0.0	0.0	0.0	0.0	25.0	0.0	5.6	0.0	0.0	0.0	0.0	0.0	2.8	0.0	11.1	36.1	0.0	13.9
	6200 - FORESTED	15	6.7	0.0	0.0	0.0	0.0	0.0	0.0	13.3	0.0	46.7	0.0	6.7	6.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	20.0
	TOTAL PIXELS	4612	845	494	162	199	144	52	289	140	56	362	29	100	73	66	30	7	12	22	19	21	222

PERCENTAGE ACCURACY OVERALL: 1893 OUT OF 4612 (41 PERCENT)

Table 40. MULTISPECTRAL CLASSIFICATION RESULTS: All Field Pixels With Field Quality Ratings of 4 or 5 and 2 Nearest Neighbor Boundary Pixels. Results of Stage 2 Classification Shown for Level I Categories. All Classes Thresheld at 210 Standard Deviations.

GROUND TRUTH \ CLASSIFICATION RESULTS (%)	NUMBER OF PIXELS	URBAN 1000	AGRICULTURAL 2000	FOREST 4000	WATER 5000	WETLANDS 6000	NOT CLASSIFIED
1000: URBAN	2584	72.9	16.4	6.5	0.5	0.5	3.3
2000: AGRICULTURAL	1559	17.8	64.3	11.2	0.1	0.6	6.2
4000: FOREST	279	6.1	9.3	74.6	0.0	2.9	7.2
5000: WATER	117	3.4	1.7	1.7	84.6	0.9	7.7
6000: WETLANDS	73	5.5	15.1	15.1	2.7	43.8	17.8
TOTAL PIXELS	4612	2185	1464	564	115	62	222

Percentage Accuracy Overall: 3224 out of 4612 (69.9 percent)

Table 41. Summary of Percentage Correctly Classified as a Function of the Number of Nearest Neighbor Boundaries (Level I only)

	Number of Nearest Neighbor Boundaries *		
	Zero	One	Two
1000: Urban	70.1	68.8	72.9
2000: Agricultural	72.3	68.3	64.3
4000: Forest	87.9	77.9	74.6
5000: Water	97.0	90.8	84.6
6000: Wetlands	34.2	53.2	43.8
Overall	72.5	70.2	69.9

* Only fields with quality ratings of 4 Or 5 were considered.

courses, and generally grassed areas. The boundary pixels assigned to the Open Urban fields were also misclassified but they were misclassified as other urban categories. Thus, at Level I, the overall urban classification was higher for two boundaries than for internal points.

The accuracy levels achievable through multispectral classification determine whether or not this technique may be effectively used to determine land use boundaries. The analysis performed here indicates that the accuracy levels observed for internal field pixels may be significantly higher (as much as 13%) higher in this analysis) than the accuracy levels achievable for pixels near to, or crossing over, land use boundaries.

2.3.3.4 Effect of Misregistration Between Spectral Bands

The last question investigated through analysis of the multispectral classification results addresses the effect of the registration between bands on the classification accuracy. The band-to band misregistration effect should have influenced most strongly those pixels which had field boundaries in the along-scanline direction. Were it not for this misregistration effect one would expect the classification accuracies for all pixels with boundaries in the along-scanline direction to be approximately equal to the classification accuracies observed for all pixels with boundaries above or below the scanline direction. Table 42 shows the percentage accuracy for all pixels correctly classified which had nearest neighbor boundaries in the along-scanline direction. Table 43 shows the results for all pixels correctly classified which had nearest neighbor boundaries above or below the scanline. While these tables do not provide conclusive evidence that the misregistration between spectral bands degraded the classification accuracy, there is some indication that higher accuracies were obtained when the boundaries were above or below the scanline direction. The only level I category which showed significantly higher classification accuracies for pixels with boundaries along the scanline direction was the land use category "water." In all other level I categories the accuracies were approximately equal or slightly better for the pixels with boundaries above or below the scanline direction than for pixels with boundaries in the open along scanline direction. The overall accuracy difference between pixels with boundaries along the scanline versus pixels with boundaries above or below the scanline was slightly better than one percent.

Table 42. MULTISPECTRAL CLASSIFICATION RESULTS: Pixels With Nearest Neighbor Boundaries in the Along Scanline Direction.

	Number of Pixels	Percent Correctly Classified Level III	Level II	Level I
1000: URBAN	1578	20.0	29.8	67.2
2000: AGRICULTURAL	1094	35.2	61.4	66.9
4000: FOREST	320	48.8	48.8	74.7
5000: WATER	202	52.0	54.0	93.1
6000: WETLANDS	119	36.1	36.1	48.7
OVERALL	3313	30.3	43.8	68.7

Table 43. MULTISPECTRAL CLASSIFICATION RESULTS: Pixels With Nearest Neighbor Boundaries Above or Below the Scanline.

	Number of Pixels	Level III	Level II	Level I
1000: URBAN	1792	18.5	28.7	70.2
2000: AGRICULTURAL	1395	33.8	63.2	66.7
4000: FOREST	335	49.3	49.3	78.8
5000: WATER	210	56.7	60.0	86.7
6000: WETLANDS	174	40.8	40.8	56.3
OVERALL	3906	29.6	45.0	70.0

2.3.4 Implications from the Results of the Multispectral Classification Analyses

The multispectral classification results reported here were somewhat disappointing in that the land use Level I classification accuracies were lower than anticipated. There are several possible explanations for the observed results. The time of year of the overpass (August 5, 1973) was not optimal for discriminating many categories. The time of day of the overpass (11:00 AM EST) may also have been a factor. The misclassification observed, however, are to a large degree understandable. Urban grassed and treed areas were identified as agriculture and forest respectively. Forested wetlands were identified as forest. There is an apparent high degree of "internal consistency" with the results presented. The size of the data set analyzed (24, 634 pixels) leads one to believe that for the most part, the results observed are "real" and not random statistical fluctuations. The results do strongly indicate the need for repetitive year round coverage to improve overall classification accuracy since no single date is optimal for all land use categories.

The accuracy levels obtained for land use Level II must be judged as unacceptable for most land use planning purposes. These accuracy levels may be improved by multidate coverage but there still appears to be a "natural" spectral similarity between many of the Level II categories. Open Urban grassed and treed areas appear spectrally similar to residential grassed and treed areas. Forested wetlands appear spectrally similar to forested areas. These categories may never be adequately discriminated with resolution similar to that available with the S-192 Skylab Multispectral Scanner.

The question of boundary delineation is a difficult question to address analytically but there is some indication that generally there is a significant drop in classification accuracy (as much as 13%) as one moves from internal field points to field boundary points. The implication of this drop is that in complex areas of small entities, or where there is spatial variability <u>within</u> a natural class, that it will be difficult to erect the physical boundaries between these entities using pattern recognition procedures unless one moves to higher spatial resolutions, perhaps of the order of those of LANDSAT D.

3.0 ENVIRONMENTAL MAPPING WITH SPACE IMAGERY: A CENTRAL AUSTRALIAN EXAMPLE

In two companion papers (Rohde and Simonett, 1975; Bale et al, 1975) we have discussed the problems of thematic land use mapping from space by focusing attention on cultural landscape line, point and areal categories as observed in SKYLAB S-190A and S-190B color IR and color imagery of the Washington D.C. - Baltimore, MD. test site. In this paper we are concerned primarily with different problems, namely those of natural plant community-landscape boundary delineation, and entity discrimination with the SKYLAB images.

The example to be discussed concerns resource mapping of a natural environment near Alice Springs, Central Australia, where both boundary detection and categorization are complicated by the inherent complexity of landscape elements. In this environment we are not dealing with the same degree of patterned regularity found in cultural landscapes; nor are we dealing with more or less discrete entities such as crop types and roads, or cultural vs. natural phenomena. Instead nature has provided in this region a continuous variation in space of the several elements: terrain, soil surface, and vegetation. We know from principles of geography and ecology that such variation is not random, and were we to study it on (or near) the ground, we could eventually decipher much of the intricacies of the patterns. When viewed from space, however, the meaning and composition of boundaries, to say nothing of the "things" they separate, are to some degree ambiguous, being both system and interpreter-dependent.

Two themes will be pursued in this study to illustrate some advantages and disadvantages of using space images in tropical arid-land resource inventories. The first concerns itself with the detection and meaning of boundaries, the second with sources of confusion during categorization.

In the discussion comparisons will be made between interpretations of the SKYLAB images and earlier studies in the same area with Gemini photography by Simonett et al (1969), and by Story, Yapp and Dunn (1976) with LANDSAT images. Comparisons of boundaries in the field with those on the space photograph, as well as comparisons using low altitude obliques and photo mosaics, have shown that even small pin pricks of space data relate to qualitative changes in the landscape. Space boundaries may also be easier to detect than are identical boundaries on photo mosaics.

Detecting a boundary and knowing that the landscape is somehow different either side of it is not the same as knowing either the nature or magnitude of that difference, nor is it to be presumed that the boundaries lie all at the same hierarchical level in a classification. Since generalizing is unavoidable in space photography due to current resolution limitations, boundaries may result from changes in one or several features of the environment, none or all of which may be significant in a particular resource inventory. While the boundary has meaning, therefore, it may not be one we wish to map.

Categorization is also related to this problem of generalization. Image discrimination functions such as texture, height, etc. have restricted values in space photo interpretation. Tone is the most versatile of the image qualities but its limitation should be appreciated. Disjunct shades of similar color on the space photo sometimes relate to dissimilar combinations of elements in the landscape. Equally serious, dissimilar colors sometimes contain a similar combination of elements but in different proportions or under different illumination. In both cases substantial errors of interpretation arise, even among experienced interpreters.

The Alice Springs area was photographed with Ektachrome MS Aerographic 70 MM film in August 1965 by the crew of Gemini V. LANDSAT I images were obtained during the period October 1972 - February 1973. SKYLAB S-190A and S-190B photographs were taken on August 12, 1973 (unuseable because of cloud cover) and on September 29, 1973. Reconnaissance-scale land mapping was completed by Perry et al in 1962, and forms the basis for comparison with all these later images. Figure 12 is an enlargement of the Gemini photograph on which the main geographic features have been identified. This area was selected for study because it is representative of the mapping problems to be encountered in very large regions of the semi-arid and arid tropics. In addition, the original photograph is of acceptable to good quality despite a substantial haze scattering in the blue and to a lesser degree even the green sensitive layers (Figure 18). The principal investigator (Simonett) and a number of his previous co-authors (Cochrane, Morain) have been to the Alice Springs area on separate occasions during the period 1968-1972 and have spent fifteen weeks in the field studying the soils, vegetation and topography. The co-operating Australian Scientists who provided the ground truth and low altitude imagery at the time of the SKYLAB imaging have very intimate field knowledge of the area. Appreciation is extended here for the co-operation by Mr. Ray Perry of CISRO and his colleagues Dr. Max Ruff, Robert Winkworth, and Robert Millington. Field work is immensely important from the point of view of categorizing areas delineated on the photo. In addition, one of us (Simonett) has conducted aerial and ground reconnaissance of the region with space-photo-in-hand for purposes of comparing boundaries, and obtaining low altitude aerial oblique and ground photos for laboratory comparisons. Finally, although the area is relatively remote, the natural environment is well known thanks to the efforts of R.A. Perry and his colleagues at CISRO Division of Land Research. Perry's (1961) pasture map of the area is particularly valuable as a source of information and comparison.

3.1 NATURAL FEATURES OF THE ALICE SPRINGS AREA

The Alice Springs study area (Figure 12), through the center of which the SKYLAB track passed, covers almost 21,000 square miles of country in semi-arid central Australia. It stretches from the James and Krichauff ranges in the south, and includes most of Missionary Plain, a large part of the Macdonnell and Chewings Ranges, all of Napperby Lake and Stuart Bluff Range and terminates at Mt. Denison in the north. Alice Springs itself is located on the lower right margin of the Gemini photo. The following brief discussion of landscape types draws heavily from the works of Perry (1961) and Perry et. al. (1962).

Physiographically four major landscape divisions are delimited on the Perry Pasture Map. These are:

1. Folded Ranges: represented on the photo by the James, Waterhouse, Macdonnell, Hann and Stuart Bluff Ranges.
2. Crystalline Uplands: Reynolds ranges, Crown Hill, Pine Hill.
3. Crystalline Ranges: Strangways Range, Mt. Chapple, Mt. Hay, Mt. Zeil, Mt. Heughlin, Redbank Hill.
4. Northern Plains: Burt Plain, Everard Scrub, Missionary Plain (the last is not included by Perry as part of this category).

All of the ranges and uplands have large bare-rock outcrops and skeletal

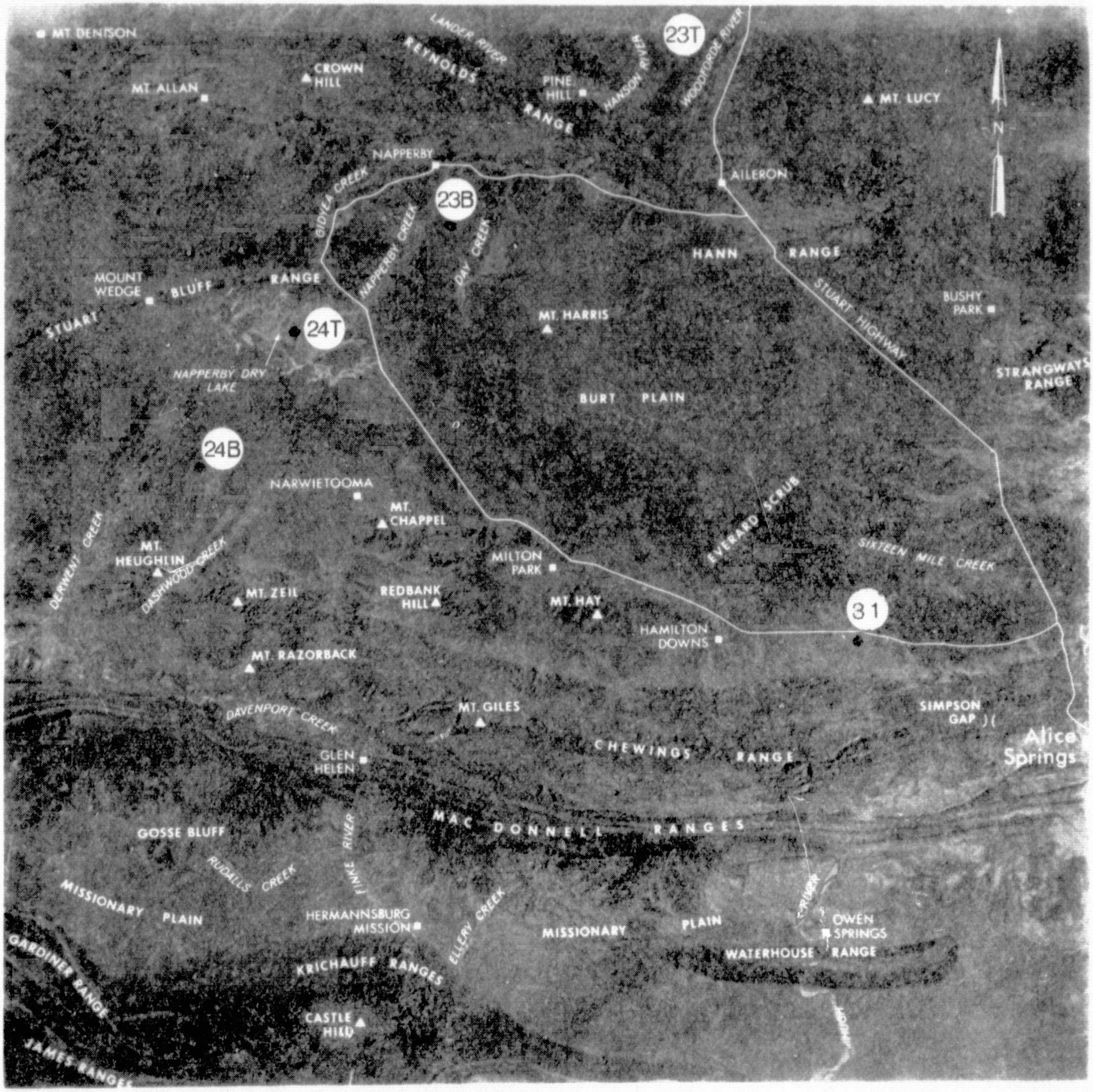


Figure 12. Location Map for the region northwest of Alice Springs, Central Australia. The numbers give the locations of air photo mosaics in Figures 24 and 25, and the location of Kunoth Wells Paddock illustrated in the enlargement of SKYLAB S-190B imagery in Figure 32.

or shallow, stony soils. In the plains area, soils are generally characterized as red sands, red clayey sands and red earths. Saline soils and unconsolidated sands are usually found along drainage lines and at Napperby Lake.

When vegetation types are superimposed on the physiographic and soil patterns, seven broad landscape types may be recognized. Figure 13 illustrates the distribution of these environments as mapped by Perry et al. (1962). Where possible the broad categories are defined in terms of vegetation even though in most circumstances the plant cover is open or very sparse. What is actually recorded on the Gemini photograph is the spectral reflectance not only of vegetation but of rock and soil surface as well.

The seven landscape types delineated on Figure 13 are:

1. Mountains and Hills
2. Alternating Hills and Lowlands
3. Salt Lake and Pans
4. Grass-Forb Pasture on Young Alluvia
5. Mitchell Grass
6. Low Trees and Shrubs, Mainly Mulga (Acacia Anuera)
7. Spinifex Sand Plains and Dune Fields

Of these landscape units the Mountains and Hills and Alternating Hills and Lowlands categories have intricate mixtures of the other five categories recognized. The scale of these mixtures is far too small to detect or map on the Gemini photo. Some discrimination of the larger entities may be feasible on say 10 X enlargements but uncertainties arising from shadowing and highlighting will prevent even modestly reliable categorization.

