AD-A045 871 ARMY ENGINEER WATERWAYS EXPERIMENT STATION VICKSBURG MISS F/G 15/4
ACQUISITION OF TERRAIN INFORMATION USING LANDSAT MULTISPECTRAL --ETC(U)
SEP 77 H STRUVE, W E GRABAU, H W WEST

UNCLASSIFIED WES-TR-M-77-2-2 NL

1 OF 2
ADA045871

TECHNICAL REPORT M-77-2

ACQUISITION OF TERRAIN INFORMATION USING LANDSAT MULTISPECTRAL DATA

Report 2

AN INTERACTIVE PROCEDURE FOR CLASSIFYING TERRAIN TYPES BY SPECTRAL CHARACTERISTICS

by

Horton Struve, Warren E. Grabau, and Harold W. West

Mobility and Environmental Systems Laboratory
U. S. Army Engineer Waterways Experiment Station
P. O. Box 631, Vicksburg, Miss. 39180

September 1977

Report 2 of a Series

Prepared for Assistant Secretary of the Army (R&D)
Department of the Army
Washington, D. C. 20310

Under Project 4A161101A91D, Task 02
Work Unit 095

ORIGINAL CONTAINS COLOR PLATES: ALL DDC REPRODUCTIONS WILL BE IN BLACK AND WHITE

Destroy this report when no longer needed. Do not return it to the originator.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report M-77-2	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER: WES-TR-M-77-2-2	
4. TITLE (and Subtitle): ACQUISITION OF TERRAIN INFORMATION USING LANDSAT MULTISPECTRAL DATA. Report 2. AN INTERACTIVE PROCEDURE FOR CLASSIFYING TERRAIN TYPES BY SPECTRAL CHARACTERISTICS.	5. TYPE OF REPORT & PERIOD COVERED: Report 2 of a series
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Horton/Struve, Warren E./Grabau, Harold W./West	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: U. S. Army Engineer Waterways Experiment Station, Mobility and Environmental Systems Laboratory, P. O. Box 631, Vicksburg, Miss. 39180	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: Project 4A161101A91D, Task 02, Work Unit 095
11. CONTROLLING OFFICE NAME AND ADDRESS: Assistant Secretary of the Army (R&D), Department of the Army, Washington, D. C. 20310	12. REPORT DATE: Sept 1977
	13. NUMBER OF PAGES: 138
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 140 p.	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

1 — AD42999

18. SUPPLEMENTARY NOTES

A-039 521

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Data acquisition
Landsat (Satellite)
Multispectral data
Spectrum analysis
Terrain classification

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Developed in the study reported herein was a semiautomated procedure for classifying Landsat radiance data in terms of preselected land-use categories. The procedure is an interim solution to the problem of mapping very large areas in terms of relatively crude categories in very short periods of time.

Operation of the procedure requires an analyst to direct a computer by means of interactive instructions to search for all the 3 by 3 pixel arrays

(Continued)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

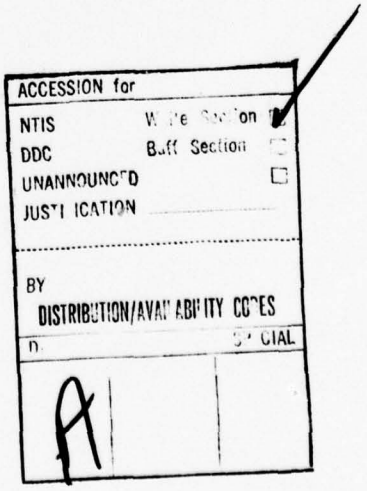
038 100

20. ABSTRACT (Continued).

exhibiting spectral signatures that conform to a selected criterion of homogeneity. The computer then retrieves these signatures from within the array of Landsat radiance values and groups them into spectrally similar clusters. The clusters are then displayed on a color coded map overlay from which the analyst must provide the final interpretation and classification.

The area selected for study is centered approximately 40 km northeast of Vicksburg, Mississippi. The area includes a representative section of loess hills, forming the eastern wall of the Mississippi floodplain, and a section of the floodplain, including an oxbow lake and a number of other floodplain features.

THE CONTENTS OF THIS REPORT ARE NOT TO BE USED FOR ADVERTISING, PUBLICATION, OR PROMOTIONAL PURPOSES. CITATION OF TRADE NAMES DOES NOT CONSTITUTE AN OFFICIAL ENDORSEMENT OR APPROVAL OF THE USE OF SUCH COMMERCIAL PRODUCTS.



PREFACE

This report is the second of a series dealing with the manipulation and interpretation of Landsat digital data. The concepts and procedures reported herein were developed under the In-House Laboratory Independent Research Program, Project 4A161101A91D, Task 02, Work Unit 095, "Feasibility of Using Landsat Spectral Data for Acquisition of Terrain Information for Multiple Purposes." The work was performed during the period September 1975-June 1976 by personnel of the Environmental Simulation Branch (ESB), Environmental Systems Division (ESD), Mobility and Environmental Systems Laboratory (MESL), U. S. Army Engineer Waterways Experiment Station (WES), under the direct supervision of Messrs. H. W. West, Project Manager, and J. K. Stoll, Chief, ESB. The study was under the general supervision of Messrs. B. O. Benn, Chief, ESD, and W. G. Shockley, Chief, MESL. Dr. H. Struve, ESB, was responsible for the development of the interactive procedure for classifying terrain types by spectral characteristics. Special acknowledgment is made to Mr. W. E. Grabau, Special Assistant, MESL, for his constructive criticisms and encouragement throughout the study. Dr. Struve and Messrs. Grabau and West prepared the report.

COL G. H. Hilt, CE, and COL J. L. Cannon, CE, were Directors of the WES during the study and report preparation. Mr. F. R. Brown was Technical Director.

CONTENTS

	Page
PREFACE	2
CONVERSION FACTORS, METRIC (SI) TO U. S. CUSTOMARY AND U. S. CUSTOMARY TO METRIC (SI) UNITS OF MEASUREMENTS	4
PART I: INTRODUCTION	5
Background	5
Objective and Scope	6
Rationale of the Procedure	6
PART II: DEVELOPMENT OF SEMIAUTOMATED PROCEDURE	9
Theory	9
Practice	22
PART III: EVALUATION AND ADJUSTMENTS	94
Evaluation	94
Adjustments	99
PART IV: CONCLUSIONS AND RECOMMENDATIONS	104
Conclusions	104
Recommendations	104
REFERENCES	105
TABLES 1-26	
PLATES 1 and 2	
APPENDIX A: DESCRIPTION OF OPTRONICS FILM READER/WRITER	A1

CONVERSION FACTORS, METRIC (SI) TO U. S. CUSTOMARY AND U. S. CUSTOMARY TO METRIC (SI) UNITS OF MEASUREMENT

Units of measurement used in this report can be converted as follows:

Multiply	By	To Obtain
Metric (SI) to U. S. Customary		
micrometres	3.937007×10^{-5}	inches
millimetres	0.03937007	inches
centimetres	0.3937007	inches
metres	5.399568×10^{-4}	miles (U. S. nautical)
metres	6.213711×10^{-4}	miles (U. S. statute)
kilometres	0.5399568	miles (U. S. nautical)
kilometres	0.6213711	miles (U. S. statute)
square kilometres	247.1054	acres
square kilometres	0.291553	square miles (U. S. nautical)
square kilometres	0.3861021	square miles (U. S. statute)
hectares	2.471054	acres
U. S. Customary to Metric (SI)		
feet	0.3048	metres
degrees (angular)	0.01745329	radians

ACQUISITION OF TERRAIN INFORMATION USING LANDSAT MULTISPECTRAL DATA

AN INTERACTIVE PROCEDURE FOR CLASSIFYING TERRAIN TYPES BY SPECTRAL CHARACTERISTICS

PART I: INTRODUCTION

Background

1. Modern military operations require vast amounts of terrain information, both for long-range strategic planning and quick-response tactical planning. There is also a host of more prosaic needs for terrain information, such as long-range planning for and monitoring of military facilities uses. The various agencies that supply the U. S. Armed Forces with terrain information have used conventional air-photo interpretative procedures for many years.[1,2] Unfortunately, needs have grown more rapidly than capabilities, and thus a requirement for faster and less costly ways of acquiring the necessary terrain information is urgent. One possible avenue is that opened by the Landsat satellites and their proposed follow-ons.

2. There are (1977) two satellites (Landsat 1 and Landsat 2) in orbit around the earth. Each satellite circles the earth every 103 min and contains a multispectral scanner (MSS) that provides radiance measurements of terrain materials for four spectral bands (0.5-0.6, 0.6-0.7, 0.7-0.8, and 0.8-1.1 μm*) for each pixel area (≈57 by 79 m) of the terrain surface. Spectral data for an area on the ground are provided every 18 days by each orbiting satellite, and the two satellites are scheduled in such a manner that actual coverage (i.e. spectral data) is

---

* A table of factors for converting metric (SI) units of measurement to U. S. customary units and U. S. customary units to metric (SI) units is presented on page 4.

obtained every 9 days.* Computer-compatible tapes (CCT's) of the spectral data and imagery made from the tapes are made available by the Earth Resources Observation Systems (EROS) Data Center** approximately 6 weeks after an area has been scanned by the satellite sensor system.

Objective and Scope

3. The objective of the research described herein was to develop a rapid and semiautomated method of classifying Landsat radiance data in terms of preselected land-use categories or terrain types. The procedure is intended as an interim solution to the problem of mapping very large areas (in excess of 100,000 km^2) in terms of relatively crude categories in very short periods of time. A major requirement was that the procedure operate with a minimal amount of processing of primary Landsat data. Ideally, it would operate by manipulating Landsat data as it is recorded by the satellite, without data decompression, calibration, or other modification.

Rationale of the Procedure

4. Any semiautomated procedure must be somewhat anagolous to the procedure(s) used by human interpreters; therefore, an understanding of the various methods by which human interpreters identify terrain and land-use information from images is necessary. There appear to be three basic techniques:

a. Identification of features or materials by spectral analysis. This process assumes prior knowledge of the absorption and reflection characteristics of the feature for which a search is being made. Identification is made

---

* At the time this report was being published, the National Aeronautics and Space Administration (NASA) was rescheduling the Landsat coverage intervals. After the proposed launch in 1978 of Landsat C, Landsat 1 will follow Landsat 2 six days later, and Landsat C will follow Landsat 2 nine days later.

** United States Department of the Interior, Geological Survey, EROS Data Center, Sioux Falls, S. Dak. 57198.

by matching the known spectrum with the spectrum as revealed by the sensor system.

b. Identification of features or materials by pattern recognition. This process assumes that each feature or material exhibits a unique pattern of colors or shades of gray. Recognition is achieved when a pattern on an image can be correlated with an object or material on the ground.

c. Identification of features or materials by association. This seems to be the major process used by the human interpreter. Despite the fact that it is widely used, it is very poorly understood. In fact, it cannot even be defined except by an example: A smooth gray area around a residential building is almost instantly recognized by an interpreter as a lawn, not because it is smooth and gray, but because it is positioned in a certain way with respect to the building. The precise same shade of gray associated with a building identified as a factory might well be interpreted as a parking lot. An immense amount of time is devoted by educational institutions to the development of the photo interpreter's ability to make associations, and an interpreter's skill is largely measured by the number of associations he can reliably make.

5. The interactive procedure uses the developed skill of a human interpreter; the human makes the primary interpretation and then specifies to a computer program those combinations of spectral values that are to be accepted as representing the terrain types of interest on the ground. The basic reason for exploiting an interactive procedure is that the human interpreter is capable of making subjective judgments on the basis of all three of the techniques described above; whereas at the present level of development, programmed interpretation can be based only on the first of the three techniques (i.e. spectral composition). The advantage is that classifications can sometimes be based on quite subtle criteria, such as those that cannot be revealed or defined by rigorous objective analysis.

6. The disadvantages, however, are compelling. Perhaps the most important is that each image must be independently subjected to the interpretative process. A classification made on the basis of an image obtained on Monday cannot be reliably carried over to an image obtained on Tuesday. Nor can classifications be reliably transposed to another

geographic region, even if close by. The reasons are many. The most important are as follows:

a. Spectral characteristics are changed by differences in atmospheric composition.

b. Spectral characteristics are changed by differences in reflectance geometry.

c. Spectral characteristics are changed by time-dependent variations in terrain conditions.

7. The skilled human interpreter can often, somehow, sort out and subjectively compensate for all of these things; but the basic digital data, on which a computer program must base its decisions, may be quite dramatically different from scene to scene. Writing a program that will automatically make the necessary adjustments and accommodations is well beyond current capability. However, it is clear that the long-range goal of programmed interpretation is a program that will make the necessary adjustments. A portion of such a program is discussed in detail in Reference 3.

PART II: DEVELOPMENT OF SEMIAUTOMATED PROCEDURE

Theory

Primary data versus processed data

8. In principle, the primary record stored on Landsat CCT's contains all of the spectral data that will be available. The primary data may be modified by calibrating them to references and thus converted to true radiance values. They may be corrected for atmospheric transmittance and other effects, but those processes merely change the absolute values; they do not change the actual informational content. Therefore, at least in theory, the primary CCT record is as useful for purposes of terrain classification as those data after calibration and other forms of processing. That is, if "natural" classes exist, they will be inherent in the primary data as well as in the processed data.

9. This suggests that correlations that relate natural classes of CCT primary data with conditions on the ground (i.e. with terrain types) are at least as valid as those obtained with processed data. If this is so, then no advantages are to be gained by correcting primary data in any way prior to classification. Thus, if interpretation is to be based strictly on the establishment of correlations between spectral composition of pixels (i.e., values from each of four wavelength bands, if the sensor is Landsat) and terrain types (not, be it noted, between spectral composition of pixels and known spectral characteristics of the terrain types), then the primary data are at least as good as any form of processed data.

Determining spectral signatures

10. To simplify subsequent discussion, the primary Landsat data defining the spectral characteristics of a pixel is called a "CCT value set." Thus, a CCT value set is an array of four CCT values, one for each wavelength band.

11. Let it be assumed that the CCT value set reflected from a particular terrain type always consists of four fixed CCT values under any given set of conditions (i.e. of atmospheric transmittance and

reflectance geometry). Now suppose that a region including several different terrain types, each always exhibiting a set of four fixed CCT values, is imaged by a multispectral scanner, such as that aboard Landsat. If the CCT values of any one of the four spectral bands were assembled into a histogram of the number of pixels versus the CCT values, the histogram would consist of a set of discrete bars, each representing the number of pixels exhibiting a particular CCT value. For example, a portion of a Landsat scene consisting of 33,600 pixels might exhibit the histogram shown in Figure 1.