LANDSCAPES NORTHWEST OF ALICE SPRINGS, CENTRAL AUSTRALIA

Base: C.S.I.R.O. Pasture Land Map, 1961

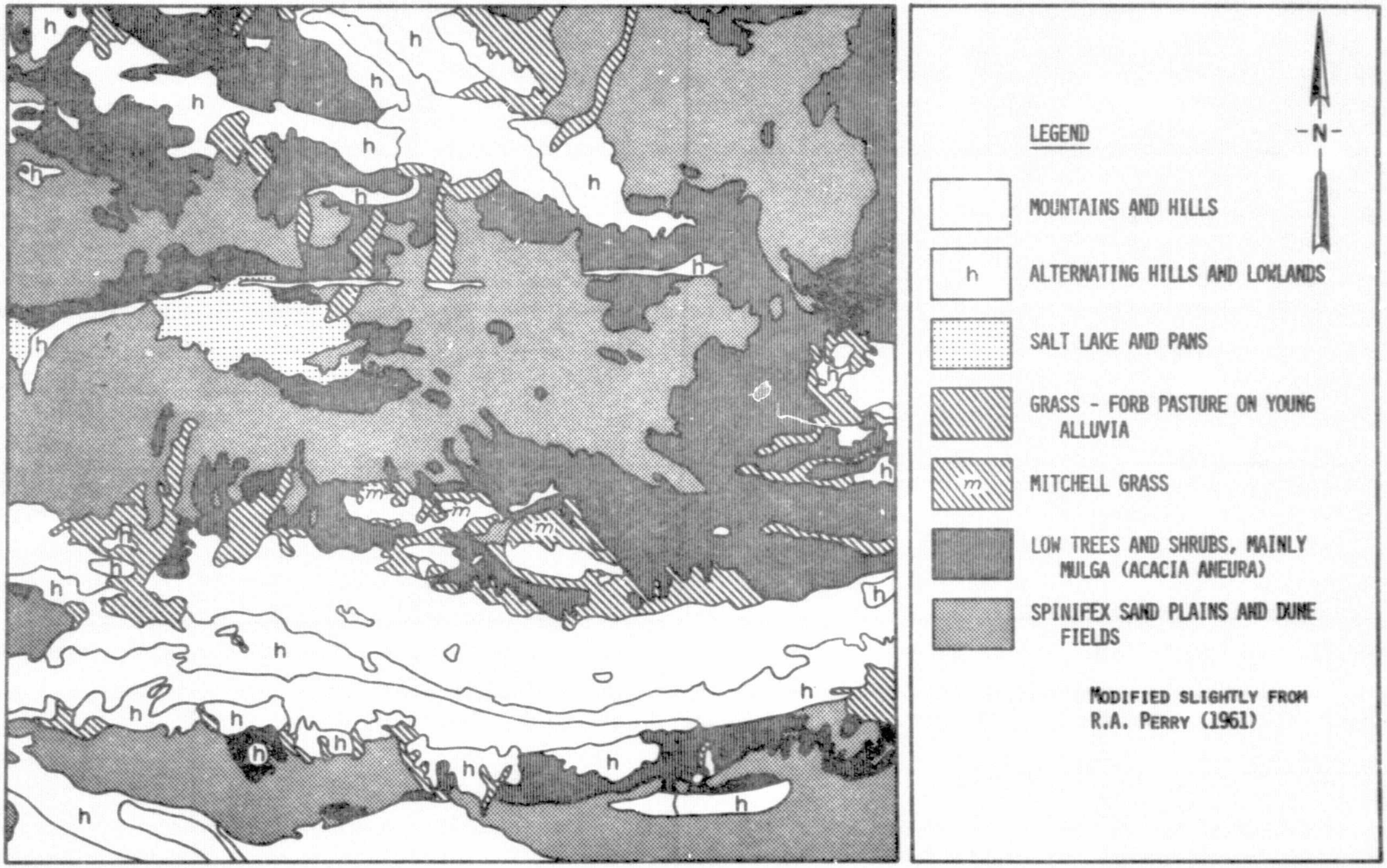


Figure 13. Landscapes northwest of Alice Springs, Central Australia. Modified slightly from R. A. Perry (1961), Pasture Land Map.

Salt Pans and Salt Lakes are largely bare of vegetation themselves (see photo 13 in Figure 21 and photo 19 in Figure 22) but are surrounded by a complex of spinifex, salt grasses, and other salt tolerant plants. Most species in this category are extremely sensitive to small changes in salt content, soil texture and drainage and for this reason distinct belts of vegetation develop around the individual pans. As with the Mountain and Hill category most of these changes occur at scales too small to map and too small to be detected on the space photograph.

Mitchell Grass country is the most restricted spatially of the grass categories recognized. The type carries mainly Mitchell grass (Astrebla pectinata) as well as the other drought-evading perennial grasses, blue bush (Chenopodium auricomum) and salt bush (Atriplex vesicarium) (see photo 4 in Figure 14). It is generally restricted to flat or gently sloping treeless plains with heavy clacareous clay soils on Tertiary or recent alluvia.

On his original map Perry recognized two types of Short Grass-Forb pasture, one type occurring on young alluvia and the other on flat or undulating country. In this report the young alluvia type is retained as a predominantly grass category, but the more extensive variety on flat or undulating country is reclassified as Low Tree and Shrub.

In alluvial areas ephemeral short grasses and forbs form the predominant ground cover with scattered low trees overhead (see photos 5 and 6 in Figure 14). Kerosene grass (Aristida browniana) is the species most commonly encountered in the footslope zone of the ranges between Hamilton Downs and Dashwood Creek. Along Gidyea, Napperby and Day Creeks northeast of Napperby Lake sparse low trees occur together with witchetty

Figure 14 (Following Page)

1. Spinifex (Triodia basedowii) with scattered wichetty bush (Acacia kempeana), blue mallee (Eucalyptus gamophylla), Hakea divaricata, and Petalostylis cassinoides, near Connors Well 100 kilometers (60 miles) north of Alice Springs; 2. Typical mulga (Acacia aneura) community with a ground cover of kerosene grass (Aristida browniana), at seventeen mile experiment site north of Alice Springs; 3. Forb-field plains with a variety of both perennial and annual chenopods, particularly Bassia spp., and composites including Brachyscome spp., and grasses mostly Aristida spp., Panicum decompositum, and Chloris scariosa, near Harry's Creek north of Alice Springs; 4. Short grass-forb pastures on Hamilton Downs station. Mount Hay is in background. Forbs increase in response to heavy grazing; 5. Kerosene grass (Aristida browniana) forb-field on young alluvium flanking the Macdonnell ranges west of Alice Springs; 6. Kerosene grass (Aristida browniana) on Napperby Creek alluvials. Low trees are of Eucalyptus suberia, Hakea divericata and iron wood (A. estrophiolata) and Atalayea hemiglauca.

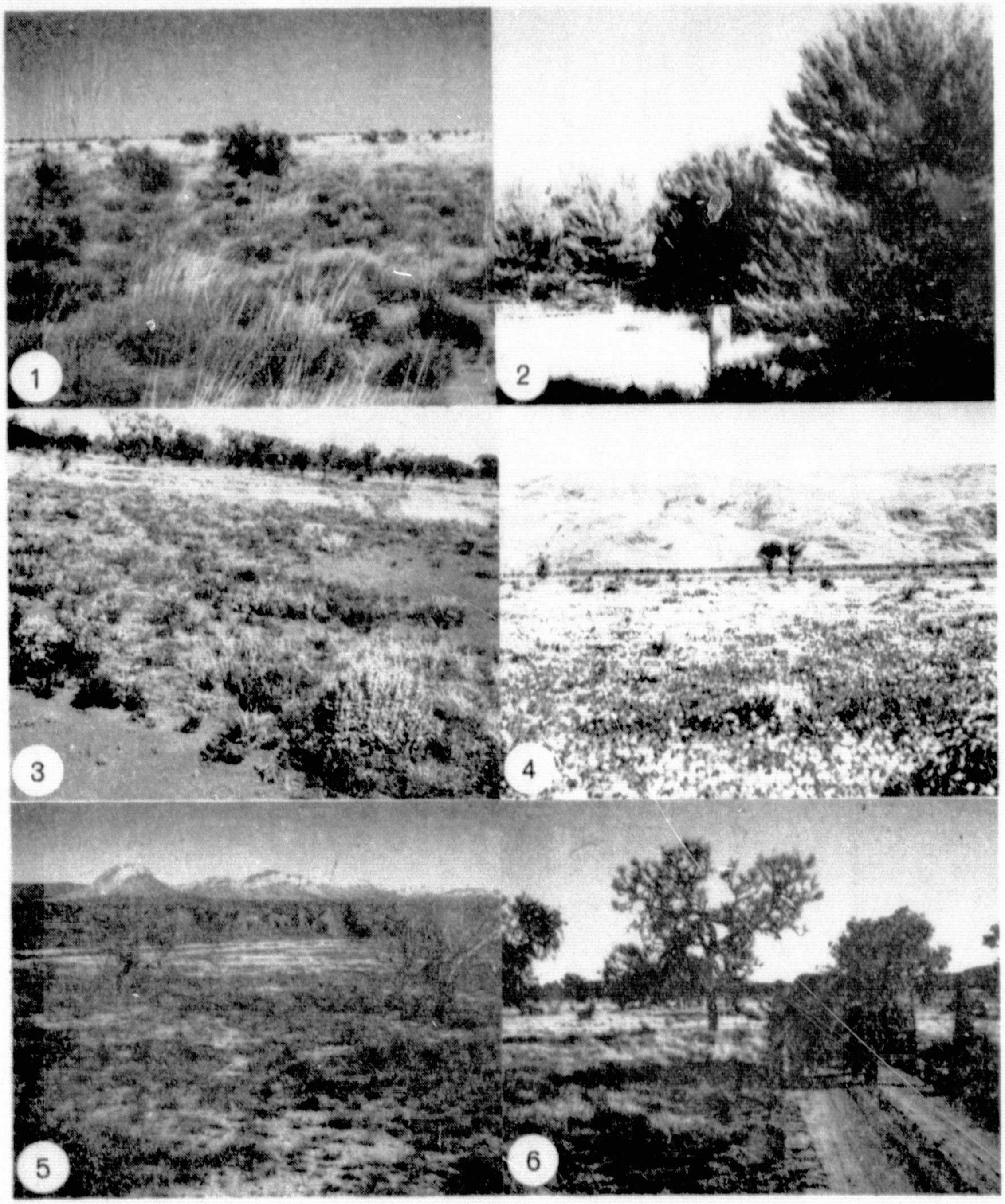


Figure 14

bush (A. Kempeana), gidgee (A. georginae), coolibah (E. microtheca), and ghost gum (E. papuana).

The Low Tree and Shrub category is overwhelmingly dominated by mulga (Acacia aneura) and is found on flat to undulating topography (see photo 2 in Figure 14) and on the flanks of mountain ranges (see photo 11 in Figure 15) on all rock types and on a wide range of soils. Associated with the mulga are: gidgee (Acacia georginae), southern ironwood (A. estrophiolata), myall (A. calicola) and witchetty bush (A. kempeana). In all of these, height, density and vigor are highly variable due mainly to the effects of drought. As a consequence an endless variety of structural subtypes exists, most of which have ill-defined boundaries.

Spinifex Sand Plains and Dune Fields occupy most of the central portion of the photo. Hard spinifex (Triodia basedowii) (see photo 1 in Figure 14) is the dominant species on flat sandy plains with smaller areas of soft spinifex (T. pungens) (see photo 8 in Figure 15) and feathertop (Plectrachne schinzii). Trees and shrubs are widely scattered except in local low spots where mulga and coolibah (E. microtheca) congregate. In dune fields hard spinifex occupies the flanks with mulga in the swales. Dune relief frequently approaches 6 meters (20 feet) or more from swale to crest with troughs 360 meters (400 yards) wide and dune flanks 130-270 meters (150-300 yards) across depending on orientation. The parallelism and linearity of these dunes give rise to alternating zones of spinifex and mulga vegetation which are easily distinguishable even from orbital altitudes.

In addition to these general boundaries a detailed vegetation map is available, from the CSIRO field staff, of Kunoth Paddock, a roughly 70 square mile area, reproduced here at a scale of 1:70,000. The location is in the grass and mulga-covered plains east of Hamilton Downs, north of the Chewings Range, and north-west of Alice Springs and Simpson Gap. (Figure 12). The vegetation map, reproduced in Figure 31, contains the following categories:

1. Riparian (Depression): River Red Gum - Curly Windmill Grass
2. Floodplains: Cotton Bush - Short Grasses and Forbs
3. Foothill Fans: Short Grasses and Forbs
4. Savannah Woodland: Short Grasses and Forbs
5. Calcareous Shrubby and *Enneapogon* Grassland
6. Mulga - Short Grasses and Forbs
7. Mulga - Perennial Grasses and Shrubs
8. Gilgai/Inter-Gilgai: Neverfail Grassland
9. Spinifex Hummock Grassland

3.2 IMAGE CHARACTER AND TRANSFORMATIONS

The basic requirement for photo interpretation is that the photo in question has differences in tone, texture, shape and size between entities. Ordinarily the photo-interpreter works with high resolutions such that texture, shape, and size convey most of the information, and differences in tone are of relatively modest importance. Air photos commonly show quite different tones for objects or aggregates of objects which we know to be the same, depending on lighting conditions and other variables; conversely, similar tones may be noted for unlike objects.

Resolutions in the several space photographs and images are roughly as follows for high and low contrast entities: Gemini 80-140 meters (262

Figure 15 (Following page)

Ground Photographs of representative vegetation types and landscapes near Alice Springs; 7 River red gum (Eucalyptus camaldulensis) along Napperby Creek. 8 Soft spinipex (Triodia pugens) with Coolibah (E. microtheca) near Rembrandt Rock southeast of Napperby Salt Lake. 9 Interbedded sedimentaries (limestone, sandstones and conglomerates) in the Macdonnell ranges. 10 Melalenca spp., swamp scrub or the Yuendumu road 11 kilometers (7 miles) southeast of Napperby Salt Lake. 11 Mulga-spinifex slopes of the Heavitree Range near Ellery Gorge in the Macdonnell Ranges. 12 Bare areas and mulga on Missionary Plain, located as site 12 in Figure 23.

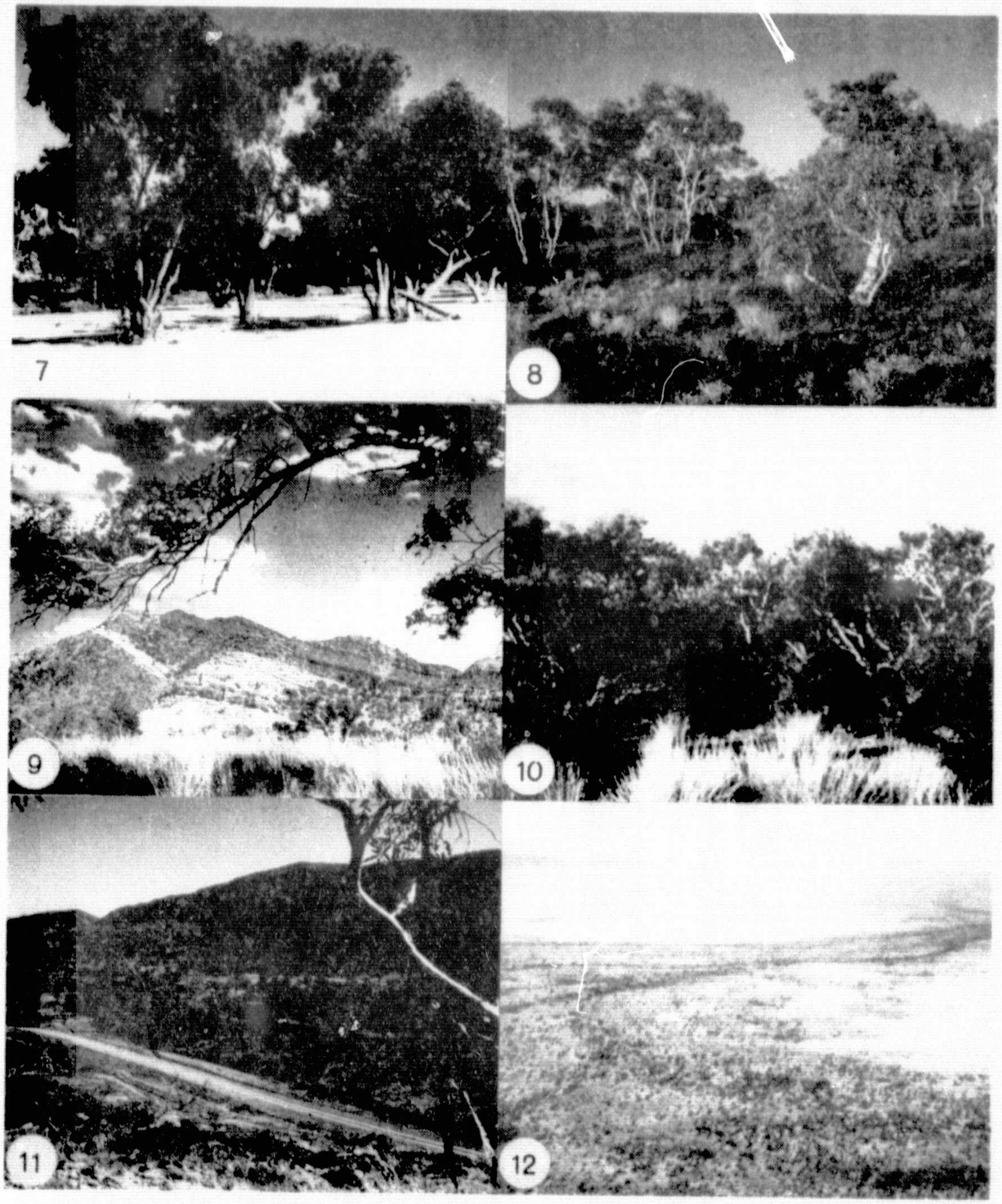


Figure 15

to 460 feet), LANDSAT 70-120 meters (230-394 feet), SKYLAB S-190A 50-75 meters (164-246 feet) and SKYLAB S-190B 13-18 meters (43-59 feet). A number of features of these resolutions should be mentioned: 1) detailed texture, shape and size clues are completely missing from all except the SKYLAB S-190B photographs; 2) Texturés, shapes and sizes at a grosser level of generalization may appear for the first time in space images; 3) tone is retained as the major clue, but because of the modest resolution many entities are mixed in resolution cells thus seriously diluting discrimination power; again the SKYLAB S190-B images are distinctly better than the others in this respect (See Figure 32, and Table 44); 4) as many as 8 discrete categories of landscape in the Alice Springs area have much the same light tone on the Gemini, LANDSAT, and S-190A space photographs, yet all are worthy of separate categorization: interpreters cannot make such separation rationally partly because the phenomena may be truly inseparable, and partly because of the haze noise in the blue and green-sensitive regions. Again, the higher spatial resolution of the S-190B is helpful; 5) as Schwarz et. al. (1969) have shown, at the 100 m. resolution level there are few environments which do not have a majority of cells containing two or more categories and there is a notable improvement at 30 m.; 6) when unknown proportions of well-, moderately-, poorly-, and totally-unknown entities are mixed in a cell, the "information" such mixtures convey is ambiguous; and 7) the grosser the resolution the more one obtains an average of the landscape which may be misleading.

3.2.1 Image Transformations

Common image transformations which may be carried out on color space imagery to enhance the quality and interpretability of the image(s) include (See Simonett and Hajic, 1976): 1) Color separations, 2) Edge enhancement, 3) Shifting the color balance with color compensation filters, 4) Density slicing in single color layers, 5) Density slicing coupled with pin-registered new color combinations, 6) Making new color combinations with multi-date, and/or multi-layer films of different spatial and spectral resolution, and 7) Digitizing multiple films, bringing to congruent spatial registration and performing a variety of operations on the images including all the former, as well as masking, and texture and spectral analysis. In addition various methods of standardizing and developing a calibrated or standardized product have been developed or refined and the last 3 years with both LANDSAT and SKYLAB imagery, particularly by Hardy and co-workers (Hardy, et al, 1975). All these techniques have been shown to be of value in improving image interpretability.

For the purposes of the present study, limited use of these techniques has been made. The principal concern has been the degree to which the unmodified, higher spatial resolution images of S-190 A and S-190 B yield incremental information over the previous levels obtained with the Gemini photography and LANDSAT images.

To give some indication of the value of a complete unpacking of the image data, color separation plates of low altitude obliques (and the Gemini space photograph) are examined in the following pages.

Most color film, including Ektachrome MS film, consists of three separate recording dyes (layers of the emulsion) each sensitive to a different region of the visible spectrum. The sensitivity (S) of each layer to a light of a given wavelength (λ) is defined as:

$$S(\lambda) = E(\lambda)^{-1}$$

where

E is the energy of monochromatic radiation of wavelength (λ) required to produce a given dye density in the individual layers when the film is developed. Figure 16 shows spectral sensitivity as a function of wavelength for each dye-type of Ektachrome MS film. As seen in this illustration each dye has a peak sensitivity at a different wavelength; thus, even though there is considerable overlap in their combined sensitivity, it is feasible to distinguish them in terms of general spectral response regions. This approximation permits us to think in terms of three colors (blue, green, red) each corresponding to a particular wavelength band.

The limits of the respective wavelength bands occur at points where the sensitivity of each given dye decreases to about 10% of its peak sensitivity. Blue is thus defined as wavelengths from 350 nanometers to 490 nm; green 490 nm to 590 nm; and red 590 nm to 690 nm. Figure 16 is a model of energy received by the film, lumped into three wavelength bands each corresponding to one of the primary colors. The actual error introduced by such lumping is small, because color photographs by their nature record only color, not the actual spectral reflectance of the original scene.

Color separation involves: 1) masking to correct for overlapping skirts of the three dye density curves (Figure 17) and 2) preparation

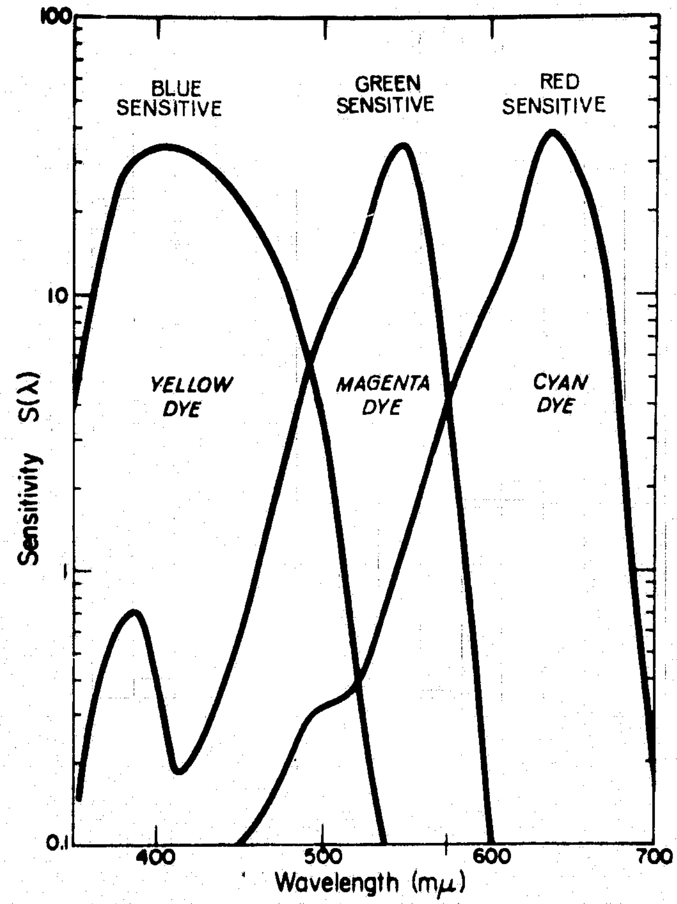


Figure 16. Spectral sensitivity of Kodak Ektachrome MS Aerographic Film (Estar Base), Type SO-151.

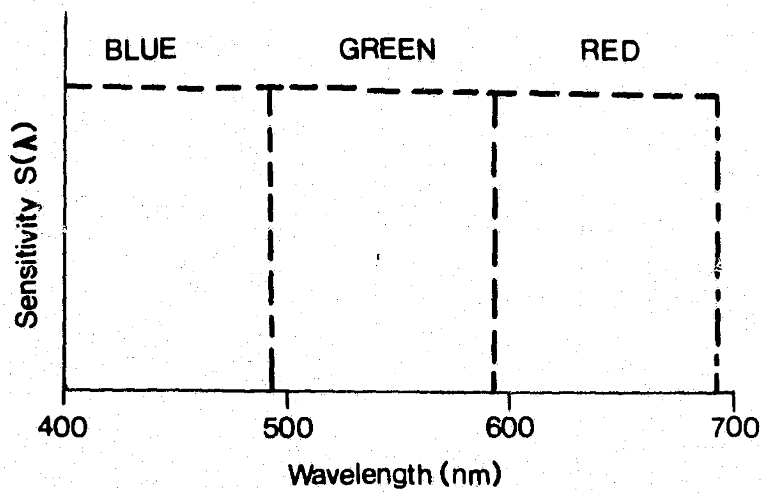


Figure 17. Crude lumped model for the blue, green and red sensitive layers of Ektachrome MS Aerographic Film after color separation and masking.

of black and white negatives from the color photo using blue, green and red filters. The particular filters used on the Gemini Alice Springs photograph, Wratten numbers 47B, 58, and 29, are illustrated in Figure 18, photos 1, 3, and 5. Masking is essential to ensure that the content of each separation plate is crudely spectrally limited. Each of the separation plates is a rough record of the amount of energy received by the camera in the corresponding wavelength band (Figure 17); thus, variations in density on any of the separation plates represent approximate relative increase or decrease in reflectance in that wavelength region and in a general way simulate the way three true multiband photographs would appear. It is important to emphasize, however, that these separation plates are not quantitative, nor are they multiband; they are approximations.

The separation plates enable some assessment to be made of some of the tonal ambiguity present in the original color photograph. This approach to the study of image content is a function of wavelength is useful in evaluating spacecraft photography, since atmospheric attenuation is a function of wavelength. Figure 19-constructed from data in Elterman (1964)-shows the theoretical trend of atmospheric attenuation versus wavelength for Rayleigh, aerosol, and ozone attenuation factors in a "clear standard atmosphere" never occurs in nature. Consequently, Figure 19 illustrates the best possible conditions ever available for spacecraft photography.

Comparison of Figures 18, photos 2, 4, and 6, with Figures 19 and 20 reveals, as expected, that attenuation is most severe in the shorter wavelengths, and this applies also in the S-190A and S-190B photographs. The blue band has very little terrain detail and is practically useless for mapping purposes. The green-sensitive layer contains considerably more

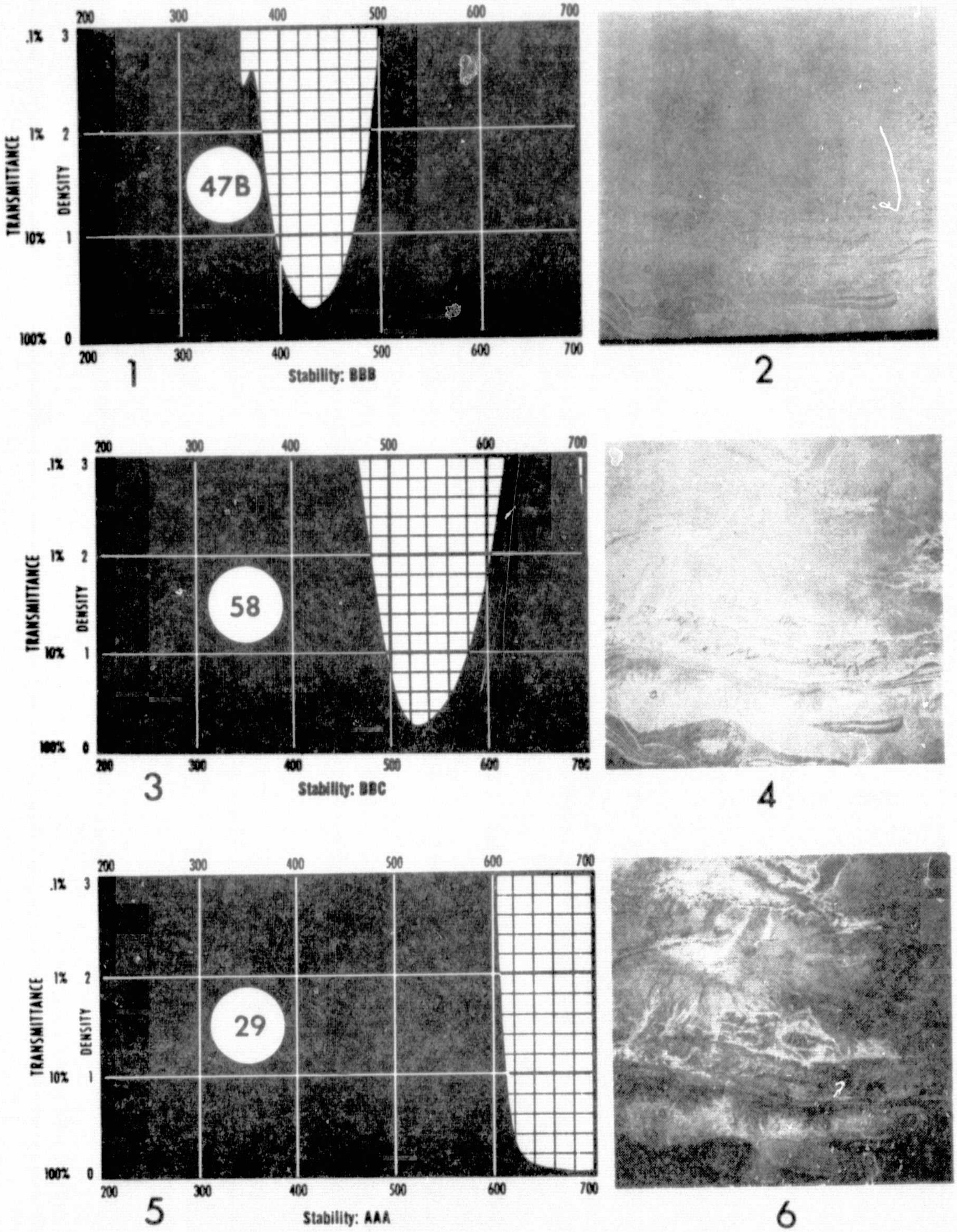


Figure 18. Reproduction of the blue, green and red separation plates of the Alice Springs Gemini space photo with the filters used for the separations.

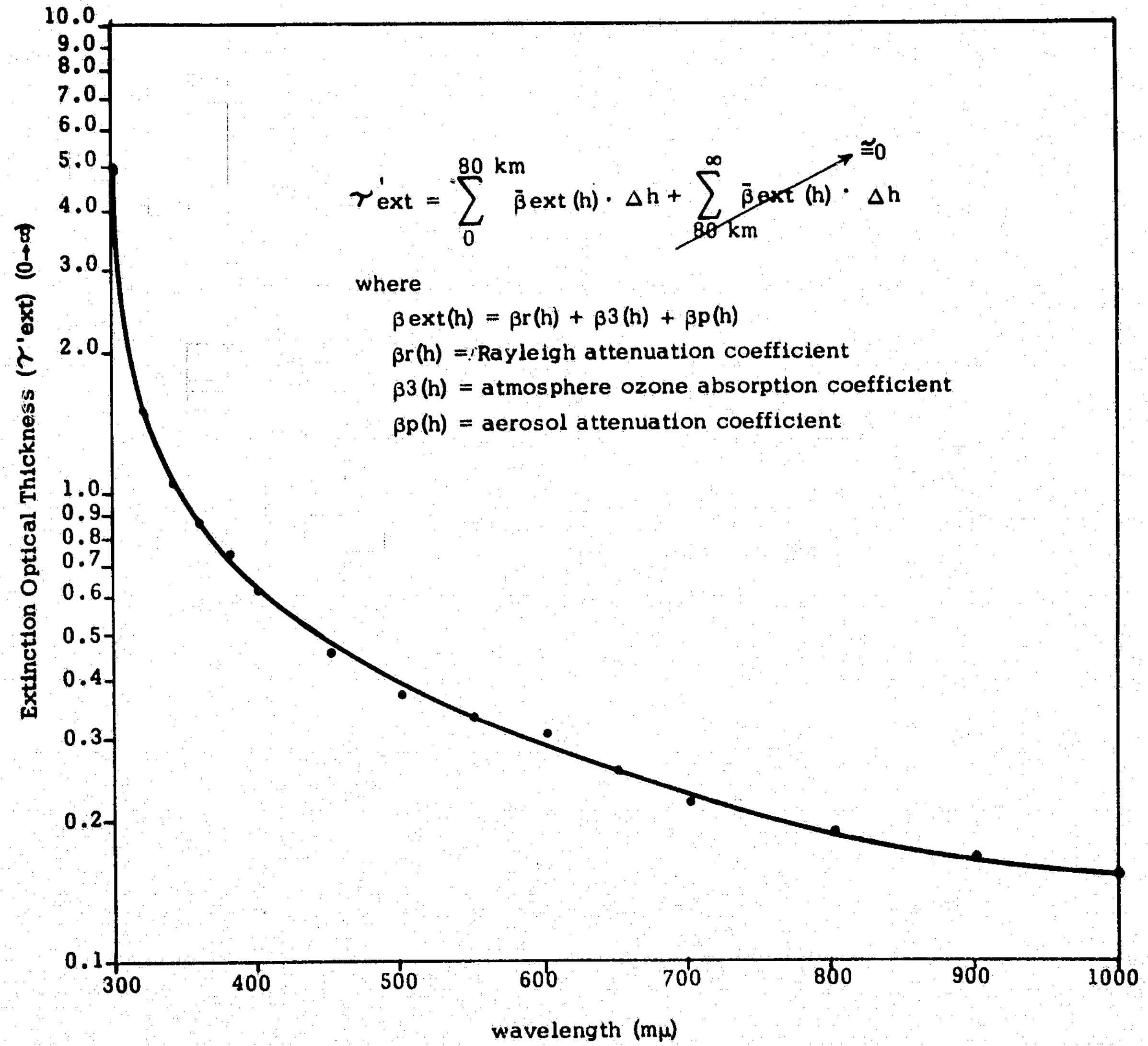


Figure 19. Theoretical trend of extinction optical thickness for a clear standard atmosphere (after Elterman 1964).

terrain detail in the form of boundaries and discrimination of areas visible on the color photograph. In effect this means that the blue band merely adds noise to the color photograph, and the same is true to some degree of the green band. The red-sensitive layer is, as expected, most contrasty with clear vegetation and soil boundaries. Figure 20 is particularly interesting in this respect. It shows the blue, green and red separation plates of aerial obliques located as marked on Figure 12. The blue separation plate shows the effect, even with short passage through the atmosphere, both of inherent low contrast (few blues occur in arid regions) only whites have high reflectance in the blue region, and contrast reduction from scattering, and consequent weak boundary discrimination. The improved level of boundary delineation possible with the green and red plates is consistent with the amount of detail recorded on the space photograph, indicating that this procedure does give a reliable guide to where information lies in the latter.

This brief account of the color separation process is given, because color separation plates were employed in the analysis of the various images, and the LANDSAT data was initially obtained as black and white multi-channel data. Color combinations of LANDSAT data as well as color positive transparency enlargements from Gemini color photography, S-190A color infrared and S-190B color photography were prepared at a scale of 1:250,000 for use in the interpretation. Samples of the various image forms are given in this paper, rather than completely ringing the changes on the various images available.

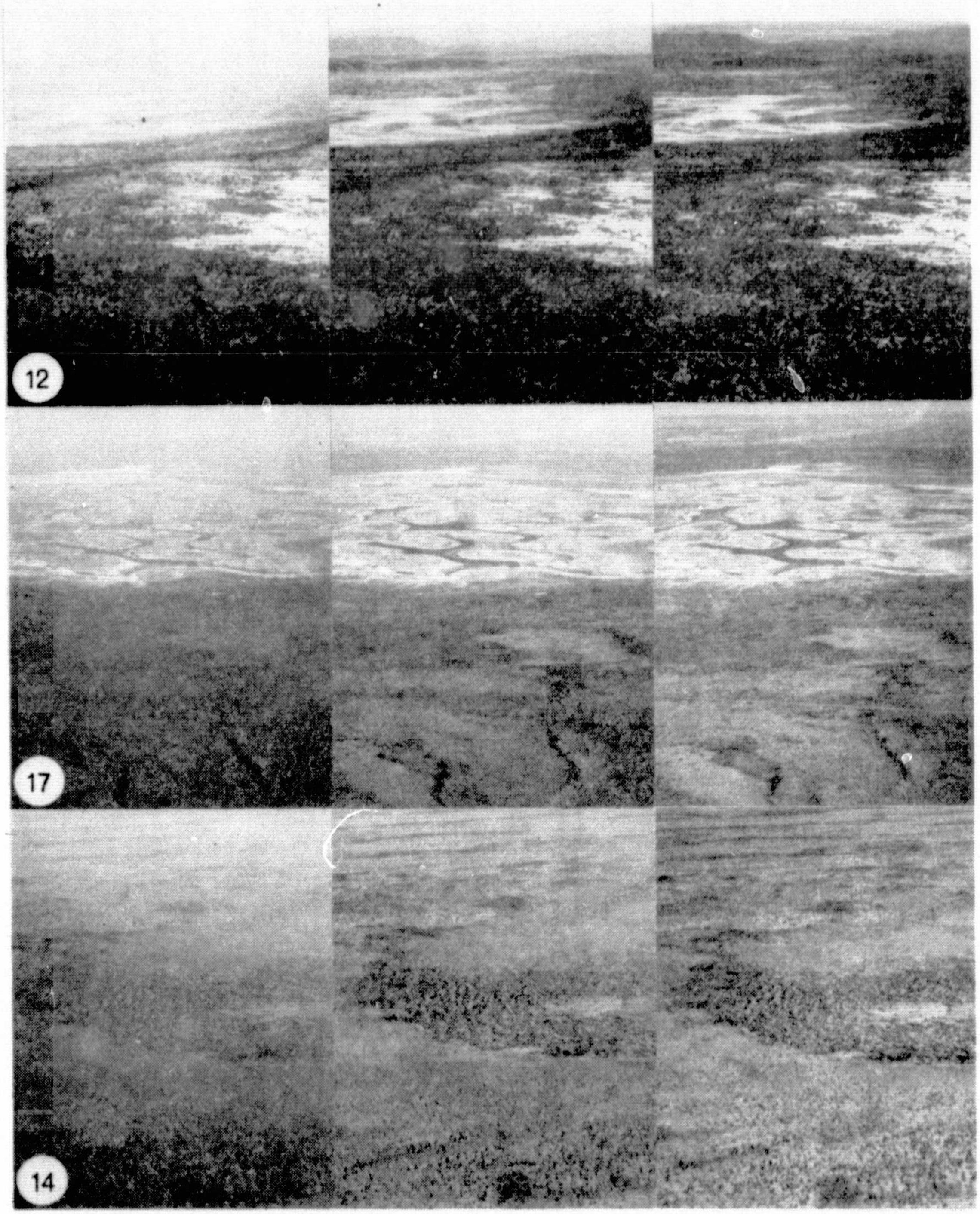


Figure 20. Blue, green and red separation plates of three aerial color oblique photographs. The location of these areas is shown on Figure 23, sites 14, 17 and 12, respectively located west and north of Napperby Salt Lake, and southwest of Alice Springs on Missionary Plain.