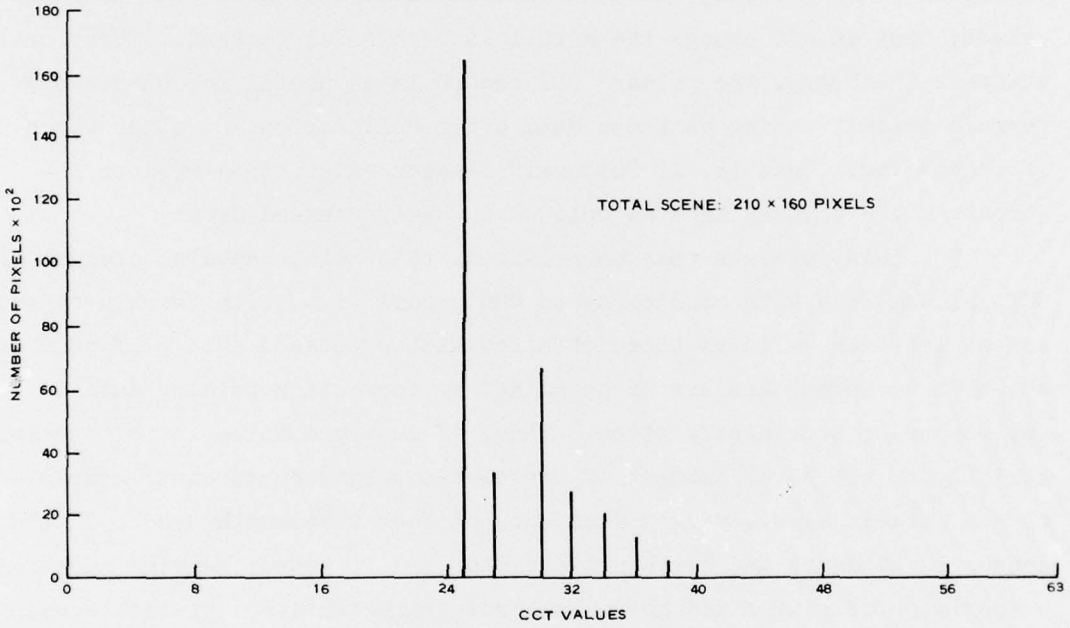


Figure 1. Idealized histogram of one wavelength band

12. However, experience has shown that CCT value sets for a given terrain type tend to distribute themselves to either side of average values rather than remain fixed values. Therefore, histogram sets of several terrain types using real world data (Figure 2) will take the form of a continuous curve rather than discrete bars. There are at least three major sources that act randomly and in concert to cause the CCT values to fluctuate.

a. The Landsat scanner has an internal error of about ±2

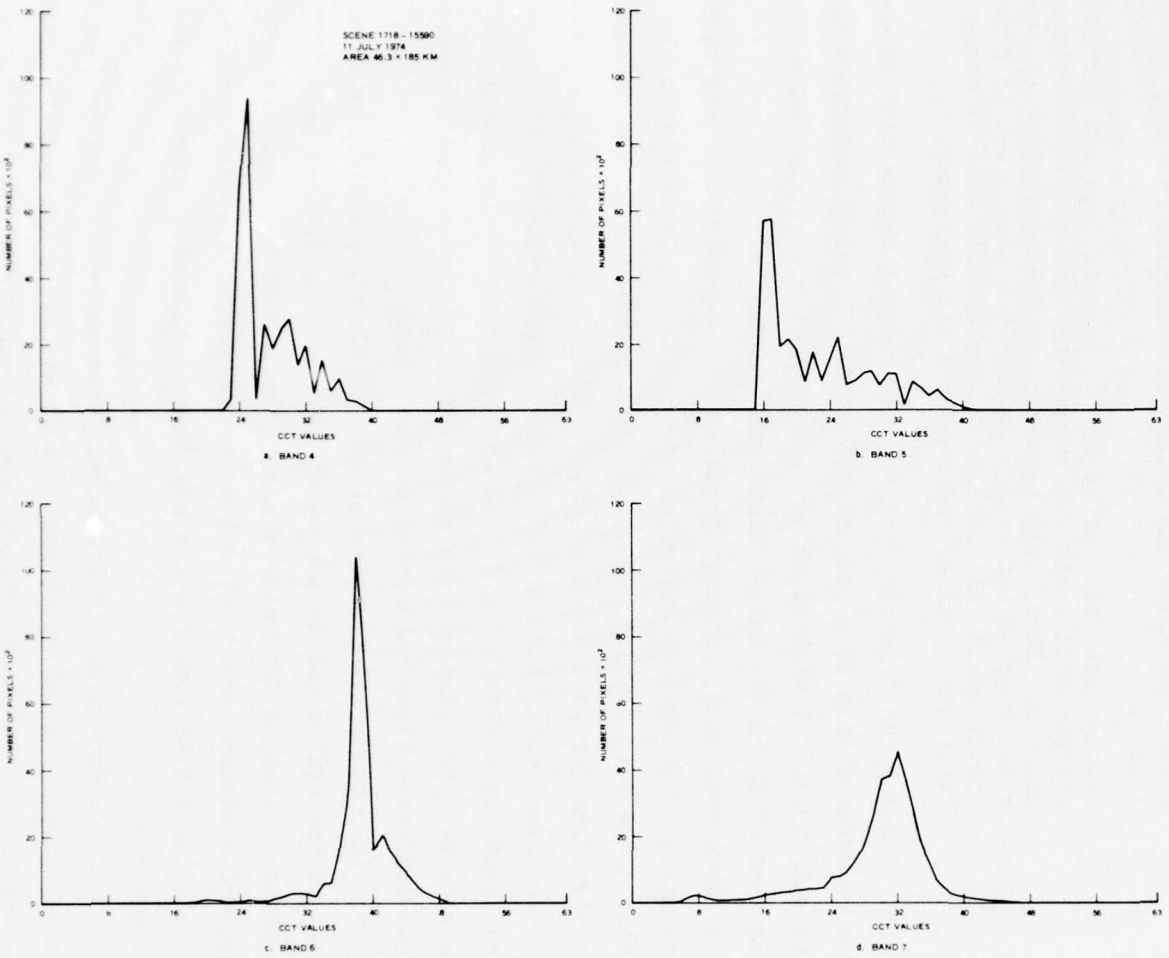


Figure 2. Histogram of wavelength bands of one quarter of Landsat scene

percent of total range. Thus, for example, a true radiance value of 30 may be recorded as any value between 28.7 and 31.3. Since the recorded CCT values are always integers, the values become 29 and 31.

b. Local variations in atmospheric transmittance and terrain conditions alter the amount of radiation reaching the sensor.

c. Pixels that fall partly in one terrain type and partly in another will yield CCT values that are somewhere between the two. Thus, they belong to no "naturalistic" classes, since they do not really represent a single set of ground conditions.[4] All Landsat images contain some of these "spurious" pixels, and most contain very many.

13. Despite the "smearing" of the theoretical values, let it nevertheless be assumed that each peak represents a CCT value characteristic of a terrain type. It might also be assumed that a distinct "shoulder," such as that at 38 in the band-4 histogram (Figure 2a), would represent a weakly developed peak, possibly because the area covered by the terrain type is small. The obvious extension of this notion is to divide the CCT value scale into classes, with the class boundaries at the valleys of the histogram (or at a break in slope, leading to a shoulder). Each peak or shoulder would then be, in effect, a modal value for that spectral band for a terrain type. Table 1 illustrates the product of such a process. The classes were chosen subjectively; others may choose somewhat different class boundaries in some cases.

14. In consideration of the discussion above, a naturalistic subdivision of the CCT values in one wavelength band is defined as a "CCT value class." Thus, a CCT value class is assumed to be characterized by a normal distribution of CCT values around a specified modal value, with the class limits some fixed standard deviation away on either side of the modal value.

15. Since each pixel in a Landsat image is represented by four CCT values (i.e. by a CCT value set), it is at this point easy to leap to the conclusion that each terrain type is represented by a unique combination of CCT value classes. However, the fact that not all wavelength bands exhibit the same number of CCT value classes (e.g., band 4 has 7, and band 6 has 9) illustrates a critical point; namely, not all

of the CCT values in a set need be unique. Only the set as a whole is unique. For example, one terrain type might be characterized by values in classes 1, 1, 1, 1 (bands 4, 5, 6, and 7, respectively), and another terrain type characterized by classes 1, 1, 1, 2. Thus, at least potentially, there are as many unique spectral "signatures" as permutations of CCT value classes. In the example (Table 1), there may be 4032 (7 × 8 × 9 × 8 = 4032) unique combinations of classes.

16. For future discussion, let an array of four CCT value classes be called a "CCT class set." Hence, a CCT class set is assumed to be the "spectral signature" of a unique combination of conditions on the ground.

Relating spectral
signatures to terrain types

17. The central questions at this point are: Which of the possible combinations are "real" (i.e. originate from an identifiable terrain condition or type on the ground), and how may the relation between the spectral signature (i.e. CCT class set) and the terrain type be established? The following two basic modes by which these questions can be answered are dependent on the state of prior knowledge:

a. An "unguided" mode with no prior knowledge of the terrain type spectral signature relation. In other words, the analyst has no knowledge of the actual distributions of terrain types on the ground or, indeed, of what types actually exist in the region. Two major variants of the unguided mode are known; these are discussed below.

b. A "guided" mode with analyst knowing what terrain types exist and the geographic locations of some presumably typical or characteristic examples of each.

18. Unguided mode, first variant. In concept, it would be possible to construct a map showing the geographic distribution of all pixels characterized by any specified CCT class set. For example (Table 1), all class 1 CCT value ranges (22-26, band 4; 15-18, band 5; 17-23, band 6; and 5-11, band 7) would be selected, and all pixels exhibiting any set of values encompassed by these ranges would be located in geographic space. These locations would then (somehow!) be examined in the field to determine what terrain condition or type, if any, was

the source. This process would be repeated for each of the possible permutations. A situation such as that defined by the Figure 2 histograms would mean 4032 such exercises. In most cases, the sheer magnitude of the task makes this approach impractical.

19. There is, however, a way of dramatically reducing the size of the task. It depends on the intuition that not all possible permutations of CCT value classes represent actual conditions on the ground. Thus, an obvious approach is to determine the number of pixels in each permutation of CCT value classes. Since most (if not all) landscapes incorporate only a very limited number of distinct terrain types, it would be expected that the result would be a relatively small number of CCT class sets having a very large number of pixels. It would also be expected that a relatively large number of CCT class sets would show a very small number of pixels. The latter condition would be expected because, purely by chance, spurious pixels (see paragraph 12c) will combine CCT values in such a way as to meet the criteria of one or another of the possible CCT class sets.

20. At this point, an arbitrary minimum number of pixels could be selected as a cutoff between presumed "real" signatures (those deriving from actual terrain types) and "false" signatures (those deriving from spurious pixels). When this procedure is actually employed, the number of CCT class sets exhibiting significantly large pixel numbers falls to relatively small numbers, normally between 20 and 70. This is still a comparatively large number to field check, but the number is much more tractable than the original several thousand.

21. It is instructive to examine the reasons for the persistence of such large numbers of CCT class sets. The basic reason is that the sensor 'sees" and records only radiation, and any change in radiation intensity is duly recorded. Therefore, any change in condition that results in a change in reflectance is perceived as a new spectral signature. On the other hand, human perception of terrain type tends to be based on intrinsic qualities of the terrain. A hardwood forest covering all sides of a hill is perceived as "forest," uniform over the entire hill. However, the satellite sees the slope away from the sun quite

differently from the slope toward the sun and thus records at least two (and in practice, several) signatures from the same hill. The consequence is that the sensor records several terrain types (i.e. different CCT class sets) for the feature that human perception and experience insist is a single homogeneous terrain type. Since it is the human needs that must be satisfied, the problem is to determine which of the sensor-perceived signatures are derived from each of the terrain types required by the human analyst.

22. For example, suppose the analyst wanted to map the distribution of only three terrain types: open water, forest, and nonforest. In effect, this means that the 4032 potential CCT value sets (see paragraph 15) or the residual 20 to 70 CCT value sets (see paragraph 20) must be allocated among only those three categories. In the absence of any information on either the location of actual areas exhibiting the three terrain types, or information on the probable signatures of the three types, nothing further can be done. Further interpretation requires some process by which the relations among terrain types and CCT value sets can be established.

23. As described above, one of the major sources of multiple signatures will be variations in reflectance geometry. This fact can be exploited to assist in the assembly of signatures into groups or arrays that derive from a single terrain type (as perceived by the human analyst). When a terrain type such as a forest covers an entire hill, the CCT value sets deriving from that terrain type will include examples from all slope angles and aspects inherent in the relations among topographic configuration, sun zenith angle and azimuth, and sensor zenith angle and azimuth. These signatures thus form a graded sequence, with the constraint that the same reflectance geometry will always produce the same CCT value set. A different terrain type will also produce a graded sequence, but the CCT value set for a given reflectance geometry will be different from that deriving from the first terrain type.

24. In theory, the procedure would involve careful placing of the geographic locations of the various CCT value sets on a topographic map. Any specific topographic slope degree and aspect would then be selected,

and all CCT value sets recorded from each occurrence of the specified slope would be noted. If there are only two distinct terrain types, then there will be two CCT value sets in the pixels deriving from that specified slope. If there are three terrain types, then there may be three CCT value sets, but only if all three terrain types occur in all topographic expressions. In any event, systematic examination will result in the assembly of the 20 to 70 CCT value sets into a very significantly smaller number. Each aggregation of CCT values identified in this manner can be regarded as a terrain type, and at this point a map of the distribution of those terrain types (which are actually only aggregations of related CCT value sets) can be drawn. A minimal amount of field checking will now establish the relation between ground condition (i.e. terrain type) and aggregation of CCT values.

25. Unguided mode, second variant. Let it be assumed that the analyst has no immediate way to establish relations among CCT value sets and terrain types as they exist on the ground. Nevertheless, there are circumstances in which the analyst would like to map the distributions of the naturalistic CCT class sets. It should be recalled that the CCT class sets are based on all four spectral values, hence no one image (i.e. no one Landsat spectral band) can be used to properly identify a class set. Even color composite images are not particularly helpful, since the human eye is not a reliable spectrophotometer.

26. The "unguided mode, second variant" procedure begins with the selection of a number of sample sites, directly from a Landsat image. The sites selected would normally be in areas of homogeneity. That is, they would be in patches in which the gray level or tone of the image (i.e. the CCT value) was reasonably uniform in all four Landsat bands for at least several pixels in all directions. This minimizes (but does not entirely eliminate) the possibility of inadvertently selecting a spurious pixel.

27. The selected pixels would then be separated from the image array, and all four CCT values for each pixel assembled into a CCT value set. Since the various CCT value sets are all assumed to be members of naturalistic CCT class sets, the remaining problem is to determine the

compositions of the relevant CCT class sets. This can be achieved by any one of several forms of "cluster analysis." A rationale for such a procedure will be discussed in paragraphs 33-43.

28. The second variant is constrained by a critical problem: sample sites must represent all of the naturalistic classes of CCT values. If they do not, then many pixels that exhibit CCT value sets will not belong to any of the recognized terrain types and so cannot be classified. Thus, it seems better to err in the direction of abundance rather than to select too few and have residual unclassified areas.

29. Guided mode. The guided mode (i.e. in which the analyst has at least some prior knowledge of the terrain types in the region as well as knowledge of the locations of at least some characteristic examples of each) is a much more direct method of exploiting the naturalistic CCT value classes. This mode depends on the fact that the analyst can establish a direct relation among terrain types and CCT value sets or CCT class sets. In principle, a pixel falling inside the boundaries of a known example of a terrain type will exhibit a CCT value set that will be characteristic of one of the naturalistic CCT class sets, as defined by histograms of the type shown in Figure 2. For example, a CCT value set of 25, 19, 39, 30 (bands 4, 5, 6, and 7, respectively) would fall into CCT value classes (Table 1) 1, 2, 6, and 6. If the class boundaries actually represent realistic natural divisions among terrain types, the CCT value set is a member of a group of CCT value sets, each value of which may vary within the limits of the stated classes. That is, the value of band 4 may vary between 22 and 26, the value of band 5 between 18 and 21, etc.

30. When this has been done, several CCT value sets will be discovered that are actually members of the same population (or CCT class set), as defined by the naturalistic classes determined from the histograms. Furthermore, since the sample sites were selected because they represented the various terrain types of interest to the analyst, the naturalistic CCT class sets associated with each terrain type can be readily identified. For example, variations in reflectance geometry might be found to result in five or six naturalistic CCT class sets

being representative of the same desired terrain type. These CCT class sets can now be used as numerical criteria by which all other pixels in the scene can be quickly evaluated. In theory, every pixel in the scene, the CCT values of which fall within the stated limits, will be evidence of the presence of the terrain type on the ground. With the identification of all pixels meeting the specified criteria, a map showing their distribution can be drawn; and that map will, in effect, be a representation of the areal distribution of the specified terrain type.

31. A variant of this method assumes the existence of naturalistic CCT class sets but does not actually use the histograms. Instead, it assumes that the CCT value exhibited by a sample pixel is the modal value of a set of values that is normally distributed about that value. An examination of the sample histograms (Figure 2) and the naturalistic classes that can be derived from them reveals that the shapes of the distributions could, in most instances, be generated by normal distributions only four to six CCT values wide (i.e. two or three values on either side of the "peak"). Therefore, it is not entirely unreasonable to assume that the value of the sample point represents the modal value of a normal distribution with a standard deviation of 1.5. The values away from the mode represent system errors, as described in paragraph 12.

32. In effect, each CCT value set is assumed to define the modal values of a CCT class set. However, if for no other reasons than the system errors, CCT value sets from different sample areas within the same terrain type may exhibit slightly different CCT value sets. These closely related sets should clearly be assembled into an aggregation that is, collectively, a CCT class set. This involves another aspect of cluster analysis that is discussed in the following paragraphs.

Cluster analysis

33. To be of maximum utility in meeting the objective stated in paragraph 3, the cluster analysis procedure must be simple, rapid, and under operator control during all phases. The primary reason for the last requirement is that, in at least one context (see paragraphs 29-32), the analyst will be trying to "force" the CCT value sets into a very limited number of quite arbitrary groupings. For example, a region

might actually display 20 or more "real" CCT class sets, but the analyst might desire that his final product consist of a map with only three or four categories. It is obvious, in such a case, that each mapped category will contain several of the naturalistic CCT class sets.

34. In practice, a CCT value set obtained for a pixel that was selected as representing a particular terrain type cannot be assumed to consist of CCT values representing the modal values of the CCT class set to which it belongs, since instrument errors and other causes can alter the theoretical values to some degree. Further, the degree of variation cannot be known a priori, and thus the assumption of a fixed standard deviation (see paragraph 31) cannot be relied upon. Indeed, there is no reason to assume that all terrain types will exhibit the same amount of variability in their reflectance characteristics, and thus it may be confidently expected that some naturalistic CCT value classes will be broader than others. An examination of the subjectively selected class ranges identified in Table 1 supports this expectation; some classes are only two CCT values wide, and others are as much as nine.