3.3 BOUNDARY DELINEATION AND VERIFICATION

3.3.1 Boundary Delineation

The delineation of boundaries and categorization of areas presented here are based on interpretation of: 1) 1:250,000, and 1:70,000 enlargements of the SKYLAB photographs and separation plates, 2) 1:1,000,000 and 1:250,000 enlargements of LANDSAT B & W & color images, and 3) 6 X enlargements of both the original Gemini color photograph of the Alice Springs area and its red and green separation plates. The initial interpretations consisted in each case of tracing all boundaries observable on the unaltered color enlargement. Three types of boundaries were mapped; those representing obvious, sharp, color differences separating grossly dissimilar entities, those representing less obvious but nevertheless distinct differences in color and density; and those differences in tone and density regarded as dubious to conjectural. The same procedure was applied to the red and green separation plates.

Following boundary delineation, the three resulting maps (original color photo, red separation, and green separation) were compared qualitatively by superposition. They were found to display remarkable similarity in their total boundary content although some differences were observed.

1. An approximately equal number of first category boundaries were drawn on both color photograph and red and green separation plates and these were strongly coincidental as to location for all image formats.
2. The second category of boundaries, those defined by moderate contrast ratios across adjacent entities, demonstrated less agreement of the color photo and red and green separations. Grassland boundaries seem to be easier to detect on the color photo and green separation, whereas the darker tone of wooded areas are better defined on the red separation.

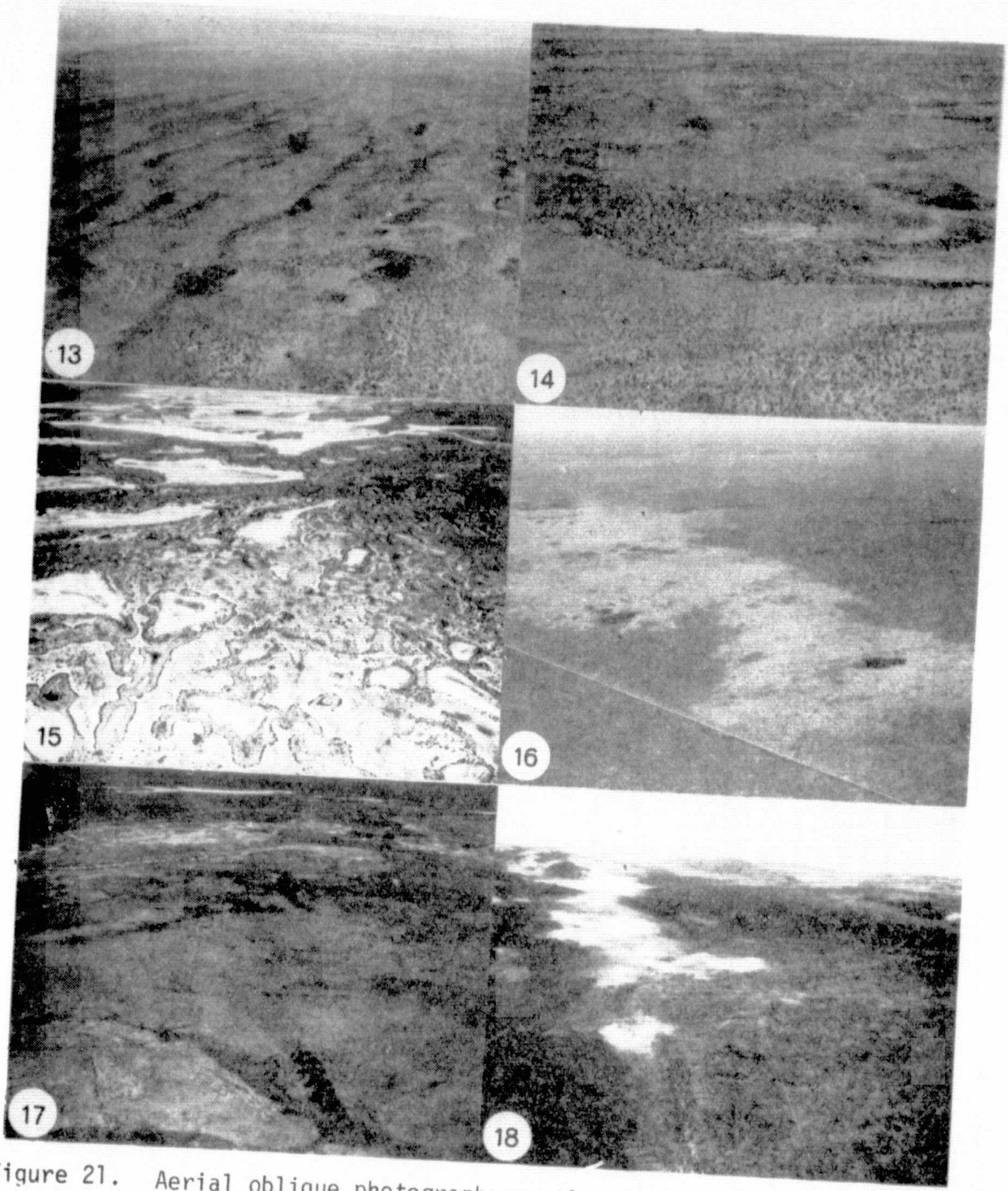


Figure 21. Aerial oblique photographs mostly near Napperby Salt Lake. The numbers are keyed for location to those given in Figure 23: 13 dunes west of Napperby Salt Lake. 14 Mulga and dunes west of Napperby Salt Lake. 15 Salt pans in troughs between irregular dunes west of Napperby Salt Lake. 16 Looking northeast from Aileron homestead. 17 Confluence of Napperby Creek and Napperby Salt Lake. 18 Looking E.S.E. from Mount Chapple to Redbank Hill.

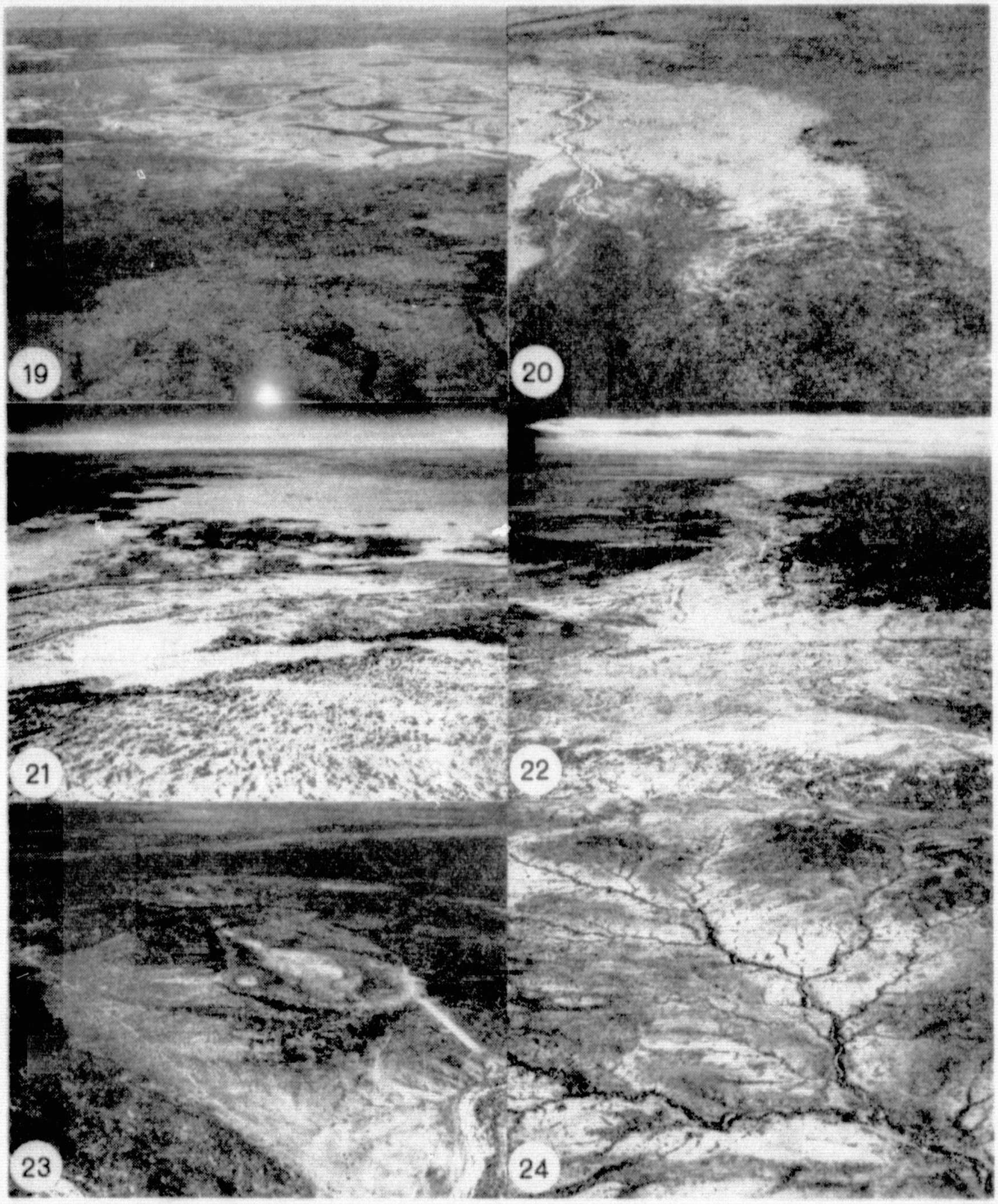


Figure 22. Aerial photographs mostly near Gidyea, Napperby and Day Creeks. The numbers are keyed for location to those given in Figure 23: 19 Looking south to Napperby Salt Lake. 20 Looking N.N.E. along Gidyea Creek. 21 Woodford River in mid distance looking east near Ti-Tree. 22 Looking south along Day Creek. 23 Napperby Station, airfield and Napperby Creek in foreground, Day Creek in distance. 24 Headwaters of Day Creek in area of dissected lateritic residuals.

3. Numerous differences in both the number and placing of third category boundaries occurred. The same observation was made by Story, Yapp, and Dunn. (1976). Since these boundaries are defined by low contrast ratios between adjacent entities, a much higher degree of subjectivity is involved in their mapping. The exact placement of any particular boundary on a separation plate is bound to shift slightly from its placement on the original color photo, especially when dealing in minor changes in entity characteristics. A subtle qualitative change between landscape types is rendered ambiguous on a color photograph because of complex interactions of atmospheric attenuation factors and the gradual change in the spectral reflective properties of the two entities involved. These influences combine to produce a low contrast ratio between the entities. When a separation plate is produced some ambiguity due to attenuation and to different reflectances in each layer is filtered out. More importantly, because the cutoff values for the information contained in the particular spectral region are relatively sharp, minor shifts in boundary location take place.

It is not surprising, therefore, that even experienced interpreters confronted with two presentations of fundamentally identical data arrive at different conclusions regarding the discrimination of subtle landscape changes. In part this also will arise from different ways of lumping and splitting. Some observers are born lumpers; other are born splitters; yet others are fuzzy-minded academics with no consistency at all. The same problems of lumping and splitting apply to all qualitative judgments by men.

It even applies to maps such as those prepared by Perry (1961), themselves substantially based on aerial photographs, which we have used as "Ground Truth" to compare with the space photograph.

3.3.2 Boundary Verification

One of the primary aims of this report is to demonstrate relationships between boundaries discernable on space photography and terrain features, and through this to gain insight into the meaning of such boundaries. The oblique photographs in Figures 21 and 22 are black and white reproductions of color photos and illustrate a range of terrain conditions through which we can begin to appreciate the nature of entities encountered and their spatial distribution.

The location of each of the obliques is plotted on the Gemini Alice Springs photograph in Figure 23. By comparing the obliques with the corresponding area on the various space images it is possible to make point-by-point comparisons of their efficiency in aiding boundary detection and delineation of "real entities". Detailed comparisons are feasible in Figures 24 and 25 which show for four regions - the locations of which are noted on Figure 12 - reproductions of an air photo mosaic based on pan minus blue 1:48,000 scale photos and the 1965 Gemini photo brought to a common scale of 1:500,000. Even more detailed comparisons are possible through comparing 1:250,000 enlargements of S-190 A and S-190 B, LANDSAT and GEMINI color images as seen in Figures 33, 34 and 35. The most detailed evaluations may be achieved through comparisons of the vegetation map of Kunoth Paddock (at a scale of 1:70,000), and a 1:70,000 enlargement of the S-190 B color photograph with low altitude vertical photographs obtained by Australian co-operating scientists (Figures 31, 32 and 36 respectively).

Figure 23. Locations of low altitude aerial oblique photographs. Numbers correspond to photographs illustrated in Figures 21 and 22.

In order to make these comparisons compact, they are collected into Table 11 which should be examined carefully in conjunction with Figures 21, 22, 23, 24, 25, 31, 32, 33, 34, 35 and 36. In Table 44 is given in turn the location of the obliques, the terrain types portrayed, the distinctness of the boundaries as seen on the color obliques and the detectability of the boundaries on the several space images. In addition the detailed examination of LANDSAT images given in Story, Yapp, and Dunn (1976) should be borne in mind.

A full comparison of each item would be wearisome. Summarizing all these checks and comparisons, we conclude that:

1. As stated in the earlier study by Simonett et al (1969), even minor juxtaposed point to point changes in tone on the Gemini photo are meaningful. However, it is not possible to decipher their meaning without detailed field work. It is encouraging to realize the very modest changes in plant communities which may be detected. Thus, quite subtle differences between crests and swales of dunes mantled mainly with spinifex are detected because of their linearity. With improved spatial resolution some ambiguities concerning the various categories are removed.
2. High contrast juxtaposed point to point changes signal that different entities are being sampled. If each entity is regarded as having its own three-dimensional probability density function for each resolution cell (the three dimensions arise from the color bands in the lumped model of Figure 17) then changes above a certain degree unambiguously indicate the presence of these entities.

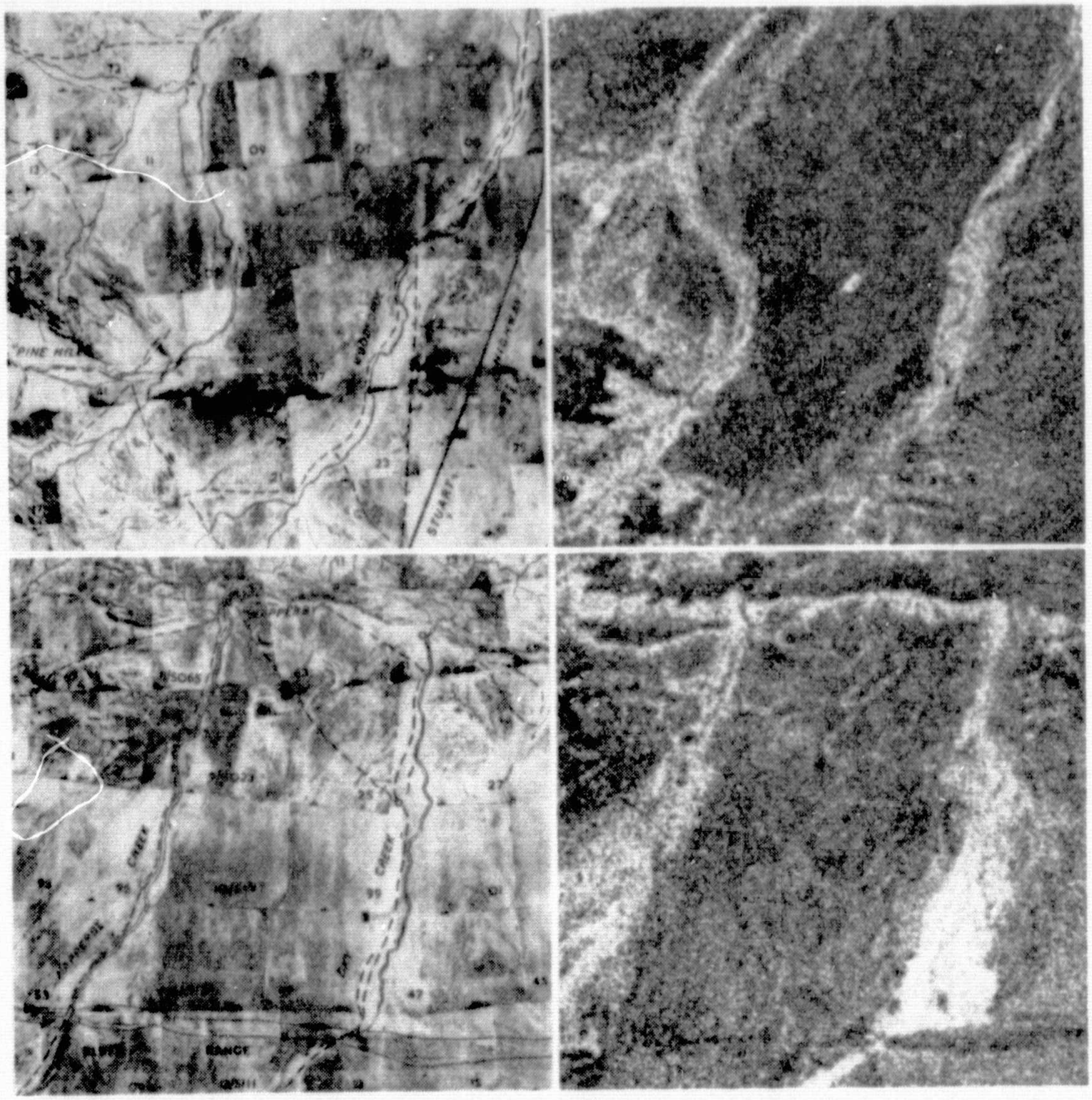


Figure 24. Comparison of air photo mosaics with red separation plate enlargements made from 1965 Gemini photography. Top, Woodford Creek area; bottom, Napperby and Day Creeks. Scale of reproduction 1:500,000. The center of each area is indicated on Figure 12 with the number 23 and letter T or B for Top or Bottom.