35. Objective criteria could, in principle, be obtained by selecting many pixels representing the same terrain type and considering them to be members of the same population, with values arranged in a Gaussian distribution. The simplest method of achieving this would apparently be to select a sample pixel representing a known terrain type and then to enlarge the sample to include an array of pixels neighboring the sample pixel. If only the immediate neighbors are included, the combined sample would consist of a 3- by 3-pixel array, which is enough to provide a sufficient estimate of mode and standard deviation if the population is internally homogeneous. That is, the assumption is valid only if the variability results from such sources as instrument errors, small and randomized variations in topographic slope, etc. If the sample areas are chosen with care, the assumption is likely to be acceptable.

36. However, it is not necessary to proceed entirely on faith. If the sample 3- by 3-pixel array is inadvertently positioned such that some of the pixels are in one terrain type and the remainder in another,

the CCT values of at least one of the spectral bands will show either a bimodal distribution or a strongly skewed distribution. Thus, the obvious step would be to subject the CCT values from each wavelength band to a normalcy analysis. If the distributions of values in all wavelength bands prove to be essentially Gaussian (i.e. normal), then it may be safely assumed that all nine pixels represent a single terrain type.

37. If several sample 3- by 3-pixel arrays were selected in the same terrain type, ideal theory would dictate that the CCT class sets, as defined by the standard deviations of the 3- by 3-pixel samples, should overlap perfectly. However, this cannot be expected in practice, because of small local variations in ground conditions, random instrument error, and the like. In view of the practical realities, similarity was defined to exist between two CCT class sets if all CCT value classes in the two CCT class sets overlap. For example, suppose that three 3- by 3-pixel arrays have been selected in a region and that the Landsat band-4 mean and standard deviations and CCT class set limits are as follows:

Site No.	Mean	Standard Deviation	CCT Class Set Limits
1	25.0	0.8	24.2-25.8
2	25.4	0.9	24.5-26.3
3	32.0	1.7	30.3-33.7

On the basis of the criteria for similarity previously defined, sites 1 and 2 are similar, because their CCT class sets overlap; but site 3 is not similar and, therefore, is presumably a member of a different population. That is, it is presumably representative of a terrain type other than that designated by sites 1 and 2.

38. It should be noted that the establishment of similarity between two sites may result in a broadening of the effective CCT class set. For example, if only site 1 in the tabulation in the previous paragraph, is considered, the CCT class set limits for the terrain type represented by site 1 are 24.2-25.8. However, since the CCT class set

limits of site 2 overlap, the two sites are defined as denoting the same terrain type. This means that the CCT class set limits representing that terrain type have been broadened to include both sites 1 and 2 and are, therefore, 24.2-26.3.

39. The implications of this are important. First, consider a natural slope smoothly curved such that a part of it is at right angles to incident solar rays and thus is illuminated at maximum intensity, while other parts are at progressively greater angles until the slope is parallel to the incident rays. The result is that the reflected radiation will exhibit a smooth continuum of values from maximum to minimum in all wavelength bands. However, the radiance values in each band will, for all practical purposes, retain the same relative ratios with respect to each other. A pixel grid placed over the slope will result in a series of CCT value sets that are generically related. If these CCT value sets are treated in the manner described in paragraphs 36-38, the product is likely to be a series of CCT class sets that sequentially overlap and form a related series, all of which represent the same terrain type. Thus, to some degree, the vagaries of reflectance geometry are compensated for by the method of defining similarity.

40. The same situation will be true for other terrain types, of course, but it must be recalled that it is the relative ratios of radiance values that determine the uniqueness of each terrain type. This attribute should allow a number of terrain types to be differentiated, more or less independently of reflectance geometry.

41. The second important implication is that it is entirely possible to have three CCT class sets, in which set 1 is related to set 2 and set 2 to set 3, but in which an independent comparison of sets 1 and 3 would not indicate a similarity. For example, suppose the analyst selected four sample sites in a terrain type and obtained CCT class sets for each (see paragraphs 36-38). Figure 3 illustrates a possible combination. If a "similarity" matrix is assembled (Figure 3), it will be noted that CCT class set 1 is similar to sets 2, 3, and 4, but that set 2 is not similar to set 3. The reason is that the CCT class sets in band 6 do not overlap in sets 2 and 3. Despite the fact that sets 2 and

Set No.	CCT Class Sets			
	Band 4	Band 5	Band 6	Band 7
1	24.2-25.8	18.3-20.3	38.2-39.3	30.0-32.8
2	24.8-26.7	18.6-20.1	39.0-40.4	30.0-33.0
3	23.6-25.0	17.8-19.2	36.8-38.8	29.2-31.4
4	24.5-26.7	18.1-19.8	37.2-39.6	30.1-32.7

Similarity Matrix

	1	2	3	4
1	X	X	X	X
2	X	X		X
3	X		X	X
4	X	X	X	X

Note: X indicates similarity

"Stretched" CCT Value Set Ranges

Band 4	24.2-26.7
Band 5	18.1-20.3
Band 6	36.8-40.4
Band 7	29.2-33.0

Figure 3. Stretched CCT value ranges for a similarity matrix

3 are not similar to each other on a one-to-one basis, they are defined in this study as similar by virtue of their mutual relationship to the entire similarity matrix. To express analytically the type of similarity shared by all four sample sites, the CCT value ranges of each site must be extended or stretched in each band to the same limits. The resulting stretched ranges (Figure 3) then represent the CCT class set of the similarity matrix as a whole.

42. Finally, it will be recalled that the limits of the CCT class sets were defined arbitrarily as being one standard deviation away from the mean on either side. There is, of course, no reason in physics why that value should possess any intrinsic merit. If the class ranges (or boundaries) are set farther from the mean, the result will obviously be that more CCT class sets will meet the similarity criteria. Thus, adjustments of the position of the boundaries of the CCT class sets will result in various degrees of "tuning" of the similarity categories.

Practice

43. The following discussion of the semiautomated procedure is

based on the section on theory (see paragraphs 8-42) and on ground truth data obtained in an 11- by 14-km study area centered approximately 40 km northeast of Vicksburg, Mississippi (Figure 4). The study area was selected to include a representative section of the loess hills, forming the eastern wall of the Mississippi floodplain, and a section of the floodplain, including an oxbow lake and a number of other floodplain features. Figure 5 is a mosaic made of conventional panchromatic air photos of this study area. The soil type categories were transferred to the photograph from the U. S. Department of Agriculture (USDA) Soil Conservation Service Soil Survey of Yazoo County[5] (see paragraph 50).

44. Plate 1 is a generalized flow diagram of the steps involved in the procedure described below (see paragraphs 62-107 for a detailed discussion).

45. Since the study area undergoes profound seasonal changes and a general procedure that would be workable in any season was sought, the intent was to use one data set (specifically the 13 October 1975 data) to develop the procedure and then to apply that procedure rigorously to the study area as represented by two additional Landsat images, one in summer (11 July 1974) and one in winter (21 February 1975). The July 1974 imagery data were obtained with Landsat 1, and the February and October 1975 images with Landsat 2. All three scenes were carefully selected to be free of both haze and clouds, thus providing optimum CCT values. Table 2 gives the characteristics of the imagery data. Note that all three data sets were of approximately the same overall quality. The only obvious differences in the data sets are in terms of solar zenith angle and azimuth. Figure 6 is an image formed from the October 1975 band-7 (0.8-1.1 μm) data, indicating the location of the study area.

Description of study area

46. The study area (Figure 5) contained several different terrain types, including agricultural lands (plowed fields, fields with crops at different stages of growth, fallow land, etc.), forested land (deciduous and mixed evergreen trees), grasslands (summer and winter pasture grasses, fields with volunteer grasses and other weedy plants), water features (rivers, lakes, canals, wetlands), and some low topographic

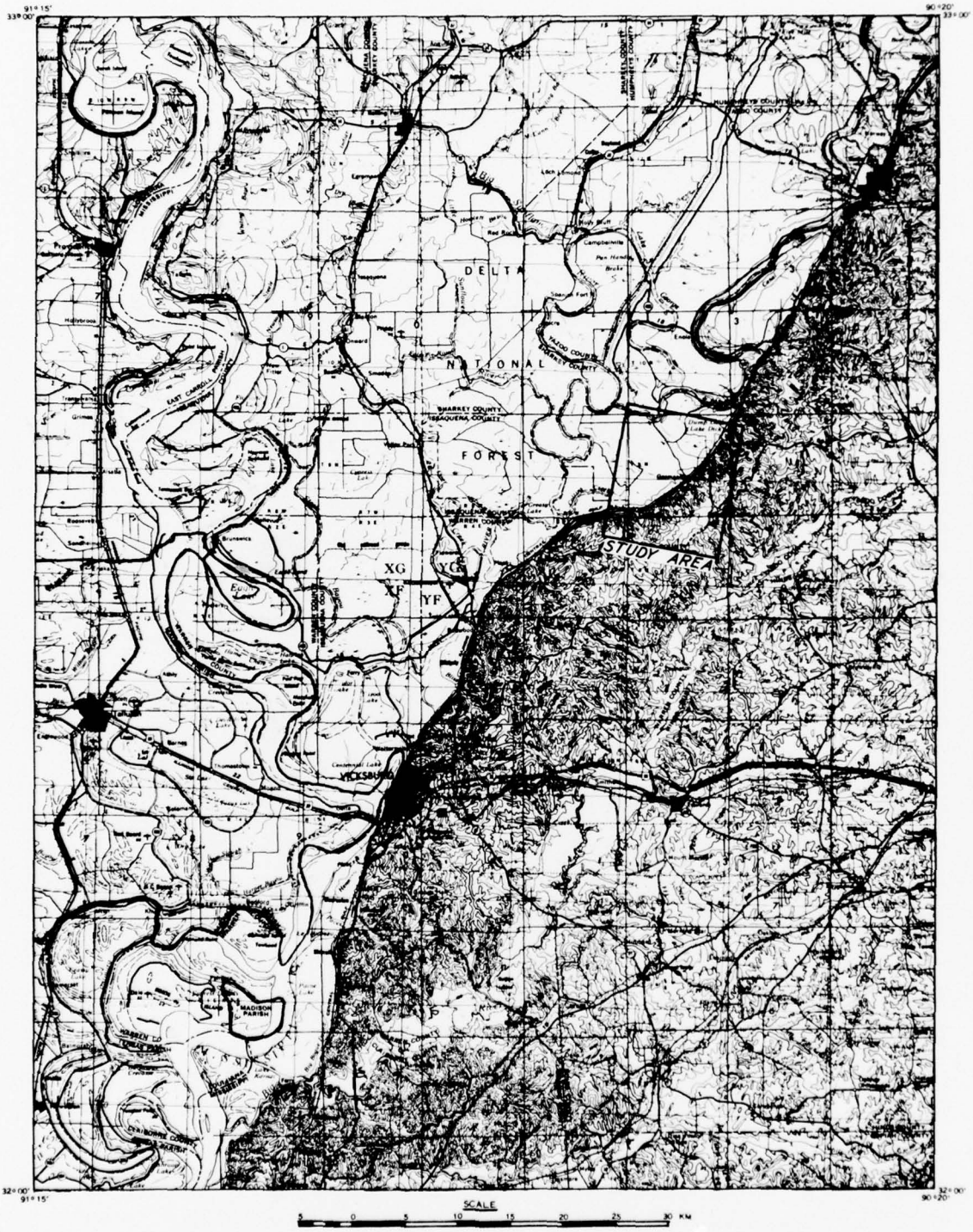


Figure 4. Regional map showing location of study area

LEGEND

USGS SOIL SYMBOLS	SOIL TYPE
ML	SILTY LOAM
CL	SILTY CLAY LOAM
CH	CLAY

Figure 5. Air-photo mosaic of study area with soil types delineated

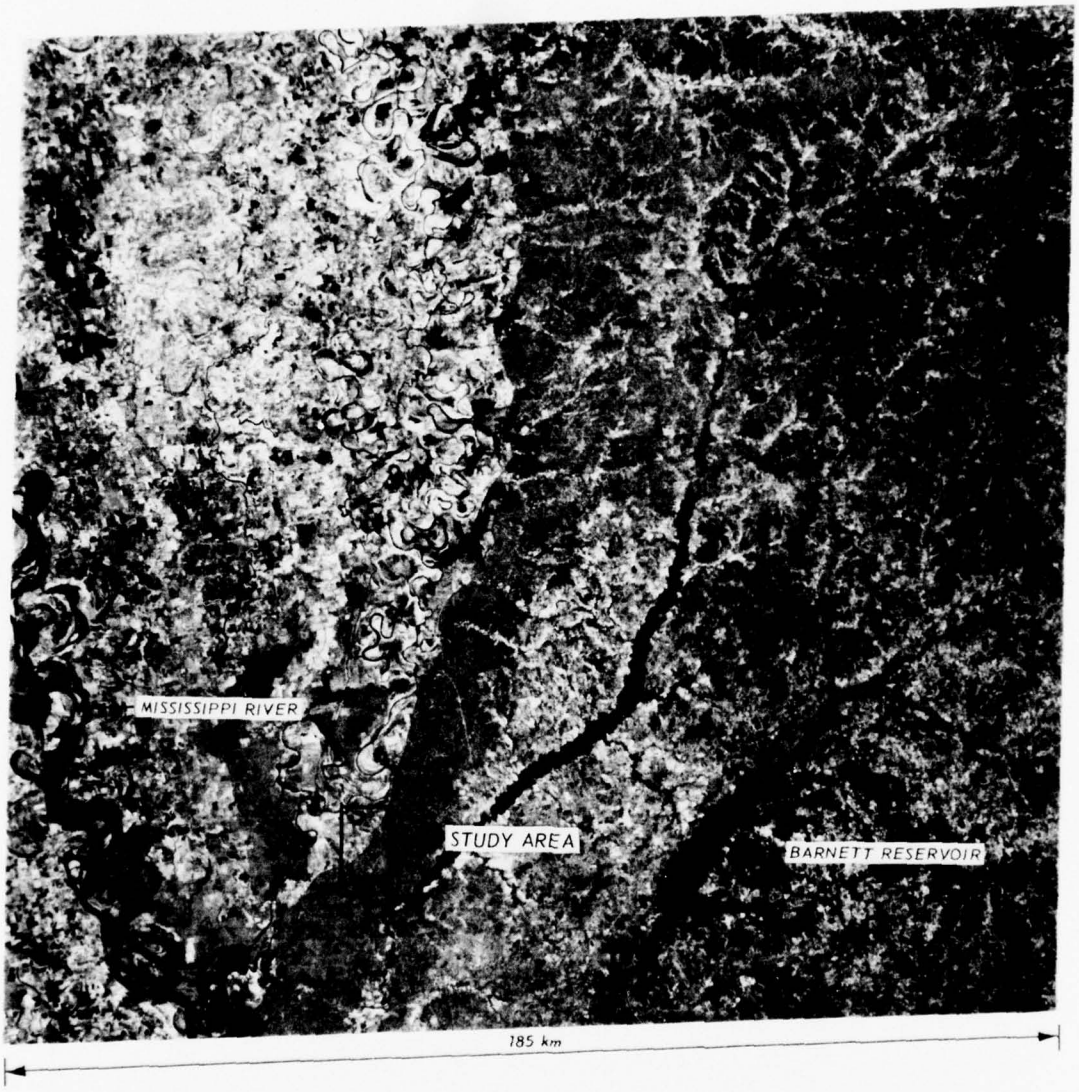


Figure 6. Image provided by the EROS Data Center constructed from band-7 data showing location of study area; Landsat scene 2264-15540, 13 October 1975

areas that are frequently subjected to backwater flooding. It covers small sections of two parts of the Coastal Plain physiographic province: (a) the Mississippi Alluvial Plain and (b) the East Gulf Coastal Plain.[6] The alluvial soils are primarily silts (ML, CL) and clays (CH) with some sand; whereas, the soils in the East Gulf Coastal Plain are primarily windblown silts (ML) with subordinate amounts of very fine sand and clay (i.e. loess material).

47. The vegetation in the area consists of crops of soybeans, cotton, and bermuda and rye pasture grasses; the forested areas contain different plant densities of oak, gum, hickory, poplar, pine, cedar, elm, pecan, and cypress.

48. The relative effects of the different terrain and climatic factors on the transfer of incident radiant energy to energy that has been emitted or reflected by the in situ terrain materials (soil, rock, vegetation material, and water) and man-made features (paved areas, buildings, etc.) have not been quantitatively determined. Nevertheless, there are some basic factors that almost surely act together in some fashion and control this transfer process. These factors are extremely critical to an overall procedure that will allow interpretation of terrain conditions for different seasonal and climatic conditions.

49. Since terrain conditions change appreciably with time and are affected significantly by local climatic conditions, it was necessary to quantitatively describe the study area at the time of each respective overflight. Data pertaining to soil, topographic and flood conditions, rainfall, vegetation, and other features were obtained and mapped at an appropriate scale for comparison with the conditions being interpreted from the Landsat spectral data. The mapped data were portrayed on a photomosaic at a scale of 1:4000 (Figure 5) that was made up from 1973 air photos obtained by the USDA.