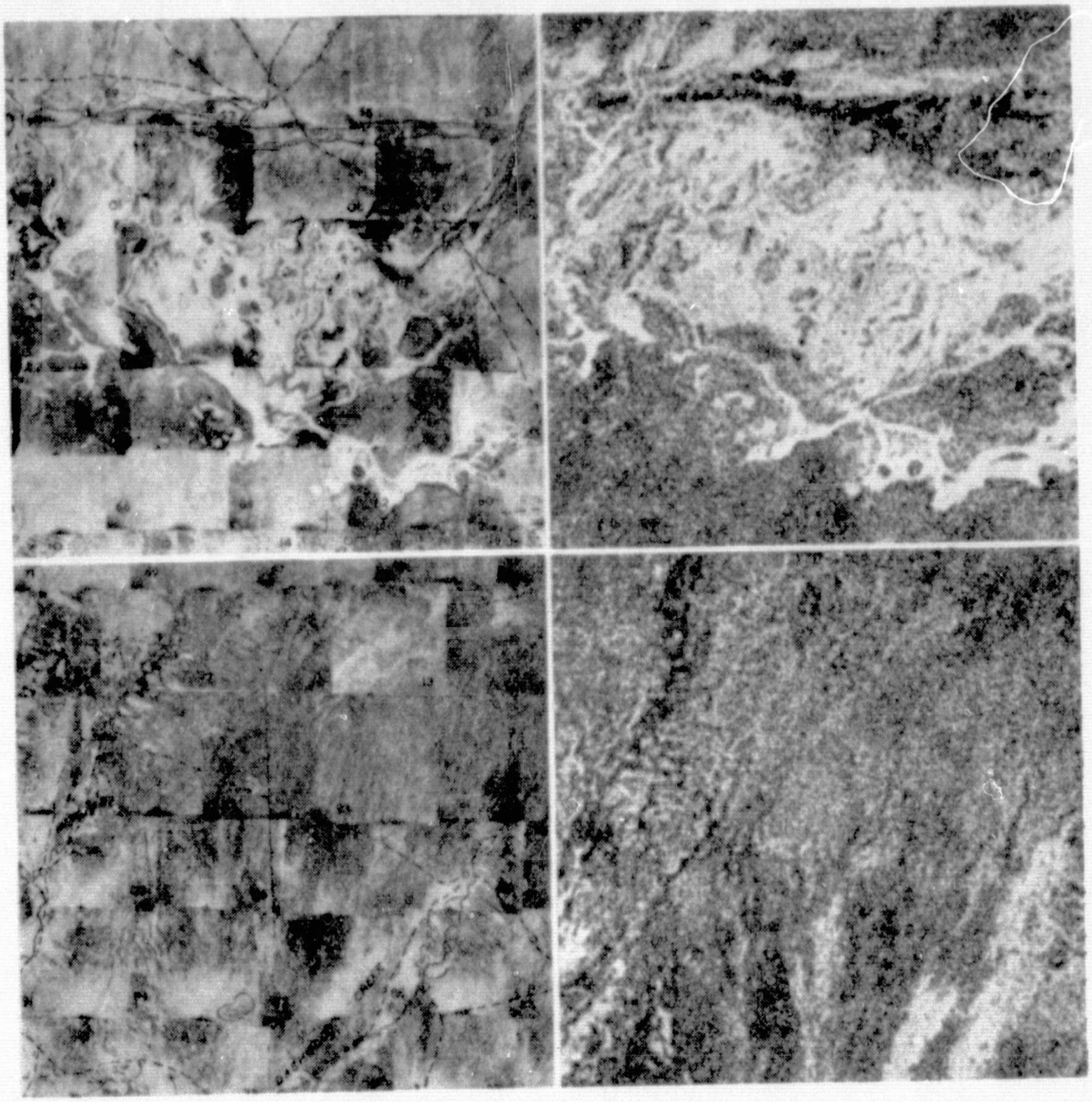


Figure 25. Comparison of air photo mosaics with red separation plate enlargements made from 1965 Gemini photography. Top, Napperby Lake; bottom, Dashwood Creek southwest of Napperby Lake. Scale of reproduction 1:500,000. The center of each area is indicated on Figure 12 with the number 24 and the letter T or B for Top or Bottom.

In short, marked changes in tone are never noise even at the resolution cell level and this applies equally with the Gemini, LANDSAT, and SKYLAB images. This is very well evidenced in the comparison between dark, light and mid grey tones on both LANDSAT Band 5 (Red) and the Gemini red separation plate near Napperby Salt Lake. Dark points are mulga or other dense clumped vegetation, light are always salt pans, and mid tones are spinifex sand plains; see for example Figures 25 and 33 where this is readily confirmed.

3. The space images enable many quite transitional or fuzzy boundaries to be integrated and detected readily in comparison to using air photos of different acquisition dates, times and hence, sun angles. Figures 24, 33, and 35 show this well in the Napperby and Day Creek area comparisons.

4. In Simonett et al (1969) it was pointed out that "In order reasonably to capture the environmental variability of this region a resolution with a 1.6:1 contrast ratio of 50 feet would be essential, though 100 feet would be acceptable. To obtain such resolution would require a system with an average of 30 line pair/mm resolution on a low contrast target and a focal length of no less than 12 inches and preferably 24 inches (Doyle, 1967). The scale of significant variation in this environment cannot be captured with a 460 foot resolution (for low-contrast features) as in the Gemini photo."

The following additional conclusions and the analysis by Story, Yapp, and Dunn (1976) strongly support the original analysis.

TABLE 44

COMPARISON OF DETECTABILITY OF LANDSCAPE BOUNDARIES ON SPACE IMAGES

PHOTO OBLIQUE AND AREA	TERRAIN TYPES PORTRAYED ON B & W OBLIQUE PHOTOS	DISTINCTNESS OF BOUNDARIES AS SEEN ON OBLIQUE PHOTOS	DETECTABILITY: (GEMINI PHOTO)	LANDSAT IMAGE COLOR AND B & W	SKYLAB S-190A COLOR	SKYLAB S-190B COLOR
Figure 21: 13 and 14. Dunes West of Napperby Lake	a) dense mulga groves (dark grey)	a) very distinct	a) Excellent; depends on length-width of grove	a) Barely detectable; depends on length-width of grove	a) Excellent	a) Outstanding detail within mulga groves; grove pattern clear
	b) spinifex on dune flanks (med. grey)	b) ill-defined	b) Not discriminable except by deductive association	b) Not detected	b) Clearly visible on Color IR image	b) Clearly visible as different community but some deduction required
	c) mulga savannah (speckled)	c) moderate to ill-defined	c) Detail lost by generalization; interpreted as spinifex sand plain (Figure 29)	c) Not detected; interpreted as spinifex sand plain	c) Readily seen	c) Easily mapped, small groves visible
Figure 21: 15 Napperby Lake	a) salt pan (white)	a) very distinct	a) Excellent for large entities; ambiguous when next to spinifex	a) Very good for large entities, especially for salt, water, bare soil	a) Excellent; especially good with salt/water/bare soil boundaries	a) Excellent: S-190 A plus S-190 B give much detail and differentiation
	b) spinifex dunes (mottled grey-light grey)	b) complex but distinct pattern	b) Detail lost by generalization; interpreted as spinifex sand plain (Figure 29)	b) Not detectable, mapped as spinifex sand plain	b) Observable, but not distinct	b) Easily discriminable
	c) short tree & shrub (mulga) (dark grey)	c) distinct to ill-defined belts	c) Not detectable	c) Not detectable, confused with spinifex sand plain	c) Observable (barely)	c) Easily discriminable
Figure 21: 16 Near Aileron Station	a) spinifex sand plain (medium grey)	a) distinct to gradational	a) Not detectable	a) Not detectable	HIDDEN BY CLOUDS Roads readily visible	HIDDEN BY CLOUDS, but roads, tracks, paddock boundaries in other locations easily observed. (See Figure 32)
	b) short grass-forb (light grey)	b) very distinct	b) Not detectable (< system resolution)	b) Not detectable		
	c) Stuart highway (white line)	c) very distinct	c) Very poor (most groves too small)	c) Not detectable		
	d) dense mulga groves (shadows) (dark grey)					

TABLE 44 (Continued)

PHOTO OBLIQUE AND AREA	TERRAIN TYPES PORTRAYED ON B & W OBLIQUE PHOTOS	DISTINCTNESS OF BOUNDARIES AS SEEN ON OBLIQUE PHOTOS	DETECTABILITY: (GEMINI PHOTO)	LANDSAT IMAGE COLOR AND B & W	SKYLAB	
					S-190A COLOR	S-190B COLOR
Figure 21: 17 Confluence Napperby Creek and Napperby Salt Lake	a) dense mulga (dark grey)	a) very distinct	a) Poor/ambiguous; entities merely inferred.	a) Not detectable as significant change	a) Good boundary delineation, but lacks the detail of S-190 B, needed for entity discrimination	a) Extraordinary detail visible, boundaries and entities clearly separable (see Fig 32)
	b) salt pan (some white/ some with water)	b) distinct/very distinct	b) Detectable as gross features; detail lost	b) Barely detectable	b) Detectable, but some boundaries difficult	b) Detail readily visible, but some units ambiguous
	c) short grass & forb w/scattered trees (light grey)	c) very distinct/ gradational	c) Good/excellent	c) Good detection	c) Good to excellent	c) Excellent detection
	d) redgum stringer	d) very distinct	d) Moderate; seen as pale line	d) Readily detectable (see Fig 32)	d) Detectable	d) Easily discriminable
	e) spinifex sand plain	e) very distinct/ gradational	e) Good/excellent when adjacent to short grass	e) Detectable, but indistinct	e) Good to excellent	e) Excellent when adjacent to short grass
Figure 21: 18 Mt. Chapple - Red bank Hill	a) hills and mountains (dark grey-in shadow)	a) very distinct	a) Excellent	a) Good to poor	a) Excellent	a) Excellent. Considerable detail within mountains on lithologic and structural differences.
	b) mulga clumps & savannah	b) distinct to ill-defined	b) Not detectable individually; interpreted as spinifex-mulga transition in Figure 27	b) Not detectable See Figure 32	b) Barely detectable	b) Individual clump and groves of mulga seen in detail. (See Fig. 32)

ORIGINAL PAGE IS OF POOR QUALITY

TABLE 44 (Continued)

PHOTO OBLIQUE AND AREA	TERRAIN TYPES PORTRAYED ON B & W OBLIQUE PHOTOS	DISTINCTNESS OF BOUNDARIES AS SEEN ON OBLIQUE PHOTOS	DETECTABILITY: (GEMINI PHOTO)	LANDSAT IMAGE COLOR AND B & W	SKYLAB	
					S-190A COLOR	S-190B COLOR
Figure 22: 19 Napperby Lake	a) spinifex islands (white/med grey)	a) very distinct when water present, less so when not	a) Fair/good for largest islands; detail apparent but not coherent	a) Fair/poor for largest islands, little detail observed, good detection of water	a) Excellent separation of various lake-edge areas	Very great detail visible - about as much as in foreground of low altitude aircraft oblique. Roads lake, plant community boundaries clearly visible (See Figure 33) Comment applies equally to a, b, c, d, e.
	b) irregular spinifex dunes (?) (med-dark grey)	b) distinct	b) Boundary with lake distinct, others not detectable	b) Boundary with lake fuzzy	b,c,d) No clear identification of plant communities, but boundaries clearer than on Gemini photograph	
	c) spinifex w/scattered low tree (med. grey)	c) distinct/gradational	c) Not detectable except as continuation of (b)	c) Not detectable		
	d) mulga clump? redgum stringer? (dark grey)	d) distinct	d) Not detectable } system resolution inadequate	d) Just detectable		
	e) Yuendumu road (med. grey line)	e) fairly distinct	e) Not detectable }	e) Not detectable (See Fig 33)	e) Road barely visible	
Figure 22: 20 Gidyea Creek	a) spinifex sand plain (med.-light grey)	a) very distinct	a) Excellent; high contrast with (b)	a) Good; moderate contrast (with b)	a) Excellent, but no internal details	Stratiform clouds Partially obscure area; otherwise a) Excellent internal details visible
	b) short grass-forb w/scattered low trees (light grey)	b) very distinct	b) Excellent; high contrast with (c)	b) Good; moderate contrast (with c)	b) Excellent, but no internal details	b) Excellent internal details visible
	c) mulga scrub (dark-very dark grey)	c) distinct/gradational	c) Good/excellent; boundary with (a) somewhat diffuse	c) Fair only, boundwith a not detectable	c) Excellent, but no internal details	c) Excellent internal details visible
	d) red gum stringer	d) very distinct	d) Poor/not detectable; < system resolution	d) Poor not detectable	d) Detectable, but not clear	d) Excellent internal details visible (See Figure 35)

TABLE 44 (Continued)

PHOTO OBLIQUE AND AREA	TERRAIN TYPES PORTRAYED ON B & W OBLIQUE PHOTOS	DISTINCTNESS OF BOUNDARIES AS SEEN ON OBLIQUE PHOTOS	DETECTABILITY: (GEMINI PHOTO)	LANDSAT IMAGE COLOR AND B & W	SKYLAB S-190A COLOR	SKYLAB S-190B COLOR
Figure 22: 21 Woodford River	a) mulga scrub (cont. dark grey)	a) very distinct	a) Good/excellent diffuse boundaries in some locations	a) Poor to fair only	NOT COVERED IN SKYLAB PHOTOGRAPHS	
	b) dune field (as in Figure 21 photo 12)	b) very distinct	b) Excellent detection of field boundary; dune detail lost; interpreted as mulga scrub in Figure 27	b) Dune detail not observed.		
	c) spinifex sand plain (med. light grey)	c) very distinct	c) Excellent adjacent to (a); merges imperceptably to mulga savannah	c) Good adjacent (a)		
	d) short grass/scattered tree (light grey)	d) very distinct	d) Excellent; gradational to mulga savannah	d) Fair		
Figure 22: 22 Day Creek	a) mulga scrub (cont. dark grey)	a) very distinct/ distinct	a) Excellent; high contrast with	a) Fair to good; high contrast.	PARTIALLY CLOUD COVERED SEE FIGURE 25 a) Good to Excellent	a,b,c) Great detail visible <u>Within</u> categories at threshold of detectivity in S-190A.
	b) short grass w/ scattered trees (med. grey)	b) distinct-diffuse	b) Excellent; diffuse boundary with spinifex	b) Fair; diffuse boundary	b) Excellent	
	c) kerosene grass w/ scattered trees (med. grey)	c) diffuse	c) Excellent; entity ambiguous with mulga scrub	c) Fair to good; ambiguous with mulga scrub	c) Excellent	

TABLE 44 (Continued)

PHOTO OBLIQUE AND AREA	TERRAIN TYPES PORTRAYED ON B & W OBLIQUE PHOTOS	DISTINCTNESS OF BOUNDARIES AS SEEN ON OBLIQUE PHOTOS	DETECTABILITY: (GEMINI PHOTO)	LANDSAT IMAGE COLOR AND B & W	SKYLAB S-190A COLOR	SKYLAB S-190B COLOR
Figure 22: 23 Napperby Station	a) mulga scrub (dark grey)	a) very distinct/ distinct	a) Excellent; intricate detail	a) Fair	a) Excellent	Under magnification to 1:30,000 Can see details equivalent to foreground view of low altitude aircraft oblique, including Red gum stringers on both sides of Napperby Creek, and eroded soils in foreground See Fig. 25
	b) mulga savannah/ short grass (speckled-med-light grey)	b) very distinct	b) Excellent	b) Clearly visible but no detail	b) Excellent	
	c) red gum stringer	c) distinct	c) Fair; ambiguous pale grey	c) Just detectable	c) Just visible	
	d) hills	d) distinct	d) Fair/poor; ambiguous	d) Detectable	d) Clearly detectable	
Figure 22: 24 Headwaters Day Creek	a) low tree/shrub-spinifex (light to medium grey)	a) gradational	a) Not detectable; detail lost generalization	a) Not detectable	a) Differences visible but uncertainties still present	a) Clearly visible, easily mapped.
	b) red gum stringers (very dark grey)	b) very distinct	b) Very poor; present but incoherent	b) Not detectable	b) Just detectable	b) Very readily mapped See Figure 25

5. The SKYLAB S-190 A images, with a spatial resolution of 75 meters for low contrast features (320 feet) is a definite improvement over the Gemini image, but nevertheless, the absence of any clear geometric identity to shapes makes identification of objects difficult. The threshold of resolution for significant entity discrimination is still not reached with S-190 A.

6. The SKYLAB S-190 B images with a spatial resolution for low contrast features of 18 meters (59 feet) lies at about the threshold for clear entity discrimination, and of understanding the basis for boundary differences. In addition fine supporting detail such as roads, tracks, location of boundary fences (shown by differential grazing density effects on pastures) is clearly visible. All of these are critical as location identifiers. Their identification in the S-190 B photograph reduces the need for additional stages in a multi-stage sample design.

Variations Between Photo-Interpreters in Relation to Image Spatial Resolution

From the nature and degree of boundary differences encountered in delineating boundaries, a number of questions arose concerning, first, the ability of experienced interpreters to accurately map unfamiliar environments and, second, the comparability of their efforts as a function of spatial resolution of the systems and image contrast.

VARIOUS INTERPRETATIONS OF BOUNDARY CONDITIONS ON SPACE PHOTOGRAPHY

Alice Springs, Australia

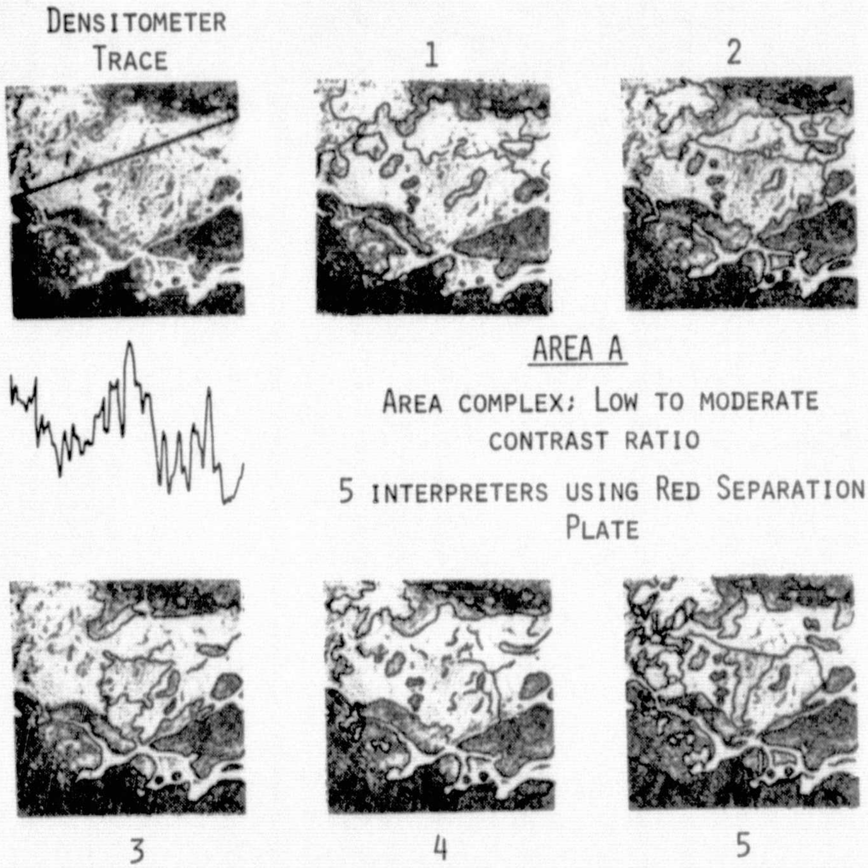


Figure 26. Interpretations of boundaries by five interpreters in an area where a complex of small entities containing low to moderate contrast ratios between entities occurs.

The initial analysis was prepared by Simonett et al (1969), and a parallel procedure was used by Story, Yapp, and Dunn (1976) in their analysis of LANDSAT images. Comparison of these examples should be made with SKYLAB images in Figure 32 of Kunoth Paddock, and those in Figures 34 and 35. To gain insight into these problems a test area containing many of the vegetation types and boundary conditions was selected on the Gemini red separation plate in the 1969 study. Twelve interpreters with no first-hand knowledge of this area were asked to perform a three category boundary delineation similar to that carried out by G. R. Cochrane who prepared the master boundary delineations.* No constraints were placed upon the interpreters as to what they should be looking for; simply that they should map as consistently as possible any boundaries they detected.

The results of these efforts clearly showed the extent of variation between interpreters and helped focus attention on the general problems of line detection. Four of the twelve interpretations, representing a fair cross section of all, were selected for comparison with the original work of Cochrane. Based on these interpretations Figures 26 through 29 illustrate the degree of variation encountered in four fundamentally different boundary situations. The name of the interpreters are keyed to the illustrations as follows: 1) G. R. Cochrane, 2) S. A. Morain, 3) W. G. Brooner, 4) F. M. Henderson, 5) D. E. Egbert.

Each of the four sets of boundary conditions contains its own problems of line detection. In Figure 26 attention is directed toward a portion of Napperby Dry Lake in which a complex pattern of salt pans,

---

*Cochrane's map was revised by Morain and checked by Simonett. We felt that this procedure of serially reconciling differences was the most appropriate.

spinifex islands, and salt grass rises occurs. Comparing the five interpretations, it is clear that no two observers saw things alike, although there is fairly high congruence of boundaries in the lower portion of the area. Toward the center of the lake, however, where low contrast ratios prevail, there is virtually no comparison between interpretations. Here is a situation, according to the film density trace, in which numerous, small, moderately contrasting elements are contained in larger, area-extensive elements with lower contrast ratio. Such a condition is confusing to interpreters because the boundaries most readily detected lie at a scale too small to map; whereas those that perhaps should be mapped at a reconnaissance scale are difficult to discriminate. Similar conclusions were reached by Story, Yapp, and Dunn (1976) in their analysis of LANDSAT images. Comparison of the SKYLAB S-190 B images of the same area (Figures 33 and 35) shows that the improved spatial resolution eliminates many of these problems, though new smaller details which may be confusing are introduced.

The greatest comparability between interpreters is found wherever the phenomena being separated are extensive and contrast sharply with neighboring types. Figure 27 depicts this set of conditions in the area from Napperby to Day Creek and except for unavoidable differences in detail, all interpreters saw essentially the same pattern of boundaries with respect to the alluvial areas.

In these lower contrast areas (between the alluvial plains) individual perceptions and mapping procedures were not subject to the same "guidance" from nature as where broad areas of high-contrast entities are juxtaposed. The net result is to produce boundaries which, if seen only in map form - and not displayed against the background of the space photograph - would be accepted.

VARIOUS INTERPRETATIONS OF BOUNDARY CONDITIONS ON SPACE PHOTOGRAPHY
Alice Springs, Australia

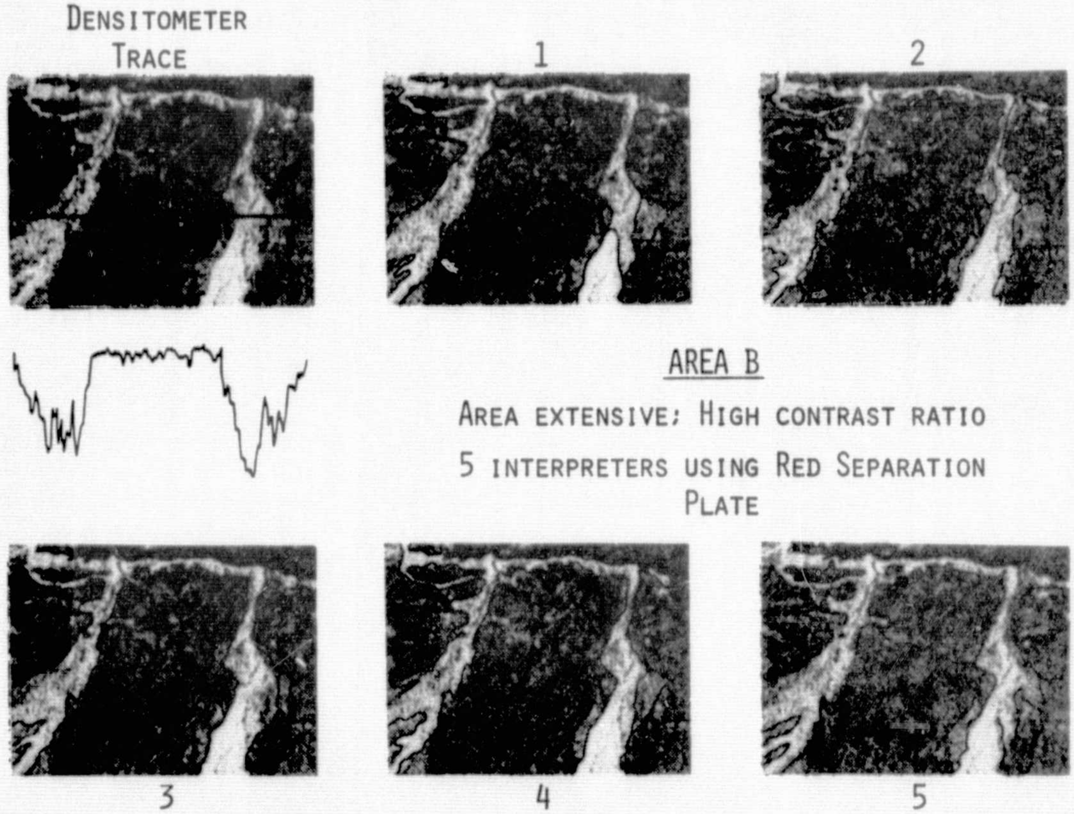


Figure 27. Interpretation of boundaries by five interpreters in an area where extensive entities are separated by high contrast boundaries.

Scientists tend to accept one anothers work in such areas! The question then arises, is this merely an artefact of the poor spatial resolution or are other factors involved.

First, to address the question of poor spatial resolution, Figure 27 should be compared with Figure 35. In the 1:250,000 enlargment of the SKYLAB S-190 B image (Figure 35) it is clear that there is justification only for separating the isolated hills near the northern (top) portion of the image as shown in Figure 27. The boundaries in the southern portion of Figure 27 are all diffuse, third-order boundaries. None appear reasonable on the SKYLAB S-190 B image. By the same token the boundaries on Figure 27 appear to be straining for separation. We conclude, therefore, that none of these third-order boundaries are real, and that the interpreters were misled by the coarse resolution into reaching for, and "identifying" non-existent boundaries. In essence then both artefacting from the poor spatial resolution and different judgments by interpreters are involved.

In a similar study with LANDSAT Story, Yapp, and Dunn (1976) found that "on the whole our correlations (with Perry et al's (1962) previous survey) are poor . . . on the face of it, the best we could expect would be reliable mapping over less than half the total survey area".

They also found similar differences in boundary delineation and entity detection to those stated above in areas of slight contrast.

VARIOUS INTERPRETATIONS OF BOUNDARY CONDITIONS ON SPACE PHOTOGRAPHY

Alice Springs, Australia

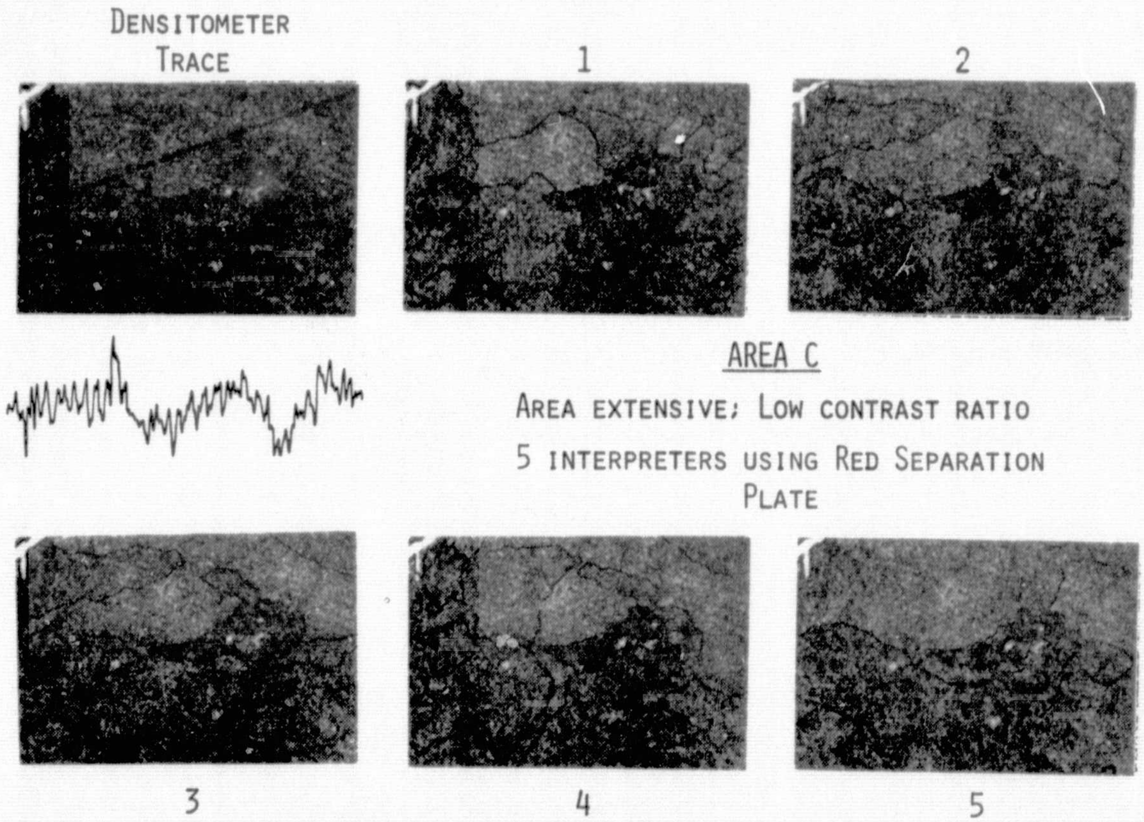


Figure 28. Interpretation of boundaries by five interpreters in an area where extensive entities are separated by low contrast ratios.

VARIOUS INTERPRETATIONS OF BOUNDARY CONDITIONS ON SPACE PHOTOGRAPHY

Alice Springs, Australia

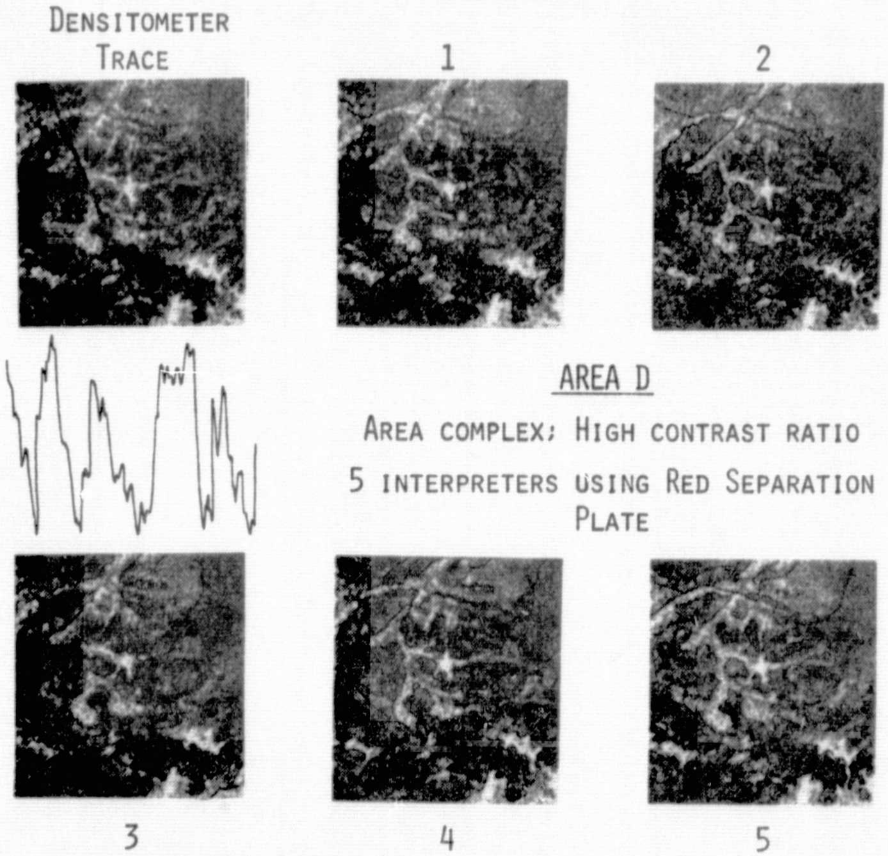


Figure 29. Interpretation of boundaries by five interpreters in an area where a complex of small sharply contrasting entities may be found.

At the opposite extreme the least comparable results were obtained in situations characterized by extensive, low contrast areas. The region illustrated in Figure 28 is predominantly a mulga scrub and spinifex landscape with linear sand dune country to the east of Napperby dry Lake. The film density trace shows clearly that almost no discrimination capability exists in this type of environment. About the only point of similarity between the interpretations is that all recognize the presence of Mt. Harris, the dark anvil-shaped area in the center of the frame.

Again, reference to the SKYLAB S-190B image in Figure 35 does not lend support to any of the interpretations, except that of Mt. Harris.

Finally, the last example of boundary conditions, Figure 29 depicts a complex pattern of highly contrasting types as shown on the Gemini photograph. The topography in this locality is hilly to mountainous which means the interpreter may inadvertently delineate shadows with other dark toned entities, and sunlit spots with light toned types. This problem has been magnified in this example because a black and white separation plate was used for analysis. The original color photo would give a more accurate view of variable illumination. Reference to Figure 35 shows that there is a significant improvement with the S-190B resolution in this instance also.

To summarize the results as given in Table 44, and in the preceding comparisons between Gemini, LANDSAT, and S-190B images, it is clear that resolutions of the order of those employed in both Gemini and LANDSAT introduce significant boundary artefacting, and incorrect interpreter boundary delineations. The S-190 B images are clearly at or near the threshold of accurate boundary delineation and the question we may now address is the apparent level of accuracy obtainable.

In order to assess this matter, reference should be made to Figures 31,

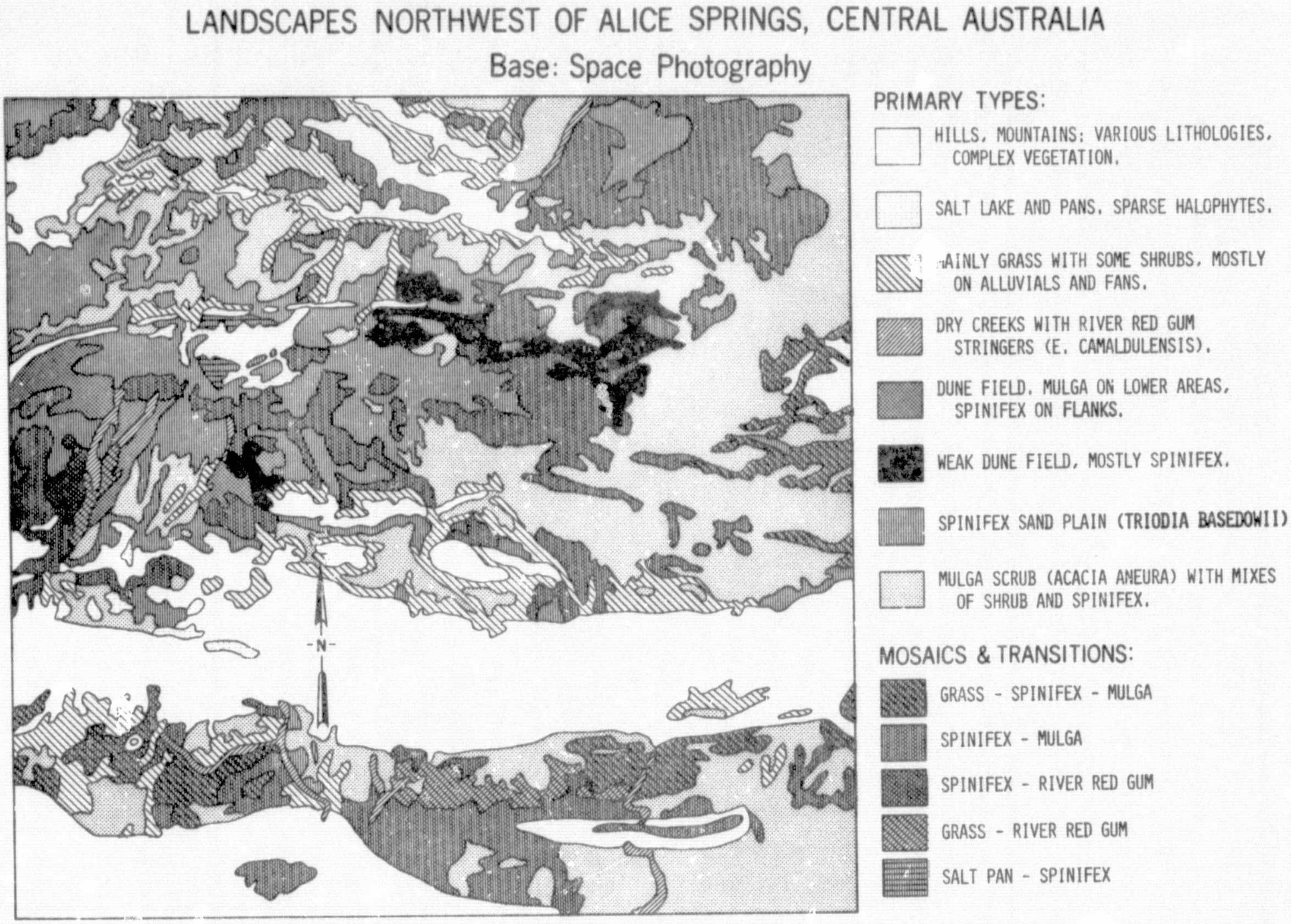


Figure 30. Landscapes northwest of Alice Springs, Central Australia. Boundaries and categories based upon Gemini space photography.