50. Soil conditions. The soils within the study area were mapped by the USDA Soil Conservation Service[5] and consist of silty loams (ML), silty clay loams (CL), and clays (CH). A soil map for the study area is superimposed on the photomosaic in Figure 5. The soils in the alluvial plain (or cropland) are composed of patches of ML, CL, and CH; whereas,

the soils in the upland forested areas are primarily ML's.

51. Topographic and flood conditions. The relief within the study area, as portrayed on a U. S. Geological Survey (USGS) 1:62,500-scale map with 5-ft (1.52 m) contours, varies from about 85-115 ft (25.9-35.1 m) above mean sea level (msl) within the cropland or agriculture region to about 115-130 ft (35.1-39.6 m) within the upland hills that occupy the remaining portion of the study area.

52. The study area borders on the Yazoo River, a large tributary of the Mississippi River, and a portion of the agricultural area is subjected to flooding during high river stages. The Corps of Engineers stage records for the Yazoo River gage at Satartia, Mississippi, indicate that flooding does occur within adjacent low-lying areas when the stage of the river reaches approximately 95 ft (29 m) msl. The river stage data from the Satartia gage for the period 10 January 1974 through 30 October 1975 (Figure 7) indicate that flooding most likely occurred within parts of the study area prior to and during two of the three overflight times (11 July 1974 and 21 February 1975).

53. Since the Yazoo River at Satartia did attain stages that would permit flooding (Figure 7) in low-lying areas during two of the three overflight times, it was anticipated that a detailed topographic map was needed in the event that it became necessary to delineate low-elevation areas within the study area. The best available map was the USGS 1:62,500-scale topographic map. The topographic map was constructed from 1963 aerial photography and does not represent the small changes and alterations that have been made to the area to make the low-lying areas more useful for agricultural purposes. Ditches and canals have been constructed within the area (Figure 8) to provide for drainage during wet periods and irrigation during dry periods. Figure 9 shows the section of the map covering the study area.

54. Rainfall. The daily rainfall records from Germania, Mississippi (site in the middle of study area, Figure 9), show that the area received some rainfall within 5-6 days of the July 1974 and February 1975 satellite overflights, and no rainfall within 13 days prior to the October 1975 overflight (Figure 10). This record is believed to be

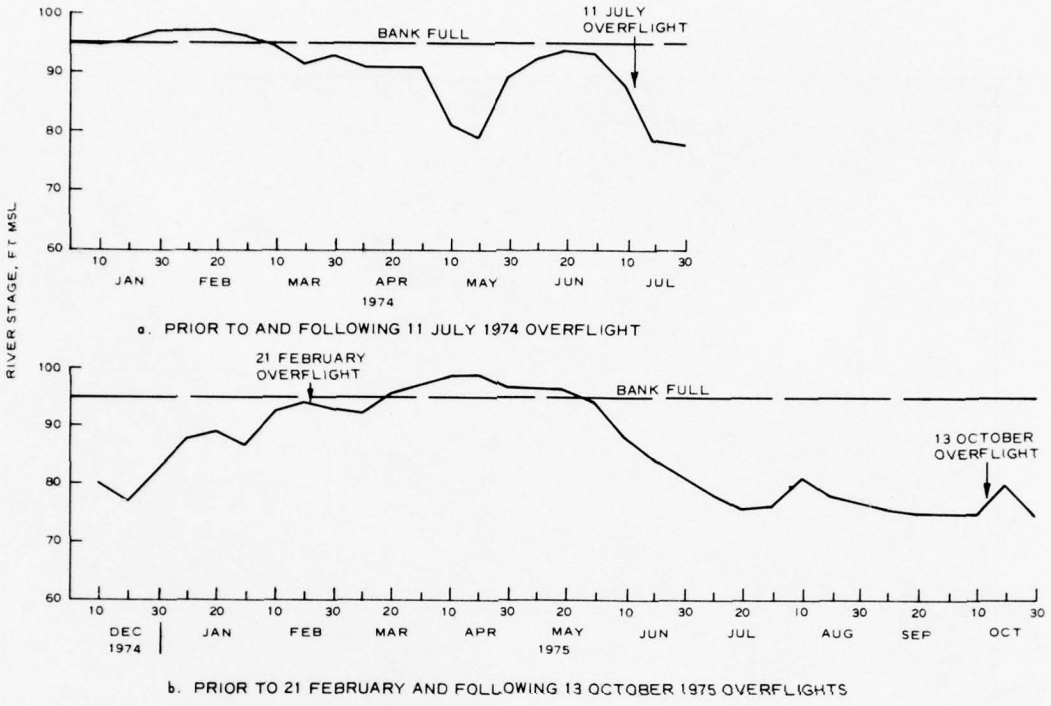


Figure 7. Yazoo River stages at Satartia, Mississippi

a. Small lake surrounded by 20-m tall deciduous trees

b. Drainage and irrigation canals

Figure 8. Canals and ditches in the study area, 8 May 1976

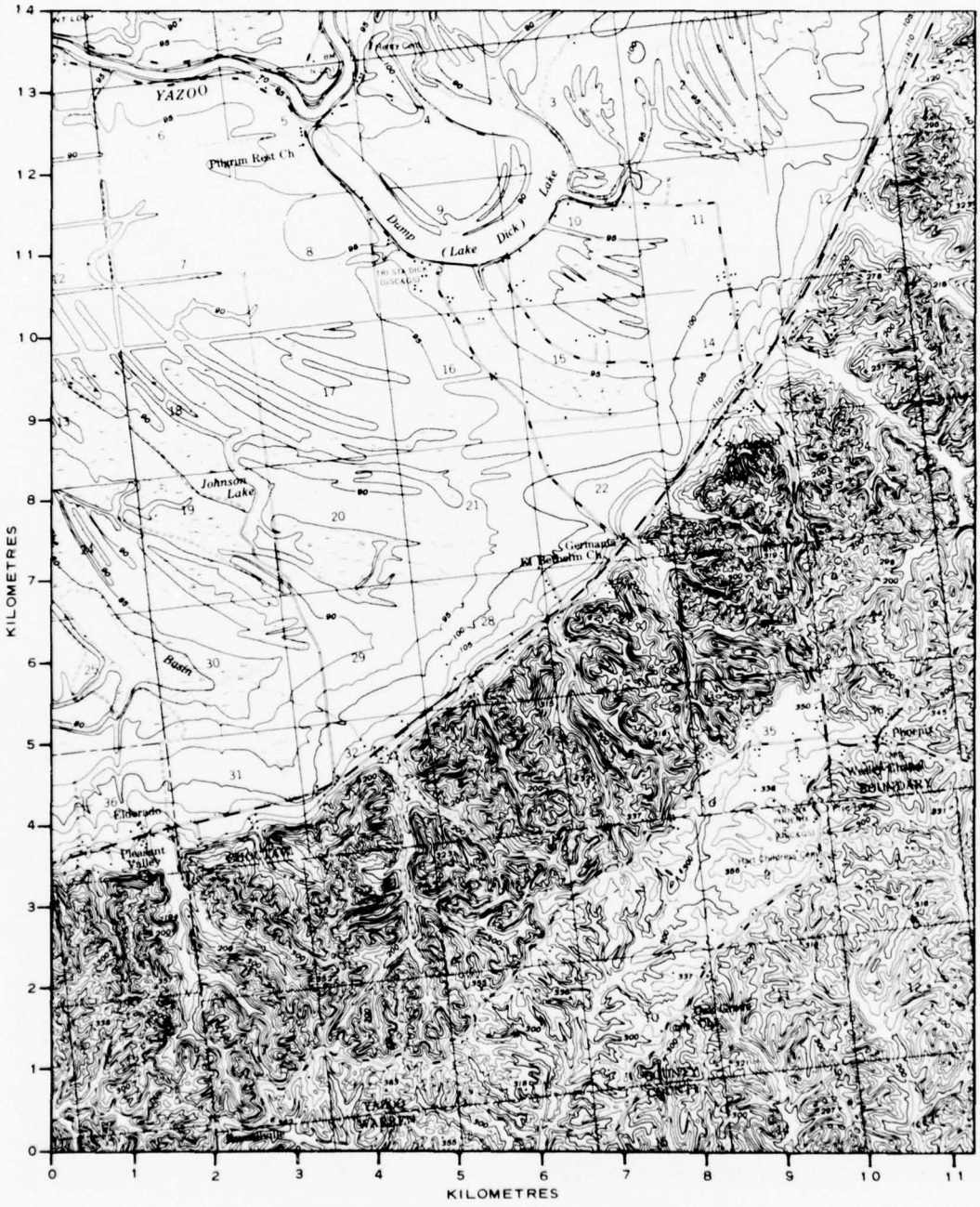
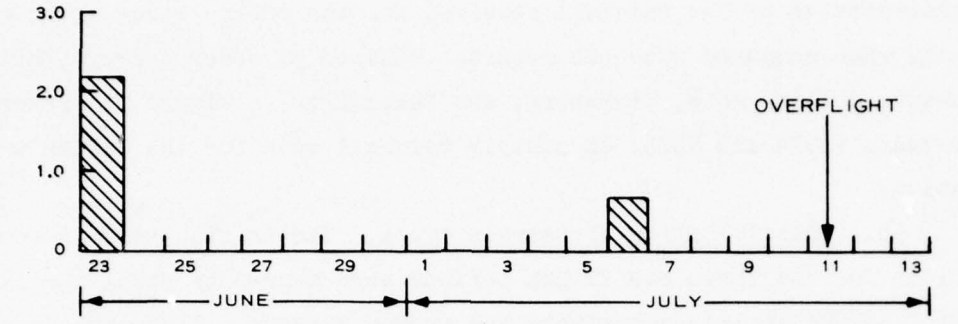
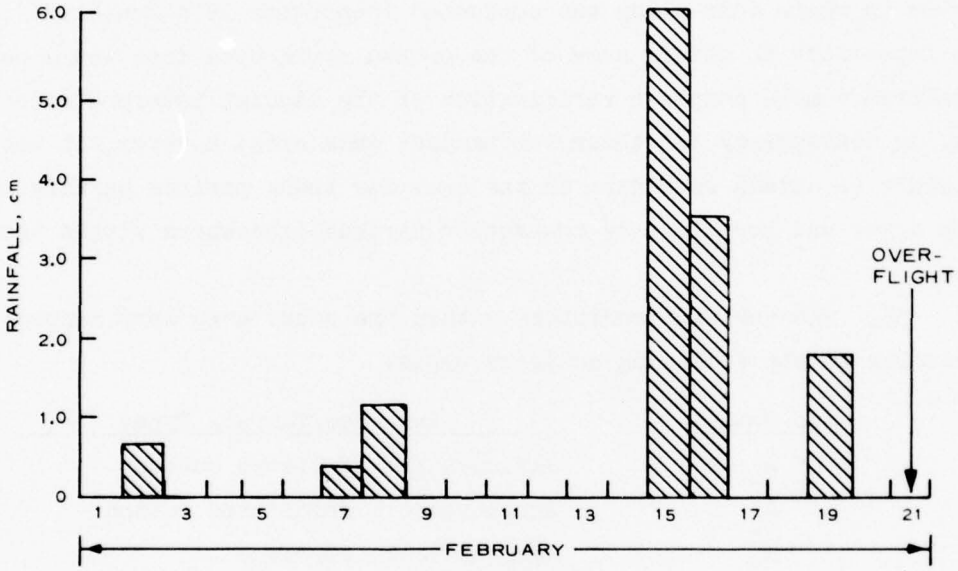


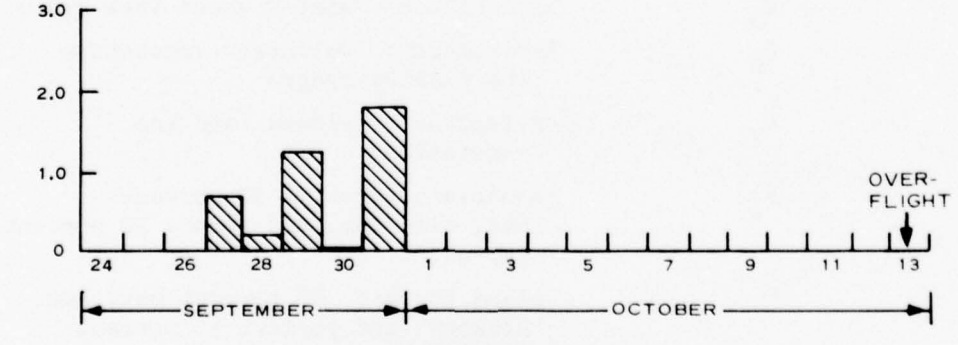
Figure 9. Section of USGS 1:62,500-scale topographic map



a. 11 JULY 1974



b. 21 FEBRUARY 1975



c. 13 OCTOBER 1975

Figure 10. Rainfall received at study site prior to the three satellite overflights

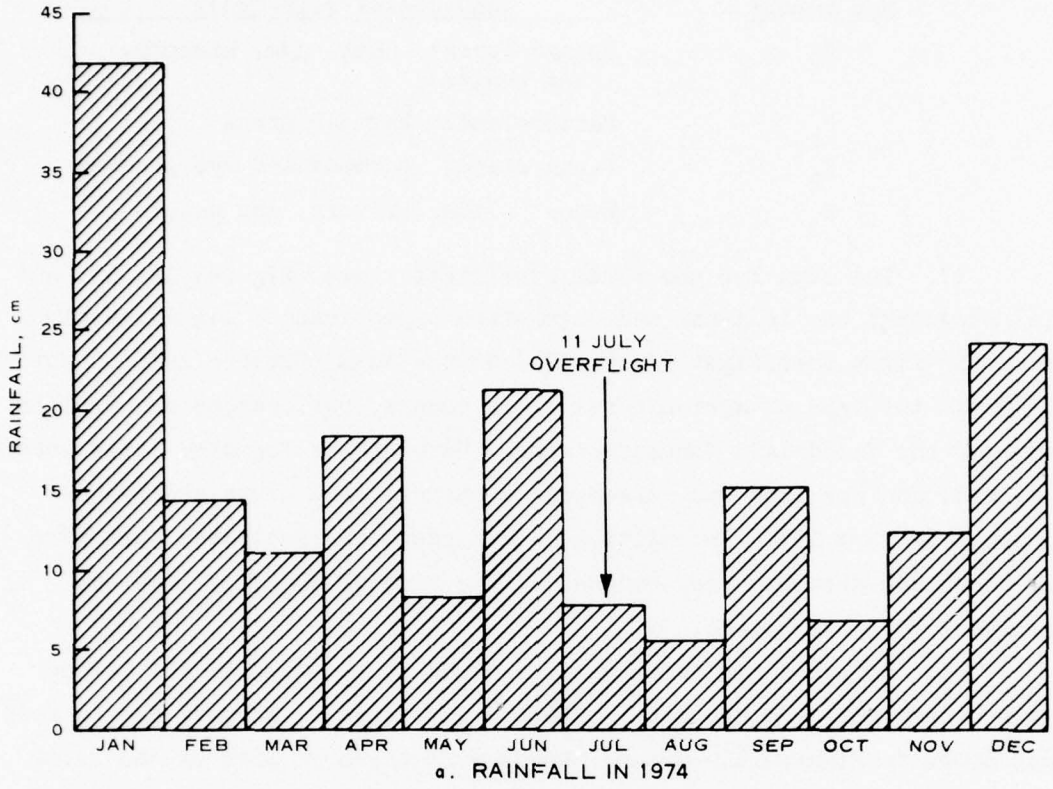
representative of the rainfall received for the entire study area, especially when compared with the records obtained at other nearby stations (Onward, Rolling Fork, Vicksburg, and Yazoo City). Figure 11 presents two years (1974 and 1975) of monthly rainfall data for the Germania station.

55. Classification of terrain types. The in situ terrain conditions for the three overflight periods were mapped by using the 1973 USDA 1:10,000 aerial photography and ground surveys. Since two of the overflights (July 1974 and February 1975) occurred before the time period in which this study was conducted (September 1975-June 1976), it was impossible to obtain some of the ground truth data that would have permitted a more complete verification of the Landsat interpretation and mapping methodology for those two Landsat data sets; however, it was possible to obtain some data on the area for these periods pertaining to crop types and locations by contacting various landowners within the area.