Figure 31. Vegetation Map of Kunoth Paddock. Scale 1:70,000. Vegetation map was prepared from ground truth data supplied by cooperating Australian scientists.

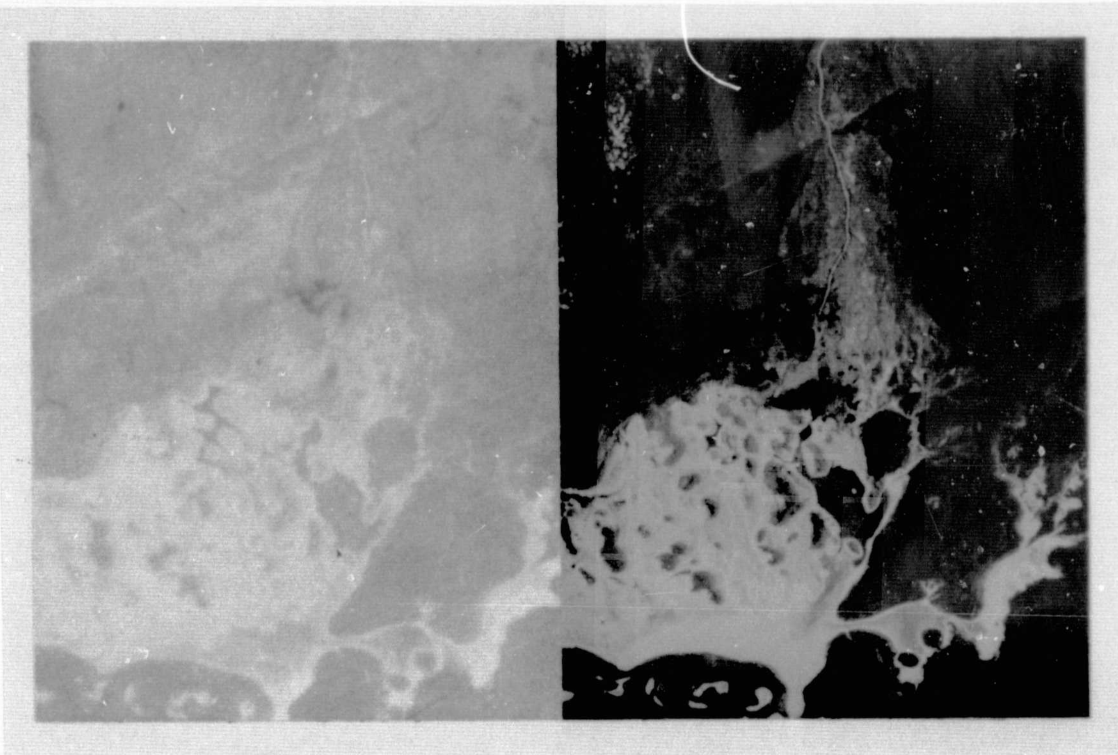
32 and 36, respectively showing the 1:70,000 Vegetation Map of Kunoth Paddock, prepared by C.S.I.R.O. personnel, a 1:70,000 color enlargement of the S-190B color image showing the same area, and sample low altitude air photos of Kunoth Paddock, obtained by C.S.I.R.O. personnel at the time of the August 12, 1973 SKYLAB mission.

Careful comparison between the three figures leads us to the following conclusions:

1. The S-190B image provides, in comparison to the Gemini and LANDSAT imagery, a quite extraordinary jump in quality of boundary detail available.
2. The level of detail available is such that enlargement to 1:20,000 or even larger would be feasible for more detailed mapping.
3. The S-190B photo suggests that the field-prepared plant community map is overgeneralized in numerous areas.
4. A significant number of the entity groups and boundaries, established by field mapping match in a general way, but not closely with those observed on the space photo.
5. An equally large number of apparently important plant community groups and boundaries, as seen on the space photo, and, as observed in Figure 36 are evidently both real and not mapped.
6. If the S-190B image had been available at the time of the field work, it would have had a significant impact in the field mapping. In short the categories established and mapped by the C.S.I.R.O. scientists on black and white aerial photographs are as subject to error as are any interpretations for the S-190B image. The S-190B image would have assisted in reducing this error.

The last of these conclusions is of considerable importance in assessing

Figure 32. Enlargement to 1:70,000 of SKYLAB S-190 B color photograph of the Kunoth Paddock area.



Band 5
LANDSAT

S-190 B
SKYLAB

Figure 33. Comparison between Band 5 LANDSAT I image, and the S-190 B SKYLAB color image of the eastern end of Napperby Dry Lake.

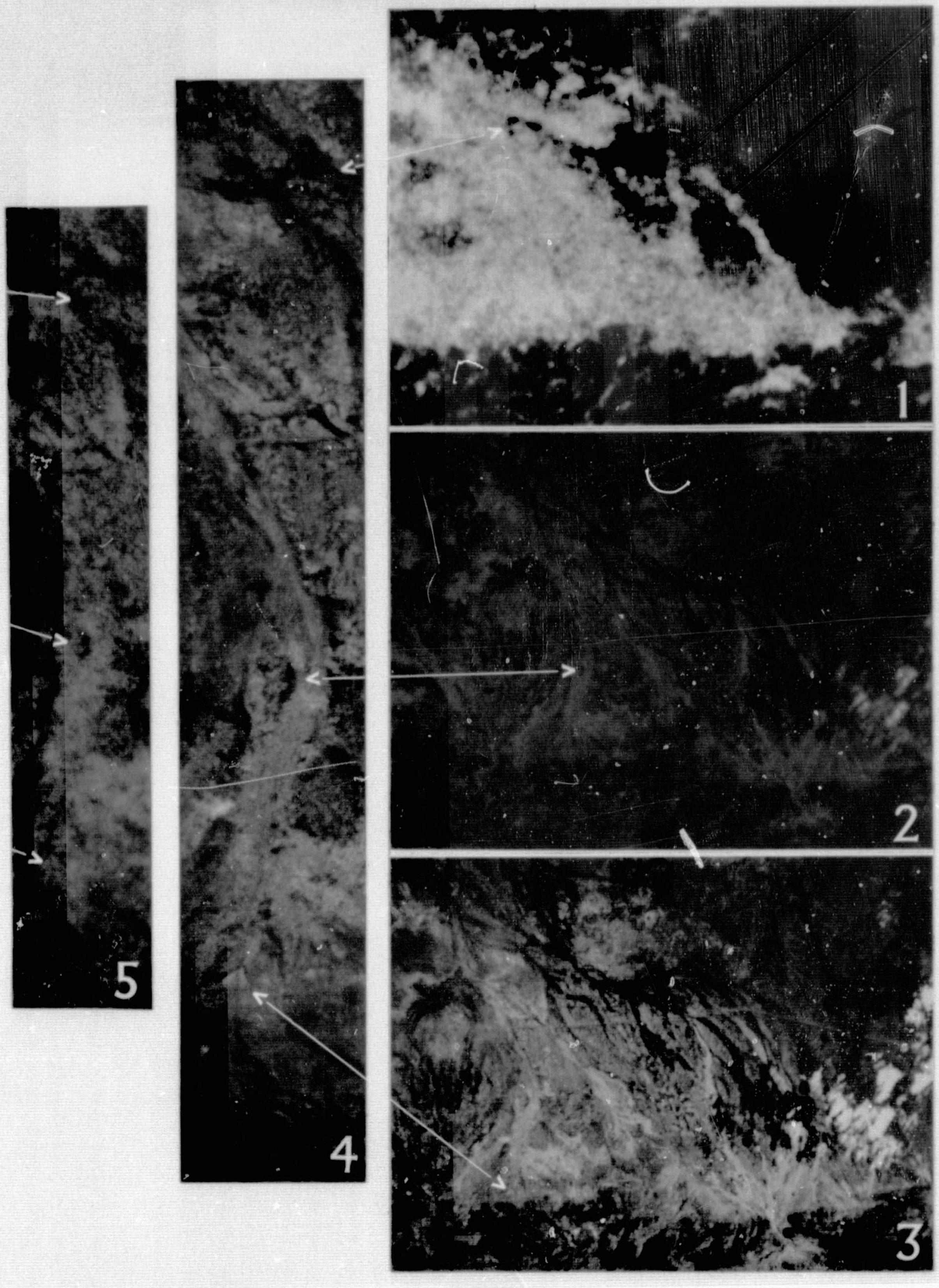


Figure 34. Comparison between 1:250,000 enlargements of the Kunoth Paddock area and further enlargements of portion of the same area.
1:250,000 1) Top: Gemini color photography; 2) Middle: S-190A color infrared image, and 3) Bottom: S-190B color photograph.
1:70,000 4) Color strip from S-190B
1:100,000 5) Color IR strip from S-190A

Figure 35. 1:250,000 enlargement of portions of the SKYLAB S-190B image, for comparison with the aerial oblique photographs of Figure 21 (15), Figure 22 (19, 30, 23, and 24) and airphoto mosaics and Gemini B & W enlargements (Figure 24, bottom; Figure 25, top)

and comparing future space images with previous conventional surveys. In this respect comments by Story, Yapp and Dunn (1976) introducing their comparative analysis of LANDSAT imagery are relevant.

"In 21 resources surveys over the past 25 years the Division of Land Use Research (C.S.I.R.O., Australia) has covered 2,135,800 km^2, or just over a quarter of mainland Australia. The reports constitute the Land Research Series of C.S.I.R.O., first published in 1962, and continuing. Survey methods are described in the reports and by Haantjens (1968) and Stewart (1968). In brief, the survey area is mapped and described in terms of homogeneous subdivisions (land systems) by means of stereo examination of black-and-white aerial photography supplemented by field work. The photos so far used have been varied in scale, with extremes of 1:16,000 and 1:85,000 for different parts of the same survey area, and quality has likewise varied from very high to extremely poor.

The paper deals with the evaluation of LANDSAT I imagery for this type of survey, with a view to evolving an inexpensive and adequate method to replace or supplement the aerial photography.

Problems of Evaluation Inherent in the Survey Methods

The reports refer to the Land Research mainland surveys as being "broad", "reconnaissance", or "general", and of "large areas", but the terms are not defined. Mapping scales vary from 1:250,000 to 1: 1,000,000, survey areas from 8240 to 373,000 km^2. land units, which are unmapped subdivisions of land systems, at least from 1.5 to 7250 km^2, in other words the terms "land

Figure 36: (On the following page) 1:35,000 Color enlargement of S-190B image compared with black and white 70 mm contact positives of color infrared vertical aerial photographs, also at a scale of 1:35,000. The photographs are portion of a strip covering the area adjacent to and including the western portion of the road between Alice Springs and Hamilton Downs (see Figure 11). The photos are arranged to run West to East, up and down the page and should be compared with Figures 31 and 32 by turning the page sideways. Note the degree of detail visible in the S-190B image. Particularly important in this respect is the identification of roads, small tracks, fence lines, stock ponds, small clumps of trees, eroded soil patches and so on. The ability to observe such small features, almost certainly means that the S-190B image will be able to be used without intermediate photographic scales for establishing sample sites for ground observations and measurements. This will lead to substantially improved precision in sampling. The special relevance of this observation is that the spatial and spectral resolution and the sensitivity expected with LANDSAT-D (30-meters, 6 channels, NE$\Delta\rho$ of .5%, see Harnage and Landgrebe, 1975) will be such that the same general quality of data will be available in computer compatible format. This will be a tour-de-force with respect to environmental inventory and monitoring.

The 70 mm Black and White images were reproduced by the Australian co-operating investigators, from color IR transparency rolls.

A, B, and C are common points for identification shown on both the Black and White and Color images.

Figure 36.

system" and "land unit" give no indication of the size of the area in question or the detail of the mapping. Their flexibility is further shown by the fact that one land system mapped by Christian and Stewart (1953) at 8 miles to the inch (1:506,880) was remapped by Speck _et al_. (1965) as 20 land systems at 4 miles to the inch (1:253,440); and land units in Carborough land system, which extends across the common boundary of the Dawson-Fitzroy and Isaac-Comet areas, number 10 and 4 in the respective reports (Speck _et al_., 1968; Story _et al_., 1967). Obviously team decisions and scale of mapping largely dictate the complexity of the land systems in any survey area and of the land units in any land system, and since the process is highly subjective, identical mapping would be very unlikely if two teams surveyed the same area independently. The survey reports do not stress this point, and some indeed tacitly present the land systems as distinct and definite by stating that they were _recognized_ during the interpretation, when in fact they were _erected_.

This subjectivity and lack of standardization are the biggest obstacles to a just assessment of the types of imagery used in Land Research survey series, and of the associated mapping."

The preceding comments from Story, Yapp and Dunn (1976) indicate the problem of comparison between Space images and Land System boundaries, and the highly subjective component in the latter. "Ground truth" as we have noted earlier is flawed. Simonett _et al_. (1969) (and repeated here in 3.3.1) noted that "problems of lumping and splitting apply to all qualitative judgements by men. It even applies to maps such as those prepared by Perry

(1961), themselves substantially based on aerial photographs, which have been used as "Ground Truth" to compare with the space photography."

We believe, to be candid, that Perry's map is unacceptable as a basis for assessing the value of SKYLAB S-190B image. Without being quite so blunt, Story, Yapp, and Dunn (1976) say that "this subjectivity and lack of standardization are the biggest obstacles to a just assessment of the types of imagery used in the Land Research survey series, and , the associated mapping." They could also have added" and, indeed, of assessing any space imagery." For these reasons, our comparisons have only been of discriminable boundaries not of the Land System units.

It is important, however, also to note that the same subjective element appears when we compare at the plant community level as shown in Figure 31 the vegetation map of Kunoth Paddock.

It is clear that "Ground Truth" is itself suspect, and that perhaps one of the greatest advantages of spacecraft imagery of the resolution of SKYLAB S-190B, in aiding surveying very large areas will be to re-vamp the procedures, methods of accuracy assessment and so on by providing multiple-scale views of an area. This is not what the conventional wisdom of using space imagery would indicate.

3.4 PRINCIPAL RESULTS OF THE STUDY

In the light of the questions posed for investigation in the introduction, what then are the principal results from these analyses. Each question may be taken up in terms in the following summary.

Question No. 6: What gains in grazing land category identification are there in comparison with either the S-190A, or earlier Gemini photography, through use of the S-190B system.

- There are significant improvements in identifiecation of grazing-land categories arising from the improved spatial resolution of the S-190B system, particularly with respect to detecting differences between plant communities on the basis of internal spatial variablity within a class. This is most noticeable with mulga communities. These abilities were in fact predicted by Simonett et al. in 1969 and are important in making space photography more useable as part of a sampling frame. The more categories detected in the first stage of a multi-stage sample, the greater the improvement in the precision of an estimate and within and between plan community discrimination and allocation of samples.

- The quality of category detection was such as to suggest substantial errors occurred in the initial ground truth mapping used to compare with the space images.

Question No. 7: What improvements are there and how significant are they with respect to landscape boundary delineation as a result of the improved spatial resolution of the S-190B.

- Paralleling the improvement in category identification there is a sharp reduction in the uncertainties of boundary placement, and a considerable increase in the number of possible additional categories and boundaries at lower hierarchical levels of classification of pasture types, and of smaller units, previously not detectable.
- Detection of boundaries of cultural features such as property boundaries, shown by differential grazing, and small tracks is greatly facilitiated with the S-190B. Thus, exact location of

sample sites is much improved, reducing the need for high flight or second stage (aircraft) data. Indeed, a 2-stage design using only spacecraft and ground observation may well be feasible. Further investigation will be needed firmly to establish this possibility.

Question No. 8: How do these improvements compare with predictions made in 1969 on the value of resolutions of 50 to 100 feet for pasture and range category separation and delineation in Central Australia?

- The improvements in pasture and range category separation and boundary delineation compare closely with predictions made by Simonett et al. (1969), based on calculations of the number of entities present in a resolution cell (see also Simonett and Coiner, 1971).
- Resolution, not additional bands, is the critical factor in such improvement. S-190A imagery, with infrared response, but with resolutions akin to the earlier Gemini photography is not adequate for additional separations. This conclusion is decidedly different for that of Colwell et al. (1974) who found that "for identification of natural vegetation types in the Colorado Plateau Test area, the EREP S-190A color IR image was judged better than all other image types tested" (including S-190B). However, they also recognized that this conclusion was arrived at because of the large natural vegetation units involved, rather than the fine scale found in the Central Australian example. We may also add that the low IR reflectivity of the Australian semi-desert plant communities at the time of imaging is also a factor. In short, there are environments where spatial resolution is critical and environments were multi-

spectral resolution is critical. Conclusions are thus environmentally modulated.

4.0 CONCLUSIONS

The most significant results from this investigation have been discussed in the body of this report. The major conclusions stated in the Executive Summary are discussed in further detail in this section.

The major conclusions are discussed first with respect to the Multispectral Scanner and the implication for future spacecraft systems and second with respect to the value of a high resolution photographic system. The conclusions for the multispectral scanner apply to both unmanned and manned systems (LANDSAT D, SHUTTLE), while the photographic systems apply mainly to the manned SPACE SHUTTLE.

4.1 S-192 MULTISPECTRAL SCANNER

The principal conclusions regarding the multispectral scanner deal with selection of spectral bands most useful for land use planning applications.