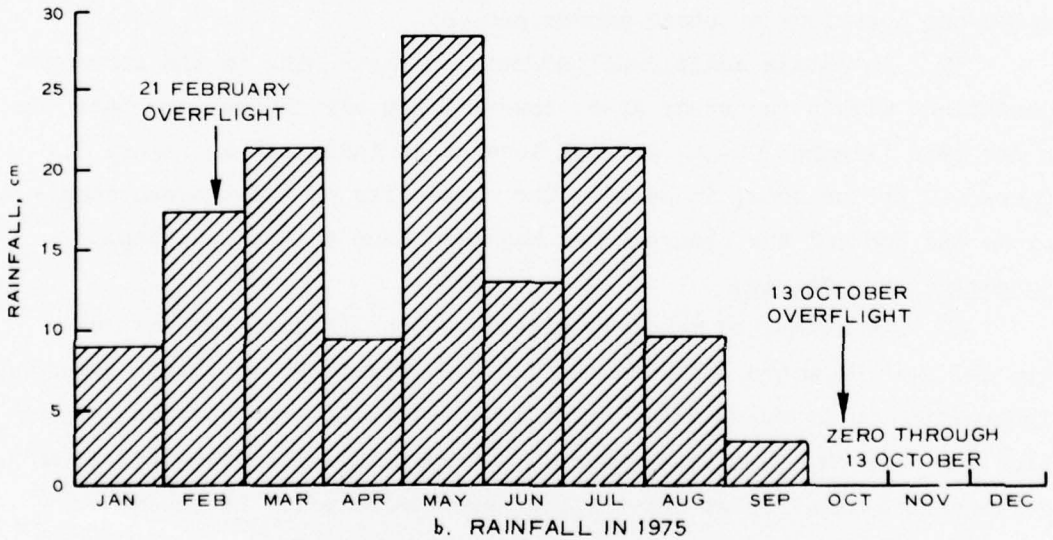
56. The terrain conditions within the study area were mapped according to the following category types:

Map Symbol	Land-Use/Terrain Types
A_1	Agriculture: foliated cotton
A_2	Agriculture: defoliated cotton
A_3	Agriculture: soybeans
A_4	Agriculture: oats
A_5	Agriculture: winter wheat (rye grass)
A_6	Agriculture: volunteer vegetation (no planted crops)
A_7	Agriculture: plowed land (no vegetation)
F_1	Floodplain forest: 90 percent oak, elm, gum, and pecan; 10 percent cypress
F_2	Upland forest: 85 percent oak, gum, hickory, and poplar; 15 percent pine and cedar

(Continued)



a. RAINFALL IN 1974



b. RAINFALL IN 1975

Figure 11. Monthly rainfall data for the Germania station from January 1974 through December 1975

Map Symbol	Land-Use/Terrain Types
F_3	Upland forest: oak, gum, hickory, and poplar
P_1	Pastureland: bermuda grass
P_2	Pastureland: bermuda and rye grasses
W	Water: lakes, rivers, and canals

57. The maps for the three overflight times (Figures 12, 13, and 14) show that the land-use and vegetation types changed significantly from the first overflight (July 1974) to the last (October 1975). Not only did the type of agricultural crops change, but changes occurred in some of the floodplain forested areas. Some of the forested areas were cleared, and the land was converted to agricultural crops and pasture grasses (Figure 15). In addition, local residents indicated that some of the study area had been flooded at the time of the July 1974 and February 1975 overpasses.

58. Detailed fieldwork was performed at the time of the October 1975 overflight, and the data collected were used to prepare the terrain map shown in Figure 14. Ground photos were taken of most of the sites at the same time; Figures 16-20 present in detail some of the significant terrain types that occurred within the study area. Figure 21 indicates the locations of these ground photos.

59. To obtain additional information pertaining to the terrain conditions within the study area, low-altitude air photos were taken on 8 May 1976 (Figures 22-30) and are located as indicated in Figure 31. These air photos show, in detail, the complexity of the terrain conditions and some of the changes that had been made to the landscape.

Automated data station

60. An automated meteorological station[7] (Figure 32) was installed near Germania (coordinates 7.8-3.0, Figure 9) to obtain detailed information on incoming solar radiation, rainfall, wind speed and direction, and air and soil temperatures. The original intention had been to collect such data for at least a two-week period prior to the October 1975 Landsat overpass. However, the equipment could not be made available at that time, and thus the actual period of record extended from

LEGEND

A_1 FOLIATED COTTON
A_3 SOYBEANS
A_7 PLOWED LAND (NO VEGETATION)
F_1 90% OAK, ELM, GUM, AND PECAN; 10% CYPRESS
F_2 85% OAK, GUM, HICKORY, AND POPLAR; 15% PINE AND CEDAR
F_3 OAK, GUM, HICKORY, AND POPLAR
P_1 BERMUDA GRASS
P_2 BERMUDA AND RYE GRASSES
W WATER

Figure 12. Terrain types on 11 July 1974 as determined by ground truth data and air-photo interpretation

LEGEND

A_4 OATS
A_5 WHEAT AND RYE GRASSES
A_6 VOLUNTEER VEGETATION
F_1 90% OAK, ELM, GUM, AND PECAN; 10% CYPRESS
F_2 85% OAK, GUM, HICKORY, AND POPLAR; 15% PINE AND CEDAR
F_3 OAK, GUM, HICKORY, AND POPLAR
P_1 BERMUDA GRASS
P_2 BERMUDA AND RYE GRASSES
W WATER

Figure 13. Terrain types on 21 February 1975 as determined by ground truth data and air-photo interpretation

LEGEND

A_1 FOLIATED COTTON
A_2 DEFOLIATED COTTON
A_3 SOYBEANS
A_7 PLOWED LAND (NO VEGETATION)
F_1 90% OAK, ELM, GUM, AND PECAN; 10% CYPRESS
F_2 85% OAK, GUM, HICKORY, AND POPLAR; 15% PINE AND CEDAR
F_3 OAK, GUM, HICKORY, AND POPLAR
P_1 BERMUDA GRASS
P_2 BERMUDA AND RYE GRASSES
W WATER

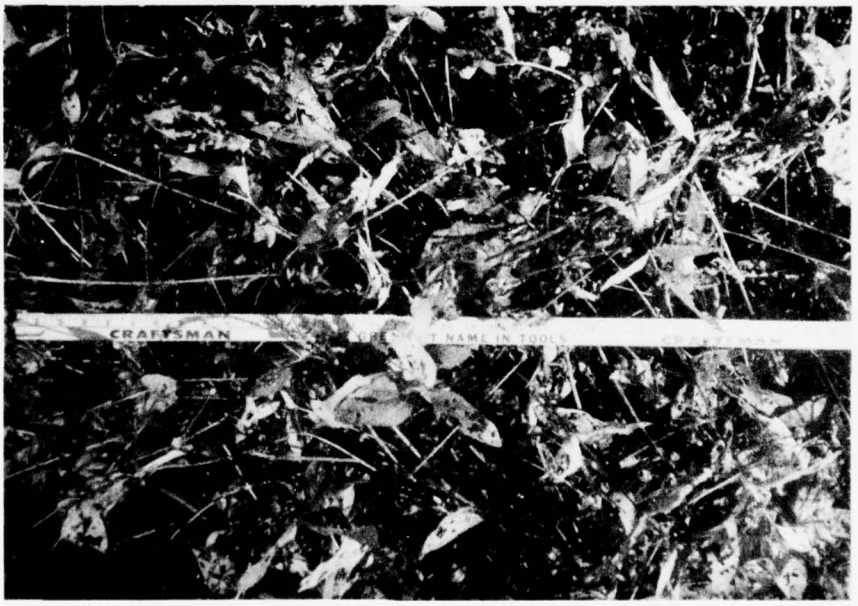
Figure 14. Terrain types on 13 October 1975 as determined from ground truth data and air-photo interpretation

a. Pastureland

b. Cropland

Figure 15. Areas shown with forest removed (since air photos dated July 1973) are currently used (8 May 1976) as pastureland and cropland on the southeast corner and eastern edge, respectively, of the large forested area on east side of Dump Lake (Figure 14)

a. General view



b. Close-up of plants

Figure 16. Mature soybeans, terrain type A3, 13 October 1975

a. General view

b. Close-up of leafy surfaces

Figure 17. Mature foliated cotton, terrain type A_1, 13 October 1975

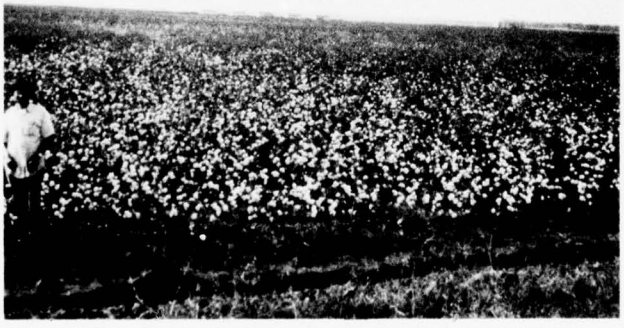
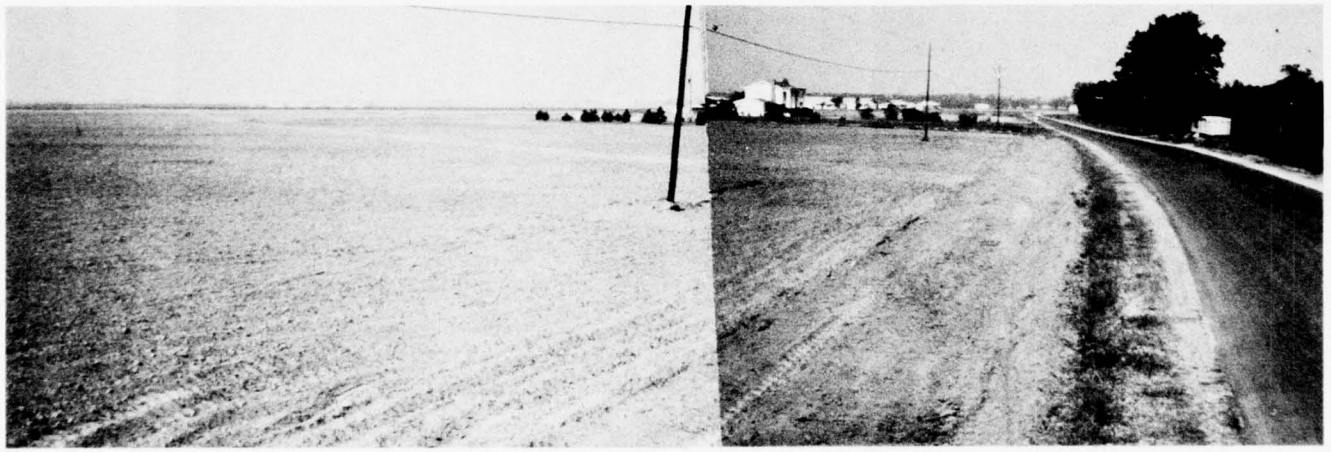


Figure 18. Defoliated cotton, terrain type A_2, 13 October 1975

a.



b.

Figure 19. Recently plowed fields, terrain type A_7, 13 October 1975

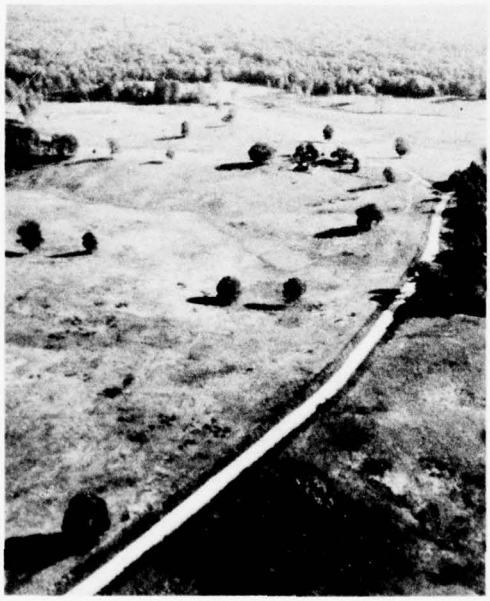
a. Recently mowed

b. Unmowed

Figure 20. Pastureland bermuda grass, terrain type P_1, 13 October 1975

Figure 21. Location and direction of view of ground photos in Figures 15-20

a.



b.

Figure 22. Pastureland (circular field in top photo), terrain type P_1
8 May 1976

Figure 23. Study area as viewed southwest of the forested area southeast of Dump Lake, 8 May 1976

Figure 24. Study area as viewed south-southeast of forested area surrounded by cropland, 8 May 1976

Figure 25. Study area as viewed east-southeast of Yazoo River and forest that lines the banks, 8 May 1976

Figure 26. Study area as viewed northeast of forest areas east side of Dump Lake, 8 May 1976

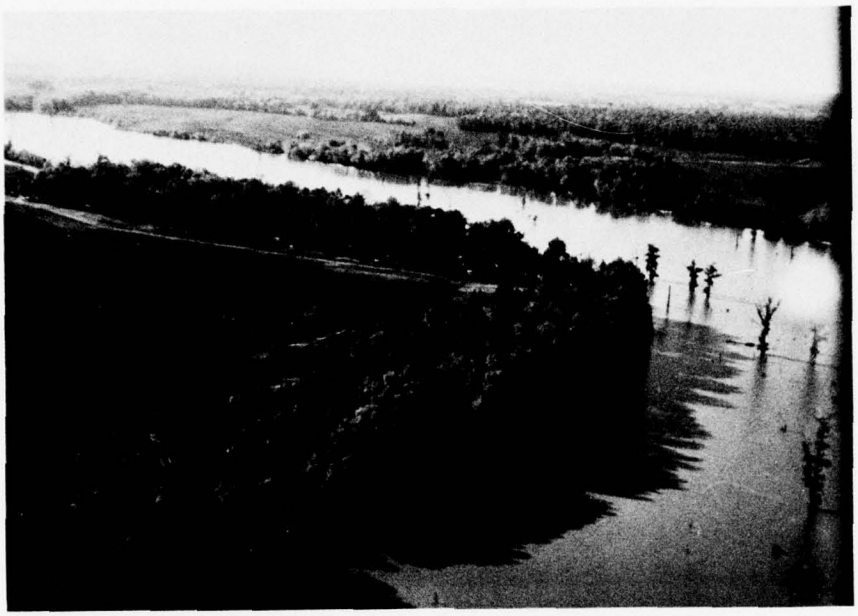


Figure 27. Study area as viewed northwest of Dump Lake and bordering forest, 8 May 1976

Figure 28. Study area as viewed northeast of complex of planted fields, pastureland, and forest, 8 May 1976

Figure 29. Study area as viewed northeast of pastureland divided by lines of trees along fence rows and drainage ditches, 8 May 1976

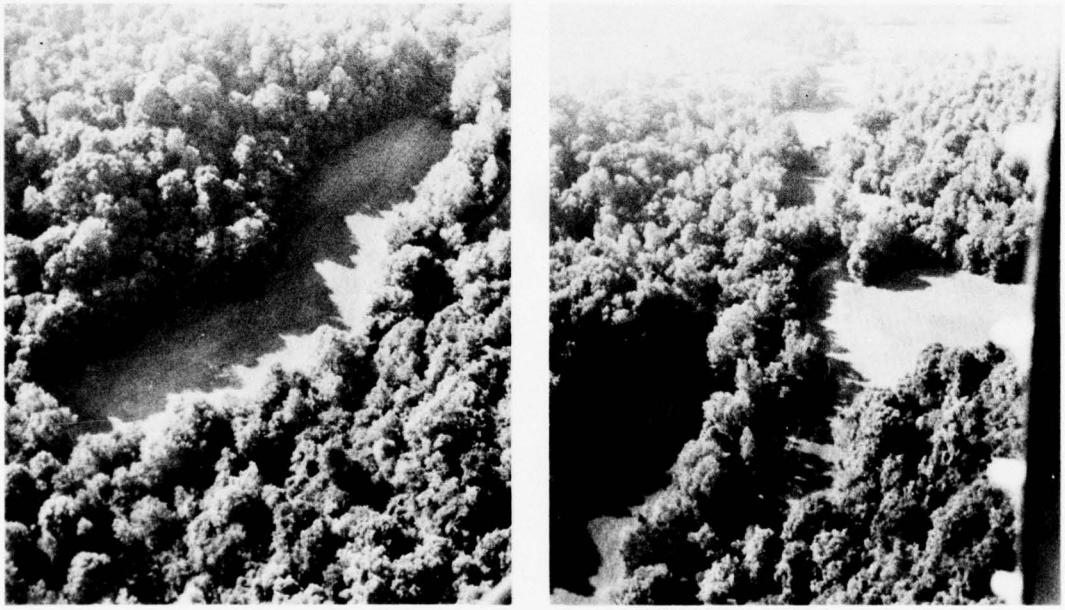


Figure 30. Study area as viewed north-northwest of small open areas (cropland) surrounded by 30-m tall deciduous forest, 8 May 1976 at 1700 hr; note effects of shadows

Figure 31. Locations and directions of view of air photos taken 8 May 1976 (Figures 22-30)

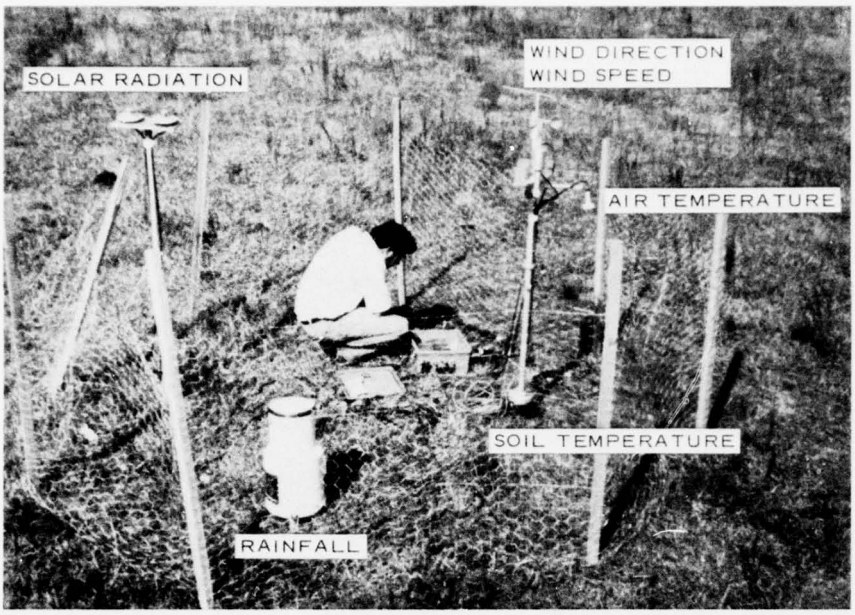


Figure 32. Automated meteorological field station as installed in study area (coordinates 7.8-3.0, Figure 5), 28 October-18 November 1975

28 October to 18 November 1975. While the data obtained are not useful for determining such conditions as soil moisture at the time of overpass, they were nevertheless helpful as a general guide to meteorological characteristics of the study area during the autumn.

Basic procedure for
guided classification

61. The actual procedure used to obtain an interpretation of Landsat imagery data was an interactive process, in which the human analyst performed a number of nonquantifiable operations from which quantitative instructions for an assortment of computer programs were obtained. The programs then very rapidly performed a number of relatively simple bookkeeping operations on the very large number of CCT values comprising a Landsat image data set. This cycle of human interpretation and machine operation was repeated several times. The general scheme is illustrated as a flow diagram (Plates 1 and 2). Individual segments are shown as "blocks" and numbered sequentially.

62. Block 1. The first step was to obtain from the EROS Data

Center (see paragraph 2) the CCT's for each of the three Landsat overpass dates selected for study (Table 2). The data for each overpass consisted of a pair of 9-track digital magnetic tapes containing multiplexed CCT values. Two tapes together comprised a Landsat scene that covered an area approximately 185 km square.

63. Block 2. The study area selected was smaller than a complete scene; therefore, some preprocessing of the CCT data was required. The problem was to extract the data for the study area from the data array covering the entire scene. In principle, it should have been possible to simply locate the study area carefully on a NASA-provided 1:1,000,000 image (one of the user product options available from the EROS Data Center) and then the corner coordinates by measuring in from the edges of the image. Unfortunately, this procedure resulted in substantial error. One of the reasons for the error was that the EROS-provided image was cropped so that it covered somewhat less area than the CCT data array. Since the amount of cropping is variable, the margins of the EROS-provided images cannot be used to precisely fix the location of a point in the pixel array.

64. The procedure followed was to write an image of the scene, using band-7 data (because of the clarity with which water surfaces are evident), and simultaneously making approximate corrections for earth rotation effects and pixel shape, as discussed in paragraphs 103-105 and in Reference 3.

65. Blocks 3 and 4. The image produced after making these approximate corrections is sufficiently close to the geometrics of conventional map projections that it was easy to locate, with only a small margin of error, the position of the study area. Since the pixel size used to produce the image was known (50 μm square), and the edges of the image represented the true edges of the CCT data array, it was easy to measure from the image margins and thus approximately fix the locations of the corner coordinates of the study area in terms of pixel and scan line numbers.

66. Block 5. The scan line and pixel numbers of the corner coordinates were used as the basis for extracting the block of pixels

representing the study area from the complete CCT data array. The extracted data array was then copied onto a new tape, which was called SA/CCT (study area CCT).

67. Block 6. The next step in the procedure was to write an image, at an enlarged scale, of the study area, using selected wavelength bands. For the current experiment, band-7 data were chosen, chiefly because water features show so distinctly. The importance of this will become apparent later in the discussion. In the current exercise, conventional air photos at a scale of approximately 1:71,700 were used for certain aspects of ground truth, and so the images were written at approximately that scale.

68. Block 7. Since a CCT value was to be correlated with a point on the ground (see paragraphs 29 and 30), the location of the point (i.e. the pixel representing that point) had to be precisely known. An error in location of several pixel dimensions (a Landsat pixel is a rectangle approximately 79.1 m long in a direction parallel to the orbital path of the satellite and 57.2 m wide in a direction at right angles to the orbital path) could have resulted in an attempt to correlate a CCT value set with the wrong terrain type. The implication of this was that the SA/CCT had to be corrected to the geometry of an air photo that was used as the basis for selecting sample items (see paragraph 29). The success of the procedure for achieving geometric accordance depended on the existence of common "transfer points."

69. A critical problem was the selection of the transfer points. They had to be such that they could be located in the Landsat image or data set regardless of the season at which the data were obtained. In the current instance, data for three widely different seasons (summer, fall, and winter) had to be used. It must be recalled that the transfer points must also be clearly recognizable on the air photo as well as in the Landsat data. The following features usually provide good transfer transfer points:

a. A small (approximately 2.5-ha) roughly circular lake surrounded by dry land or contiguous vegetation.

b. Intersection of two large (i.e. 50-m wide) streams.

c. The junction of a large stream and a lake.

d. Intersection of two wide roads or airfield runways.

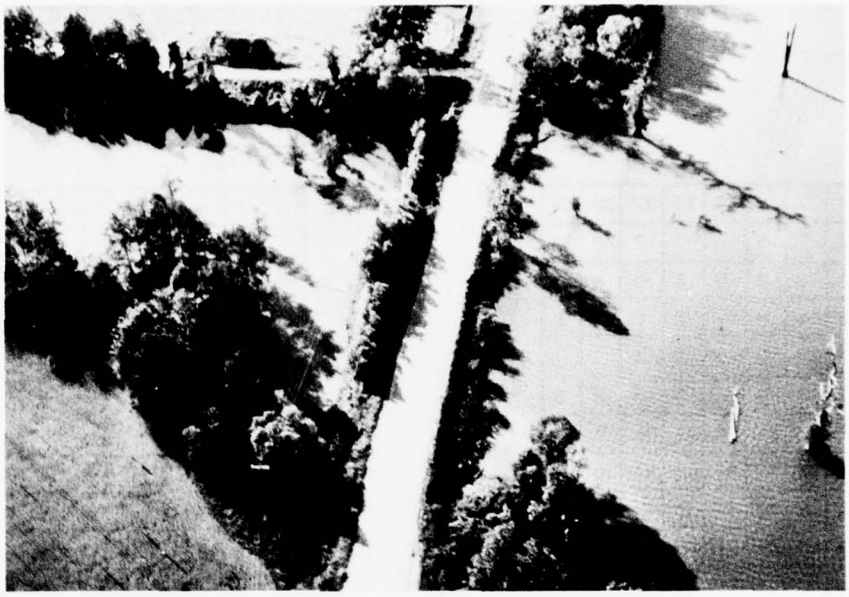
70. The large-scale image of the study area (Block 6) and an air photo of approximately the same scale were studied carefully. Then three transfer points, A, B, and C, were selected (Figure 33). They are not ideally located; it would have been better had they been more widely separated. However, as it happens, the southern end of the study area offered few opportunities for suitable transfer points. Low-altitude oblique air photos (Figure 34) illustrate the kinds of features at transfer points A and C. The features were identifiable in the Landsat data sets for July 1974 as well as February and October 1975. Figure 35 presents a 10- by 10-pixel matrix of the band-7 CCT values around transfer point A. Note that in each data set the junction of the stream with the lake bridge shows as a higher value than the "lake" pixels and that the small wide spot in the stream, just before it enters the lake and south of the junction, shows as a value significantly lower than the surrounding dry land. The junction is, of course, readily identifiable on the air photo.

71. Block 8. The problem now was to stretch (or "rubber sheet") the CCT pixel array in such a way that pixels representing transfer points are located at the same coordinates as their correlative points on the air photo. The procedure was to assign coordinate positions to the air-photo locations of the transfer points by first designating one of the transfer points as the origin of a convenient but arbitrary Cartesian coordinate system and then determining the X and Y coordinates of the other points with respect to this origin. Likewise, the transfer points within each Landsat data set were located by determining the number of pixels (N_X) in the X direction and (N_Y) in the Y direction of each transfer point. The map coordinates of the transfer points were then related to the pixel coordinates within each Landsat data set by the following two equations:

$$X_i \cos \theta_f - Y_i \sin \theta_f = N_{Xi} D_{Xi}$$

$$X_i \sin \theta_f + Y_i \cos \theta_f = N_{Yi} D_{Yi}$$

Figure 33. Location of transfer points (⊙) and sample sites (+) in study area



a. Large stream intersecting a lake (transfer point A)

b. Small lake surrounded by trees (transfer point C)

Figure 34. Types of features at transfer points A and C (Figure 33), 8 May 1976

7	7	7	7	8	9	15	27	38	34
9	9	7	7	6	7	7	13	25	33
24	24	11	6	7	6	6	6	8	13
41	41	28	14	8	8	7	8	7	8
42	42	43	32	17	8	7	7	7	7
38	38	42	40	28	(17)	15	10	8	8
21	21	21	20	20	34	37	31	18	12
29	29	32	32	39	42	40	37	36	31
28	28	32	36	38	37	38	40	39	38
35	35	40	40	39	39	39	41	41	41

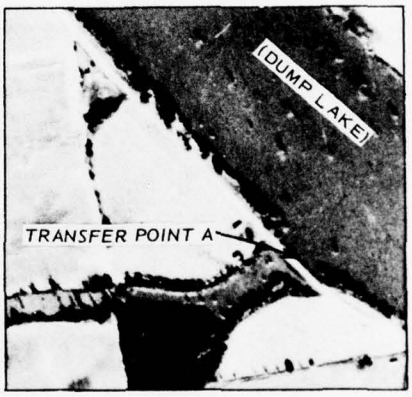
a. 11 JULY 1974

2	2	2	2	2	3	4	7	11	10
7	2	1	2	2	2	2	3	10	14
16	8	2	1	2	1	2	1	3	5
17	17	11	4	2	2	2	3	3	2
20	19	18	13	4	1	1	1	2	1
18	19	18	18	12	(4)	2	1	2	2
13	12	13	11	7	10	10	5	2	2
11	7	9	12	19	20	19	16	12	8
12	8	11	16	19	16	16	18	16	18
19	14	13	16	16	17	16	17	17	16

b. 21 FEBRUARY 1975

3	1	1	1	1	3	9	18	21	21
13	3	1	1	2	2	3	7	17	20
17	13	5	1	1	2	1	2	4	12
18	17	13	5	1	1	1	1	1	3
18	19	18	14	8	1	1	1	1	1
18	19	18	16	13	(6)	2	2	2	1
13	12	12	8	9	14	13	7	1	1
13	13	14	17	18	17	17	16	15	12
13	16	18	17	17	17	17	17	17	18
18	15	18	18	17	17	18	17	17	18

c. 13 OCTOBER 1975



d. AIR PHOTO

LEGEND

WATER

(4) TRANSFER POINT A

Figure 35. Band-7 CCT values around transfer point A (all three overpasses) with air photo for comparison

where

X_i and Y_i = the X and Y air-photo coordinates of each respective transfer point

θ_f = Landsat flight path angle with respect to true north

N_{Xi} = number of X-coordinate pixels of the transfer point within the Landsat data set

D_{Xi} = width of the Landsat pixel

N_{Yi} = number of Y-coordinate pixels of the transfer point within the Landsat data set

D_{Yi} = length of the Landsat pixel

Equations representing each respective transfer point were written and then simultaneously solved for the flight path angle and the size (D_X and D_Y) of the pixel within each Landsat data set. The values of θ_f and pixel size for the three Landsat data sets are listed in Table 3. The average values for θ_f, D_X, and D_Y were then computed and used in the equations above to bring the terrain reference system and CCT value data for each data set into geometric accordance.

72. The number of transfer points needed to bring Landsat data sets into geometric accordance depends on the size of the area being considered and the relief within the area. It is estimated that in most regions in the United States between 12 and 20 points would probably be required for an area the size of a NASA scene ($\approx$34,000 km^2). Reference 8 describes a procedure for bringing Landsat data into geometric accordance using up to 20 transfer points.

73. Block 9. The next problem was to select the sample sites (see paragraph 29) to be used in the "guided" interpretation mode. However, before making the selection, at least a few examples of each terrain type had to be actually located. In the present exercise, this was accomplished by a combination of air-photo interpretation, using the photomosaic in Figure 5, and ground truth data, as described in paragraphs 46-60. The initial classification was comprised of only three terrain types: open-water surfaces, wooded land (or forest), and non-wooded (or open) land. On the basis of all available information, 30 sample sites were selected (Figure 33) that, it was hoped, would include every significant variation of condition in the study area. As noted in

Figure 33, a Landsat pixel is an area about 1.1 mm long and 0.8 mm wide, with the long dimension north-south. Thus, a 3- by 3-pixel array (see paragraph 36) is a rectangle 3.4 mm N-S and 2.4 mm E-W. Each sample site was selected at the center of an area of homogeneity as perceived on the air photo and in an area as large as a rectangle 3.4 by 2.4 mm (Figure 33). By the use of available ground truth data and air-photo interpretation, the terrain conditions at each site during each satellite overpass were established (Table 4).

74. Blocks 10 and 11. A 3- by 3-pixel array of values around each sample site was isolated from the SA/CCT and subjected to a "normalcy" analysis to determine actual homogeneity. Table 5 gives the mean and standard deviations of the CCT values for each wavelength band for the October 1975 overflight. As noted, the standard deviations vary considerably; for example, the standard deviations are quite large (greater than 2.0) for sites 17 and 29. Also, Figure 33 shows that there are situations such that a 3- by 3-pixel array cannot quite be fitted into an area of complete homogeneity. These sites would, of course, be expected to exhibit more variance than others. Ideally, the variance exhibited by a 3- by 3-pixel array should be no larger than the variance produced by normal instrument error. Since the normal instrument error is about $\pm$2 percent, the ideal standard deviation would be $\pm$1.26 CCT units (maximum CCT value of 63 multiplied by 0.02, the normal Landsat instrument error). The assumption is that any 3- by 3-pixel array exhibiting a standard deviation of $\pm$1.26 would be likely to represent an absolutely homogeneous terrain type. In practice, such homogeneity cannot be expected because of small-scale randomized variations in surface geometry, soil moisture, etc. Since there is no way of telling ahead of time how large such variations are likely to be, the analyst should be permitted to control the amount of variance that is used to judge the homogeneity of a sample site, with the proviso that the variance should in no case be less than about $\pm$1.3 ($\pm$1.26 rounded to the nearest 0.1). This controlled parameter, which was used to determine the acceptability of the 3- by 3-pixel arrays, will henceforth be called the "screening window." Because some degree of site variance was anticipated, $\pm$1.5 CCT

units were initially accepted as the screening window. Those sample sites that exhibited standard deviations of greater than ± 1.5 CCT units in one or more of the four spectral bands were eliminated from further consideration. The assumption behind this decision was that a larger standard deviation would be prima facie evidence that two or more terrain types were represented in the sample. This decision drastically reduced the number of samples used in analysis; for example, only 17 sites remained out of the original 30 for the October 1975 data.

75. The specification of a particular screening window as the criterion for homogeneity is, of course, a specification of the amount of variance that will be permitted in a sample site. If the acceptable screening window is made larger, more sample sites will be found to be acceptable. Table 6 illustrates the effect of increasing the screening window on the number of accepted sample sites.

76. Examination of the photomosaic (Figure 33) sheds little light on the reasons for the variances or lack thereof, since some sites that appear on the photomosaic to be very homogeneous actually exhibit quite large variances. It seems that subjective analysis of air photos will not guarantee the selection of homogeneous sample sites. Therefore, the best approach would be to select a large number, expecting that many will prove unsuitable, at least initially.

77. Block 12. The initial terrain classification was to consist of three categories: water, forest, and nonforest (see paragraph 73). Thus, each of the sample sites that survived the normalcy analysis (i.e. that met the screening window criterion (Blocks 12.1 and 12.2, Plate 2)) must be assumed to be a member of a "cluster" of sample sites, each of which represents, in some degree, one of the three terrain types. Thus, the first truly subjective decision must rest on the apparent representativeness of the selected sample sites (Block 12.3, Plate 2). If only two or three survive the normalcy screening, and the terrain is known to be complex in any sense, then it may be that the number of selected sample sites is too small. In this event, the obvious step is to enlarge the screening window so that additional sample points are included (Table 6).