1. The 1.09-1.19 um band proved to be very valuable for discriminating a variety of land use categories, including within-category discrimination of agriculture, forest, and urban classes. This particular spectral band (band number 9) was selected as providing the best discrimination of any single spectral band when analyzing 1) urban classes, 2) agricultural classes, and 3) forest classes of data. Additionally, spectral band 9 was selected second in the analysis considering all Level I land use categories and was ranked first in the "overall" performance rating. The only analysis in which this spectral band was poorly rated was the wetlands analysis. Spectral band was rated tenth in the wetlands analysis. It should be noted however that spectral band number 10 (1.20 - 1.30 um) was rated first in the wetlands analysis. Data correlations between these two bands may be responsible for the low rating of spectral band number 9 in this analysis. This spectral range, which is not available on LANDSAT I or II, should be seriously considered for LANDSAT D.
2. The thermal infrared channel proved useful for discriminating between urban and vegetated categories. Spectral band number 13 the thermal infrared band (10.2 - 12.5 um) was not rated highly during any of the individual Level I

analyses however it was rated third in the analysis designed to identify those bands useful for discriminating between general Level I categories. The usefulness of the thermal infrared band for land use management purposes appears to be in discriminating urban areas from vegetated areas.

3. The 1.55-1.75 um band (spectral band number 11) proved very useful in combination with the 1.09-1.19 um band (spectral band number 9). This pair of spectral bands was selected amoung the top five bands in four of the six analyses performed during this investigation. The data content of these two spectral bands appears to be sufficiently independent to warrant consideration of using this pair of infrared bands on future satellite systems.

Additionally, it was concluded that:

4. Misregistration between spectral bands, even by as little as 1/2 pixel, may degrade classification accuracy.
5. Accuracy of identification of boundary or border pixels was as much as 13% lower than the accuracy for identifying internal field pixels. The implication of this drop is that in complex areas of small entities, or where there is spatial variability within a natural class, that it will be difficult to erect the physical boundaries between these entities using pattern recognition procedures unless one moves to higher spatial resolutions, perhaps of the order of those of LANDSAT D.
6. Multidate imagery may be necessary to accurately discriminate land use categories both at Level I and Level II. Throughout all phases of the digital analysis of the S-192 Multispectral Scanner data ambiguities were observed between

the spectral characteristics of different land use categories. Many of these ambiguities were rationally explainable and will most likely be resolved only through analysis of multidate repetitive coverage.

7. In order of overall ranking, the most useful six spectral bands were found to be:

   1) Spectral band number 9 (1.09-1.19 um),
   2) Spectral band number 3 (0.52-0.56 um),
   3) Spectral band number 6 (0.68-0.76 um),
   4) Spectral band number 1 (0.41-0.46 um),
   5) Spectral band number 11 (1.55-1.75 um),
   6) Spectral band number 13 (10.2-12.5 um).

4.2 S-190B Photographic System

The principal conclusion with respect to the S-190B camera system is-as would be expected-that the higher resolution of the S-190 B system in comparison to previous space photography (Gemini, Apollo), to the S-190 A system (SKYLAB), and to LANDSAT imagery, significantly increases the range of additional discrimination achievable. While evidence is available that the infrared layer of the S-190 A system and the IR bands of LANDSAT enables some identification not feasible (e.g. forest category separation, separation of small water bodies from forest, etc.) with color film in the S-190 B data, the high resolution is more generally useful than additional or alternative bands.

There is no reason why future photographic systems could not include color infrared film in association with high resolution. The advantages of high resolution of the order of the S-190 B system are as follows:

1. More categories can be identified with lower ambiguity. This conclusion applies not only to the natural plant communities in Central Australia, but also to the land use mapping carried out in the Baltimore-Washington area by Rohde and Simonett (1975), and by Bale et.al. (1975). Similar conclusions have been reached by numerous investigators, especially those working in areas of finely-fragmented land use types on the fringes of metropolitan areas.
2. Boundary delinieation, small-area field delineation and natural and land use category delineation is more precise and clearly matches that obtainable with high altitude aircraft data.

   In the Baltimore-Washington area:

   - Good quality information in map form can be expected from S-190 B imagery at both Level I and Level II as defined by

U.S. Geological Survey Circular 671.

- Variability in the results for test of Level III and Level IV, suggest that the spectral coverage of both color and color infrared film are needed for accurate identification and mapping in forest and agricultural classes.
- The experience of the interpreter markedly influences the quality of interpretation of both S-190A and S-190B imagery to a higher degree than is true of high altitude aircraft imagery.
- The information content of S190B is much more suitable for the land use and regional planner than is LANDSAT data, in areas of high urban-rural density. In other areas with larger natural plant community groups and with less intensive settlement, LANDSAT data and S190A may well be adequate.
- S-190B cannot acceptably substitute for all aircraft image uses by land use and regional planners, because among their responsibilities is the co-ordination of smaller jurisdictions with much more detailed data needs.
- SKYLAB S-190A and S-190B imagery can be used to update existing aircraft-based land use maps at the county level at a substantial cost savings (see Rohde and Simonett, 1975).
- SKYLAB S-190A and S-190B imagery can be used to revise, update, and improve the delineation of forest type boundaries at the county level in Maryland (see Rohde and Simonett, 1975).

In the Alice-Springs area:

- The SKYLAB S-190A images, with a spatial resolution of 75 meters (246 feet) for low-contrast features, is a definite improvement over Gemini photography and LANDSAT data, but nevertheless, the absence of any clear geometric identity to shapes makes identification of objects difficult. The threshold of resolution for significant entity discrimination is still not reached with S-190A.

- The SKYLAB S-190B images with a spatial resolution of 18 meters (59 feet) for a low contrast features lies at about the threshold for clear entity discrimination, and of understanding the basis for boundary differences.
- Careful item-by-item comparison with low altitude vertical and oblique photography shows that the detail observed in the S-190B images is at least comparable to that in high altitude aircraft photography, and in some respects is about the same as that observed in low altitude obliques.

3. Small cultural details, and natural landscape details can be observed in the S-190B photographs, thus reducing the need for an intermediate aerial photographic stage in <u>some</u> multistage sample designs. A major problem with Gemini and LANDSAT images, as part of a multi-stage sample design, is the difficulty of identifying exact locations so that ground samples may be directly related to the image. As a result aircraft imagery is required to bridge between the space photograph and ground sample. The S-190B imagery is of sufficient resolution that many ground locations may be accurately identified on the images. Thus:

- Fine details of a cultural type may be identified in most instances - individual isolated houses, farm lands, tracks, fence lines, small clumps of trees, erosion patches. This is commonplace, not exceptional, in both the Central Australian and Maryland environments and is clearly a fundamental property of imagery of about 60 feet spatial resolution (for low contrast targets). This fine supporting detail is critical in providing the location identifiers needed to eliminate some intermediate stages in a multi-stage sample design.

- Fine natural landscape details may also be observed. Examples that readily come to mind are the patchy, very small depressions and melaleuca scrub just east of Napperby Dry Lake in Central Australia. The most notable example is the ability to detect individual mulga groves, thus not only giving certain identification of the plant community, but also aiding in precise location.
- Areas of detailed and fine variations in lithology are readily observable, as in the northern portion of Kunoth Paddock. This improves the likelihood of correct location of ground samples.

4. Because high resolution in the first-spacecraft-stage of a multi-stage sample design enables better boundary delineation and category separation, greater precision is obtained in estimates of natural production of volume (timber, grazing land, agricultural production, etc.) than with LANDSAT data. While no actual estimates were made in the present study it is clear from comparisons with other work, notably that by Nicholls et. al. (1973) that the jump to S-190B resolution is very important in both improving the <u>precision</u> of estimates, and in <u>reducing</u> the cost and necessity for intermediate level aircraft flights.
5. The high-resolution of the S-190B greatly increases the analyst's confidence in reconnaissance to semi-detailed land, plant community or natural system mapping. This increased confidence arises from

the following factors:

- The level of detail available is such that enlargement to 1:20,000 or even larger is feasible for some very detailed mapping.
- Comparison with aerial photography shows that many features are clearly enough observed that the S-190B images may be used for updating existing land use or plant community mapping.
- Where there are differences between existing "ground truth" data and boundaries derived from the S-190B images the latter are clearly enough seen that the "ground truth" in some instances may require modification.

6. For many applied areas, the improved resolution brings the S-190B into the realm where existing data requirements of land-oriented agencies, may either be met, or the degree of modification to accept S-190B data is generally acceptable. Support for this view comes not only from the present investigation but also from other studies. Thus, in the earlier analyses by the authors it was observed that the trend towards planning for larger areas with a regional perspective (now evident in the planning community) was facilitating the acceptance and use of spacecraft and satellite-derived imagery (Bale et. al., 1975). Similarly, McKim, Merry, Cooper, Anderson and Gatto (1975) found in preparing land use maps with SKYLAB S-190B imagery for input data in hydrologic models, that the results "compared favorably with those obtained from high altitude aircraft photography." It seems clear from these and other examples that there will develop increasing acceptance of space data of the resolution of the S-190B system. It is only where

overlapping jurisdictions (state, regional, and local) require that the finest level of detail be employed, that the S-190B data cannot be used.

From these conclusions it is clear that there is an important role to be played on SPACE SHUTTLE for wide-area coverage, high resolution camera systems in land use and other resource analyses. It is strongly recommended that a camera system of the spatial resolution of the S-190B system - or slightly better - be carried at all times on SPACE SHUTTLE.

REFERENCES

Anderson, J.R., Hardy, E.E., and Roach, J.T. (1972). A Land-Use Classification System for Use with Remote Sensor Data. U.S. Geologic Survey Circular 671:1-16.

Anderson, T.W. (1958). Introduction to Multivariale Statistical Analysis, Wiley.

Bale, J.B., Rhode, W.G., Goehring. D., and Simonett, D.S. (1975). Evaluation of SKYLAB EREP Data for Land Resource Management, Technical Report, NASA Contract NAS 9-13314, Earth Satellite Corporation, Washington, D. C.

Biehl, L.L. and Silva, L.F. (1975). Evaluation of SKYLAB Data for Land Use Planning, Paper AAS 74-145, American Astronautical Society. 1-38.

Branch, M.C. (1971) City Planning and Aerial Information. Harvard City Planning Studies, No. 17, Harvard University Press, Cambridge, Massachusetts.

Brooner, W.G. and Wolfe, A. (1974). Maryland Land Use Classification Scheme. Maryland Tech. Series Pub. No. 209. Maryland Department of State Planning.

Christian, C.S., and Stewart, G.A. (1953). Genral Report on Survey of Katehrine-Darwin Region, 1946. CSIRO Australian Land Resources Series No. 1.

Colwell, R.N. (1961). The Analysis of Earth Resources on Enhanced Multiband Space Photography, in Sixth Symposium on Remote Sensing, 18, University of Michigan, Ann Arbor.

Colwell, R. N., Pettinger, L.R., Poulton, C.E., and Welch, R.I. (1974). A Comparison of SKYLAB and ERTS Data for Agricultural Crop and Natural Vegetation Interpretation, Technical Report, NASA Contract NAS 9-13286, Earth Satellite Corporation, Berkeley, California.

Dixon, W.J. Ed. (1973). Biomedical Computer Programs, University of California, Berkeley, California.

Doyle, F.J. (1967). Mapping from Satellite Photography. Unpublished text of a lecture delivered under the asupices of the American Society of Photogrammetry, NSF Visiting Scientist Program.

Eberlein, R., Vanderburg, G.J., Rosenfeld, A., Davis, L.S. (1974). Edge and Line Detection In ERTS Imagery: A Comparative Study, Technical Report, NASA Grant NGR-21-002-351, University of Maryland, College Park, Maryland.

Elterman, L. (1964). Atmospheric Attenuation Model, 1964, in the Ultra-Violet, Visible, and Infrared Regions for Altitude to 50km: Environmental Research Papers No. 46, Optical Physics Laboratory Project 7670, Air Force Cambridge Laboratories, Hanscom Field, Massachusetts.

Estes, J.E. and Senger, L.W. (1974). Remote Sensing: Techniques for Environmental Analysis. Hamilton Publishing Company, Santa Barbara, California. 340 pp.

Haantjens, H.J. (1968). The Relevance for Engineering of Principles, Limitations, and Developments in Land System Surveys in New Guinea,in Proc. 4th Conf. Australian Roads Resources Board 4, (2) 1593-1612.

Hardy,E.E., Skaley, J.E., Dawson, C.P. and Weiner, G.D. (1975). The Evaluation of Enhanced SKYLAB Photography for Land Use Inventory in New York, in SKYLAB Earth Resources Experiment Package, Regional Planning and Development Conference, 1-27. Purdue University, West Lafayette, Indiana.

Harnage, J. and Landgrebe, D., Eds. (1975). LANDSAT-D Thematic Mapper Technical Working Group, Final Report, JSC-69797, NASA Lyndon B. Johnson Space Center, Houston, Texas.

Holmes, J.H. (1967). Problems in Location Sampling. Annals Association of American Geographers, 57 (4) 757-780.

Holmes, J.H. (1970). The Theory of Plane Sampling and Its Application in Geographic Research, Economic Geography, 46 (2) 379-392.

Hoffer, R.M., and Fleming, M.D. (1975). Definition of an Optimal Wavelength Band Combination for Forest Cover Type Mapping using SKYLAB S-192, in SKYLAB Earth Resources Experiment Package, Regional Planning and Development Conference. Paper read but not mimeographed. Purdue University, West Lafayette, Indiana.

McKim, H.L., Merry, C.J., Cooper, S., Anderson, D.M. and Gatlo, L.M. (1975). Applications of Remote Sensing for the Corps of Engineers Program in New England, in Proc. of 10th International Symposium on Remote Sensing, Ann Arbor, Michigan.

Morain, S.A., and Williams, D.L. (1973). Estimate of Winter Wheat Yield from ERTS-1, Third Earth Resources Technology Satellite Symposium, NASA SP-351, 1, 21-28. Goddard Space Flight Center, Washington, D. C.

Nichols, J.D., Gialdini, M. and Faakkola, S. (1963). A Timber Inventory Based Upon Manual and Automated Analysis of ERTS-1 and Supporting Aircraft Data Using Multi-Stage Probability Sampling, Third Earth Resources Technology Satellite Symposium, NASA SP-351,1,145-159. Goddard Space Flight Center, Washington, D. C.

Perry, R.A. (1961). Pasture Lands of the Alice Springs Area: Map (1:1,000,000), Division of Land Research and Regional Survey, CSIRO, Melbourne, Australia.

Perry, R.A., Mabbutt, J.A., Litchfield, W.H., Quinlan, T., Lazarides, M., Jones, N.O., Slayter, R.O., Stewart, G.A., Bateman, W., and Ryan, G.R. (1962). Lands of the Alice Springs Area, Northern Territory, 1956-57, CSIRO Australia Land Resources Series No. 6.

Rao, C.R. (1965). Linear Statistical Inference and its Applications, section 8e, Wiley.

Rohde, W.G. and Simonett, D.S. (1975). Land Use Studies with SKYLAB Data, in SKYLAB Science Experiments (G.W. Morgenthaler and G.E. Simonson, Eds.) 127-185. American Astronautical Society, Science and Technology Series, Volume 38. Tarzana, California.

Sattinger, I.J. (1975). Analysis of Recreational Land Using SKYLAB Data, in SKYLAB Earth Resources Experiment Package, Regional Planning and Development Conference. Purdue University, West Lafayette, Indiana.

Schwarz, D.E., Simonett, D.S., Jenks, G.F., and Ratzlaff, J. (1969). The Construction of Thematic Land Use Maps with Spacecraft Photography, Technical Report U.S.G.S. Contract 14-08-0001-10848, University of Kansas, Lawrence, Kansas.

Silva, L.F., and Biehl, L.L. (1975). Study of the Utilization of EREP Data from the Wabash River Basin, in SKYLAB Earth Resources Experiment Package Regional Planning and Development Conference. Mimeo. 1-36. Purdue University, West Lafayette, Indiana.

Simonett, D.S., Cochrane, G.R., Morain, S.A. and Egbert, D.D. (1969). Environment Mapping with Spacecraft Photography: A Central Australian Example. U.S. Geological Survey Interagency Report NASA-172. Washington, D.C.

Simonett, D.S., and Coiner, J.C. (1971). Susceptibility of Environments to Low-Resolution Imaging for Land Use Mapping, in Proc. of the Seventh International Symposium on Remote Sensing of the Environment. 373-394. University of Michigan, Ann Arbor, Michigan.

Simonett, D.S., and Hajic, E.J. (1976). Comparisons of Qualitative and Quantitative Image Analysis, in Remote Sensing of Environment (J.R. Lintz and D.S. Simonett, eds.) in press. Addison-Wesley, Reading, Massachusetts.

Speck, N.H., Weight, R.L., Van de Graaff, R.H. M., Fitzpatrick, B.A., Mabbutt, J.A., and Stewart, G.A. (1965). General Report on Lands of the Tipparrary Area, Northern Territory, 1961, CSIRO Australian Land Resources Series No. 13.

Speck, N.H., Wright, R.L., Sweeny, F.C., Perry, R.A., Fitzpatrick, P.A., Nix, H.A., Gunn, R.H., and Wilson, I.B. (1968). Lands of the Dawson-Fitzroy Area. Queensland, CSIRO Australian Land Resources Series No. 21.

Stewart, G.A. (1968). Ed. Land Evaluation. Papers of a CSIRO Symposium organized in cooperation with UNESCO, Macmillan of Australia.

Story, R., Galloway, P.W., Gunn, R.K., and Fitzpatrick, H.A. (1967). Lands of the Isaac-Comet Area, Queensland CSIRO Australian Land Resources Series No. 19.

Story, R., Yapp, G.A., and Dunn, A.T., (1976). LANDSAT Patterns Considered in Relation to Australian Resources Surveys, Remote Sensing of Environment. In press.

Swain, P.H., Robertson, T.V., and Wacker, A.G. (1971). Comparison of the Divergence and B-Distance in Feature Selection, Information note 020871. Laboratory for Application of Remote Sensing, Purdue University, West Lafayette, Indiana.

Swain, P.H., and King, R.C. (1973). Two Effective Feature Selection Criteria for Multispectral Remote Sensing in Proc. of the First International Joint Conference on Pattern Recognition, 536-540. Washington, D. C.

Vegas, P.L. (1974). A Detailed Procedure for the Use of Small-Scale Photography in Land Use Classification. NASA Technical note TN-D-7542. 43 pp.

Westerlund, F.V. (1972). Remote Sensing for Planning, Urban Planning Development Series No. 10, Department of Urban Planning, University of Washington, Seattle, Washington.

Wobber, F.J. (1970). Orbital Photography Applied to the Environment, in Photogrammetric Engineering, 36 (8), 852-864.

Wigton, W.H., and VonSteen, D.H. (1973). Crop Identification and Acreage Measurement Utilizing ERTS Imagery, Third Earth Resources Technology Satellite Symposium, NASA SP-351,1, 87-92. Goddard Space Flight Center, Washington, D. C.