78. For the October overpass data, it was decided that the selected sites adequately represented the desired terrain types. The next step was to test the selected samples for similarity. Ideally, all of the samples denoting one terrain type should exhibit a relatively close degree of radiometric similarity. In practice, this means that the CCT values of all of the samples designating a desired terrain type should show a sort of family resemblance. Furthermore, the CCT values representing the samples in a "terrain" type should tend to "cluster" about a common value, and the cluster for each terrain type should be discrete, i.e. should not overlap.

79. The first step in the testing process was to select a "variance" window (Block 12.4, Plate 2). Ideally, the smallest variance window should be equal to or larger than the maximum instrument error, and thus, when dealing with Landsat data, it should be equal to or greater than $\pm$1.26 CCT units. However, the selection of the variance window is entirely subjective and must be made by the analyst. Generally, it appears to be better to first choose an initial value relatively close to the limiting minimum. This variance window (which is both positive and negative) is then algebraically added to the mean value of all four spectral bands of the selected sites (Block 12.5, Plate 2). The resulting values are then rounded to the nearest integer, simply because the CCT values are always integers. Table 7 presents the effects of this process, using three different variance windows. The two integer values that bracket each mean value represent class limits for each spectral band. In effect, the concept is that all CCT values that fall between or on the limiting values are assumed to be equivalent. For example, CCT values of 15, 16, or 17 in spectral band 4 (Table 7) would be assumed to be members of a population that included site 1, given a variance window $\pm$1.3 CCT units. However, note that increasing the size of the variance window tends to broaden the width of the class. In physical terms, this means that a wider range of CCT values will meet the limiting criteria and are thus categorized as belonging to the class.

80. The next procedural step (Block 12.6, Plate 2) was to construct a "cluster matrix" (Figure 36) for the 13 October data, using as

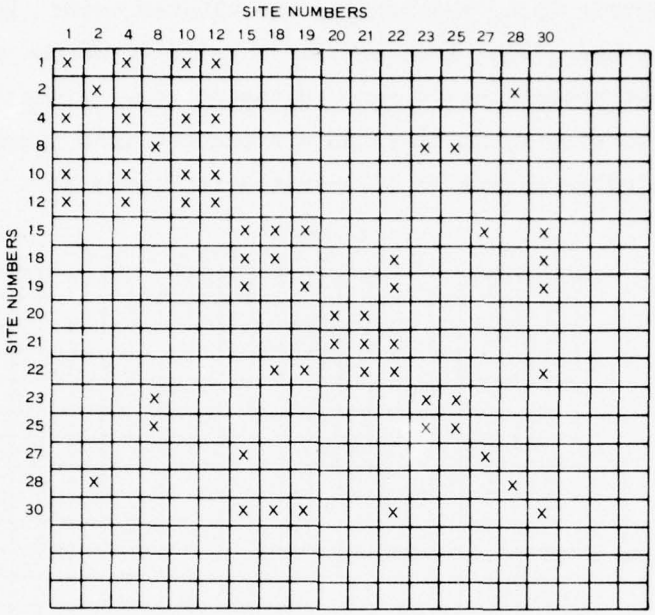


Figure 36. Cluster matrix, 13 October 1975 data, with variance window of ±1.5 CCT units

a first choice a screening window and a variance window of ±1.5 CCT units. The matrix is simply an orthogonal array with the identification numbers of the selected sites arranged in numerical sequence along both abscissa and ordinate.

81. To "fill" the matrix, "X's" were placed at each matrix position that corresponded to a similarity of spectral signatures between sites. Note that the CCT class limits in Table 7 for sites 1 and 4 indicate overlap on every band and hence are similar. Thus, an X was placed in the cluster matrix at the (1,4) and (4,1) positions. Obviously, the diagonal elements (e.g. (1,1), (2,2), etc.) all contain X's since each site is always similar to itself. The procedure was continued until every permutation of pairs of sites had been compared.

83. The problem at this point is to determine whether there is any order in the array of X's (Block 12.7, Plate 2). The procedure involves the rearrangement of the sequence of site numbers so that similarities are brought together into related "clusters." The process of rearrangement is called "diagonalization," and results in diagonalizing

the cluster matrix (i.e. rearranging the cluster matrix into diagonal form). For example, site 1 is similar to site 4, so the matrix is rearranged to put sites 1 and 4 into juxtaposition. Sites 10 and 12 are also similar to site 1, so they too are brought into juxtaposition. The result is a cluster matrix as illustrated in Figure 37.

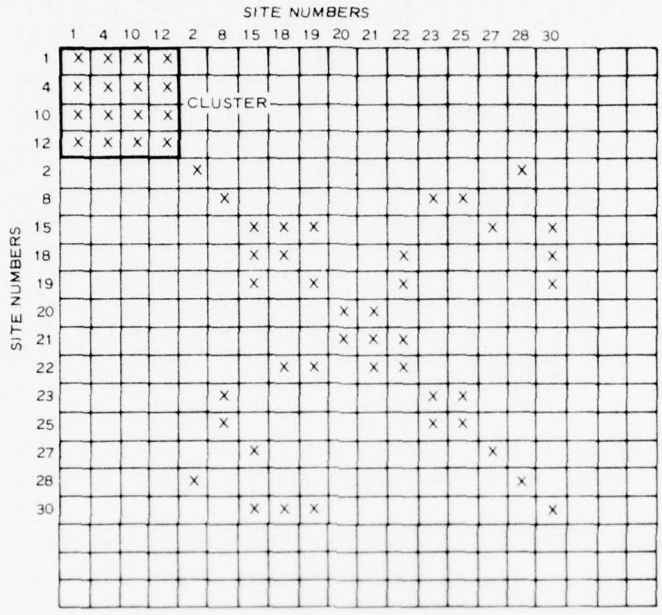


Figure 37. Cluster matrix, 13 October 1975 data, with sites 1, 4, 10, and 12 arranged in a cluster

83. Considering the rearranged matrix, it is clear that sites 1, 4, 10, and 12 are not similar to any other sites, and they can therefore constitute a cluster. Site 2 is similar to site 28, and only 28. Rearrangement to bring these two into juxtaposition results in the matrix in Figure 38. From the second rearrangement (Figure 38), it can be seen that sites 23 and 25 are similar to site 8, and they can then be brought together (Figure 39). It is now evident that sites 18, 27, and 30 are similar to site 15. Bringing these together yields the third arrangement (Figure 40). Examination of this new arrangement reveals an incomplete cluster formed by sites 15, 18, 27, and 30. It also discloses that site 22 "overlaps" the cluster, being similar to three of the four elements in the cluster. Moving site 22 into juxtaposition with the

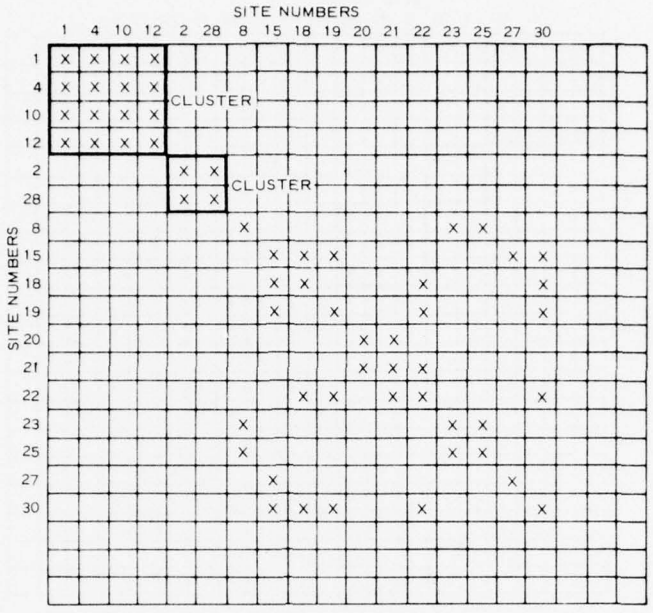


Figure 38. Cluster matrix, 13 October 1975 data, with sites 2 and 28 brought into contiguity

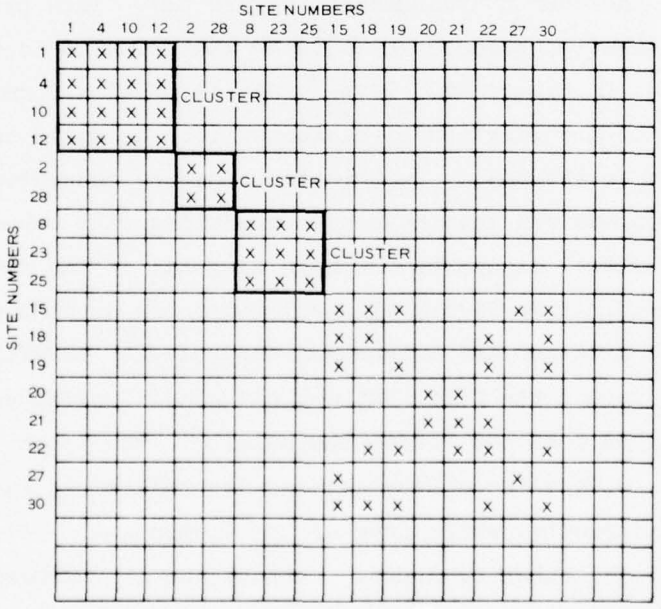


Figure 39. Cluster matrix, 13 October 1975 data, with sites 8, 23, and 25 brought into contiguity

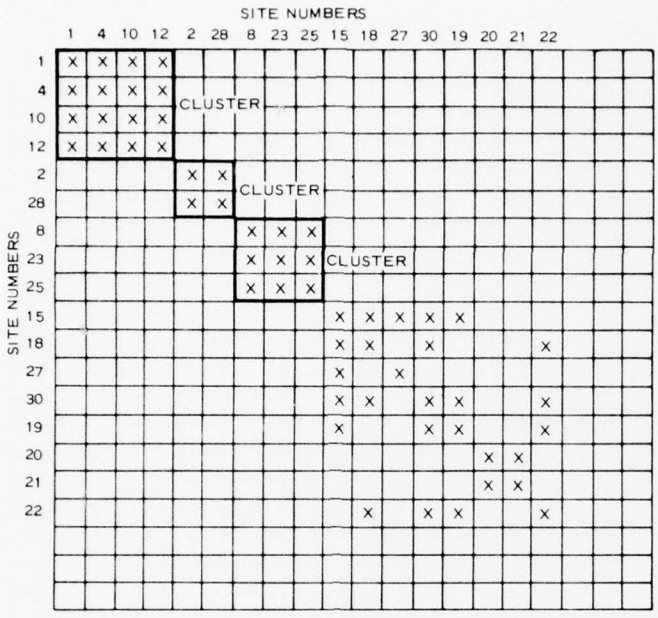


Figure 40. Cluster matrix, 13 October 1975 data, with sites 15, 18, 27, and 30 brought into contiguity

cluster yields the fourth arrangement (Figure 41). This process is called diagonalization because any cluster that results is arranged along the diagonal, as in Figure 41. Given a matrix of this form, a cluster can be defined as a collection of elements (i.e. occupied matrix positions) that can be internally (within the cluster) connected by off-diagonal connections but that have no connections with external (outside the cluster) elements except by the main diagonal of the cluster matrix.

84. The matrix in Figure 41 was interpreted as consisting of four clusters, three distinct and unique and one somewhat indefinite and incomplete. However, the fact that not all matrix positions in the cluster are filled does not lessen the utility. The point is that all of the elements of the incomplete cluster are related to each other, even if only indirectly, while none of its elements are related to elements in any of the other clusters. Further, no arrangement of its elements will yield a configuration of two or more discrete nonrelated clusters.

85. Of course, the site numbers can be arranged in a different

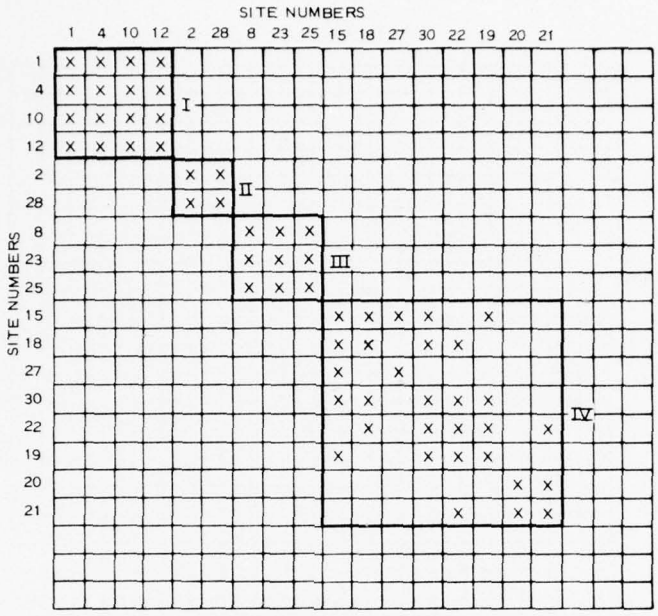


Figure 41. Cluster matrix, 13 October 1975 data; one of several possible final arrangements showing four discrete clusters

sequence (Figure 42). However, note that exactly the same four clusters, one of which is incomplete, emerges. Thus, the precise rationale by which a site number sequence is generated is unimportant.

86. In the example just described, fortune has smiled: the diagonalization process has resulted in the formation of a cluster matrix containing discrete clusters. However, as will later be seen, this is not always the case. The procedure (Block 12.8, Plate 2) for dealing with cases in which discrete clusters cannot be formed is discussed in paragraphs 110-112.

87. The initial terrain classification was to consist of only three categories: water, forest, and nonforest (or open) (see paragraph 73). Thus, each of the sites in the cluster matrix (Figure 41) was examined to determine whether the sites in each cluster represented a single terrain type (Block 12.9, Plate 2). By reference to Table 4, Table 8 can be readily compiled. The comparison of Table 8 with Figure 41 reveals that the four forest sites have been neatly placed into one of the clusters and the open sites are distributed among the

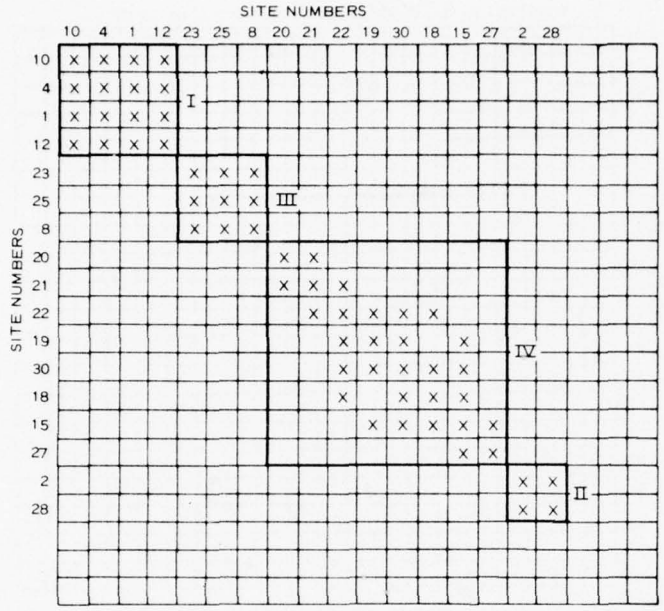


Figure 42. Cluster matrix, 13 October 1975 data, with alternative arrangement of site numbers yielding same four clusters as in Figure 41

remaining three clusters (Table 9). Each cluster then consists of sites representing a single terrain type (Block 12.9, Plate 2). If more than one terrain type had been present in any one of the clusters, the procedure described in Blocks 12.20, 12.21, and 12.22 in Plate 2 would have been applied. No case in this study required this alternative. One of the nonforest terrain types, however, is designated by more than one cluster (Block 12.10, Plate 2). The reason for this can be visualized more readily if the CCT class limits bounding each cluster in each wavelength band (Block 12.11, Plate 2) are reduced to graphic form, (Figure 43). The CCT values for the limits bounding each cluster are derived by determining the largest and smallest value in each wavelength band in each cluster. For example, Table 10 contains only those sites used to develop the cluster matrix in Figure 42, with the sites sequenced in the same order as in Figure 41. In band 4, since the smallest value is 14 and the largest is 18, the class limits of cluster I are 14-18. This is plotted as a horizontal bar in Figure 43, as are all of the other cluster classes.

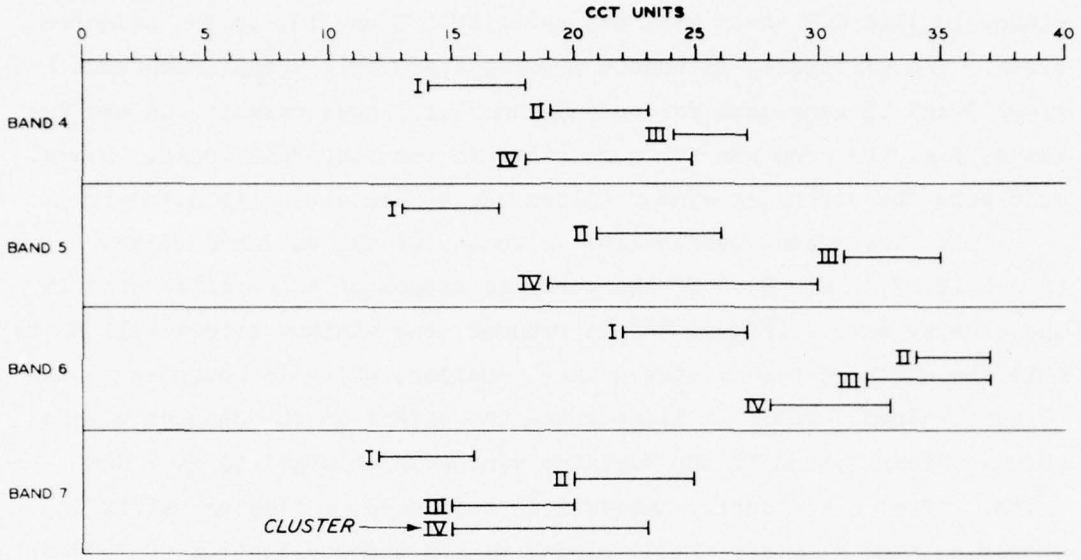


Figure 43. Graphic representation of CCT classes derived from cluster matrix (Figure 42), 13 October 1975 data, with screening window and variance window of ±1.5 CCT units

88. The nonforest clusters (i.e. II, III, and IV) were grouped separately because each failed the similarity tests in at least one spectral band. Clearly, it would be desirable to find a method of arranging the sites such that there is one cluster per terrain type. At least two ways exist in which this might be accomplished: the size of the variance window might be changed (Block 12.23, Plate 2), or the size of the screening window might be changed (Block 12.24, Plate 2), or both might be changed.

89. If the screening window is reduced, the number of selected sites will also be lessened (Table 6). On the other hand, if it is increased, sites will be added with the possibility that one of them will exhibit a CCT value in the critical wavelength band that will fill the gap between clusters (Figure 42). In the example, the analyst would look for a value that would fill the critical gap between the class limits of clusters II, III, and IV in bands 5 and 6. However, note that the class limits of clusters II and III overlap only band 6. Thus, ideally, the sites that are added should also incorporate values to fill these gaps.

90. From Table 6, it is apparent that widening the screening

window to ±1.6 CCT units will add two sites, 9 and 13, to the selected sites. The difficulty is that a reference to Table 4 indicates that both sites 9 and 13 represent forest and thus will serve only to confuse the issue, i.e. the need was for more sites in the nonforest areas. Hence, enlarging the screening window (Block 12.24, Plate 2) will not help.

91. The second possibility is to change the variance window (Block 12.23, Plate 2). If the variance window of those sites used in the cluster matrix (Figure 42) is reduced, the minimum effect will be to make the width of the cluster classes smaller, which is opposite to the effect desired. Table 11 illustrates the effect on the classes within each wavelength band if the variance window is enlarged to ±1.7 CCT units. After a similarity analysis is completed, a cluster matrix is assembled, the clusters are developed by the diagonalization procedure, and the cluster matrix in Figure 44 emerges. Although cluster C is a greatly extended cluster, it nevertheless meets all of the criteria for a cluster and also places all nonforest sample sites into a single cluster. Thus, the objective is satisfied. Table 12 is the result of arranging the class limits data with the associated clusters.

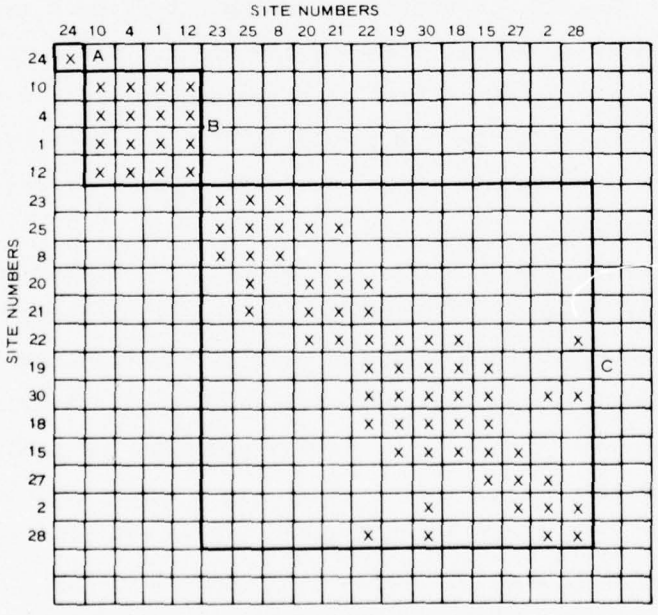


Figure 44. Cluster matrix, 13 October 1975 data, with clusters derived by using variance window of ±1.7 CCT units

92. The water site (site 24, Figure 33) did not survive the normalcy analysis in the October 1975 data. Because water so strongly absorbs band-7 radiation, the CCT value of site 24 in band 7 should be so low that it will not be similar to any other site. The data in Table 5 indicate that the mean CCT value of site 24 in band 7 is 2.2, well below any of the values exhibited by any other site, and thus this expectation is realized. Since including it among the selected sites cannot disturb the normalization process (even though it does not meet the normalcy criterion), site 24 was arbitrarily included in the diagonalized cluster matrix (Figure 44) as an independent one-element cluster. The water cluster could have been added to Figure 41 in the same way, but it was deferred to this point in the procedure so that the discussion in paragraph 87 would be less encumbered.

93. The next procedural step was to determine whether the terrain classifications were to be "inclusive" or "exclusive." Inclusive classifications are those in which the classes include all possible contingencies; every patch of the ground must be placed in one or another of the categories. Therefore, the analyst at this point must decide (Block 12.12, Plate 2) which type is required. In the present instance, the system is inclusive: each patch of ground must be either water, forest, or nonforest. Exclusive classifications are those in which only certain aspects of a landscape are included in the classification, and thus a certain proportion of the ground will remain outside the class criteria; some patches will be unclassified. In such a case, the classification actually consists of the specified classes plus an additional class in which all things not included in the criteria defining the specified classes are placed (Block 12.17, Plate 2).

94. The example cluster class limits presented in Table 12 are an exclusive set. However, since the classification system under discussion is an inclusive system (the three clusters represent three terrain types that, by definition, must cover 100 percent of the area, i.e., each point on the ground must be either water, forest, or nonforest), the question at this point is, how are inclusive cluster class sets derived from the exclusive sets in Table 12?

95. The procedure (Block 12.13, Plate 2) adopted in this study determines by a "cluster separation logic" the combinations of bands that affect the separation of clusters and then derives by a set of "cluster group stretching rules" the inclusive cluster class sets that represent and preserve the integrity of each cluster of the cluster matrix. A consequence of this procedure is that often more inclusive cluster class sets can be determined than there are original clusters. Furthermore, since there is not enough information available to reliably select the optimum set by objective procedures, a subjective decision-making process had to be used.

96. From the graphic form of the Table 12 data presented in Figure 45, the cluster separation logic used to determine combinations of bands affecting cluster separations for the October 1975 data and the symbolic representation for that logic is as follows:

a. Separation 1: Clusters A, B, and C can be separated by using band 6. The separation logic for this cluster group is represented by the symbol $(A_6^1\ B_6^1\ C)$.

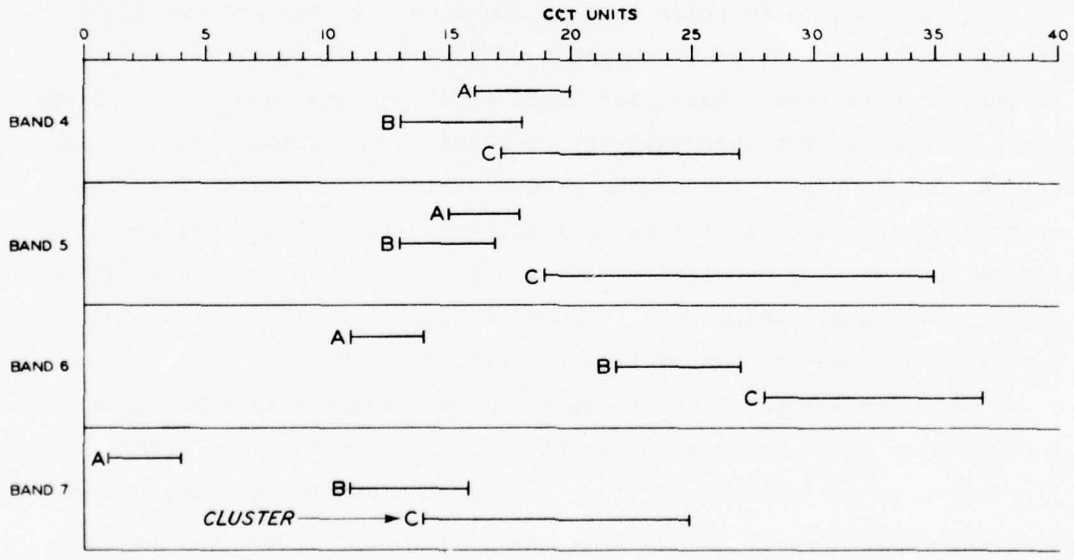


Figure 45. Ranges of values comprising clusters, 13 October 1975 data, with variance window of ±1.7 CCT units

b. Separation 2: Cluster A can be separated from clusters B and C by using band 7. Clusters B and C can be separated from each other by using band 5. The separation logic for these two cluster groups is represented by the symbol $(A) \frac{|}{7} (B \frac{|}{5} C)$.

c. Separation 3: Cluster A can be separated from clusters B and C by using band 7. Clusters B and C can be separated from each other by using band 6. The representation for this separation logic is $(A) \frac{|}{7} (B \frac{|}{6} C)$.

d. Separation 4: Cluster A can be separated from clusters B and C by using band 6. Clusters B and C can be separated from each other by using band 5. The representation for this separation logic is $(A) \frac{|}{6} (B \frac{|}{5} C)$.

97. Once the cluster separation logics had been determined, the class limits of the clusters were expanded (or stretched) by means of a set of cluster group stretching rules (Figure 46) to include all possible values in all wavelength bands. The rules are divided into three sets: an omnigroup, intragroup, and intergroup set of rules. The omnigroup rules apply to all clusters regardless of their cluster groupings; the intragroup rules, only to bands affecting separation within a cluster group; and the intergroup rules, only to bands affecting a separation between cluster groups.

98. Table 13 presents the consequences of the cluster group stretching rules as they were successively applied to the cluster separation logic $(A) \frac{|}{7} (B \frac{|}{5} C)$ (i.e. Separation 2, paragraph 96). Table 14 illustrates the results of the group stretching rules on all the separation logics. At this point in the process, it was still undetermined which of the class sets (Table 14) would produce the best map of the area of interest. In this context, the best map would be that one in which the maximum number of pixels was correctly identified as to terrain type. In the example used up to this point, it is extremely difficult to estimate which will produce the best product, since all of them are capable of clearly separating all three clusters. The only obvious way of resolving the problem was to actually produce preliminary maps (Blocks 12.14 and 12.15, Plate 2) of a small region and check the results against the actual terrain and against an air photo that was reliably interpreted. Figures 47-50 are the maps resulting from using

Omnigroup Rules

Rule 1: The lower class limits in bands not affecting cluster or cluster group separations are extended to 0.

Rule 2: The upper class limits in bands not affecting cluster or cluster group separations are extended to 63.

Intragroup Rules

Rule 1: The lower limiting CCT value in the band affecting separation is extended to 0.

Rule 2: The upper limiting CCT value in the band affecting separation is extended to 63.

Rule 3: In the band affecting separation, the upper and lower class limits of a cluster are extended until the gaps between adjacent clusters are closed. If the gap between an adjacent cluster is an even number of CCT value units, divide the values equally. If the gap is an odd number of CCT value units, assign the extra value to the lower of the two clusters.

Rule 4: If the cluster group consists of only one cluster, extend its lower and upper class limits to 0 and 63, respectively, in every band not affecting an intergroup separation.

Intergroup Rules

Rule 1: The lowest CCT value in the band affecting separation is extended to 0.

Rule 2: The highest CCT value in the band affecting separation is extended to 63.

Rule 3: In the band affecting separation, the lowest and highest CCT class limits of a cluster group are assigned to each cluster of that group. For example, clusters X, Y, and Z are in the same cluster group and have class limits of 21-25, 23-29, and 19-24 in band 6, respectively; then the band 6 class limits assigned to X, Y, and Z are 19-29.

Rule 4: In the band affecting separation, the upper and lower limits of a cluster group are extended until the gaps between adjacent cluster groups are closed. If the gap between an adjacent cluster group is an even number of CCT value units, divide the number equally. If the gap is an odd number of CCT value units, assign the extra value to the lower of the two cluster groups.

Figure 46. Cluster group stretching rules

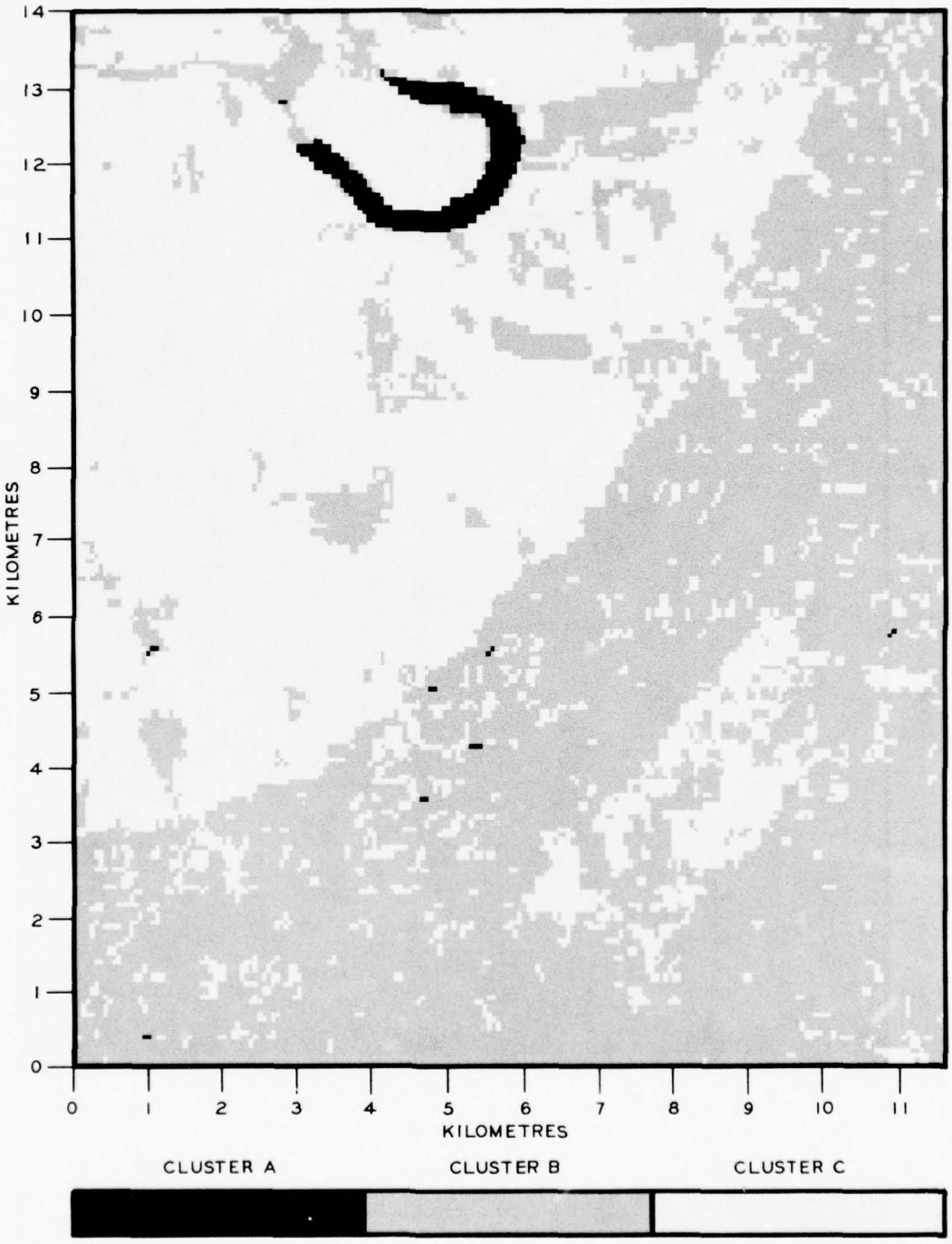


Figure 47. Study area as interpreted by cluster class set 1 (band 6, 13 October 1975 data)

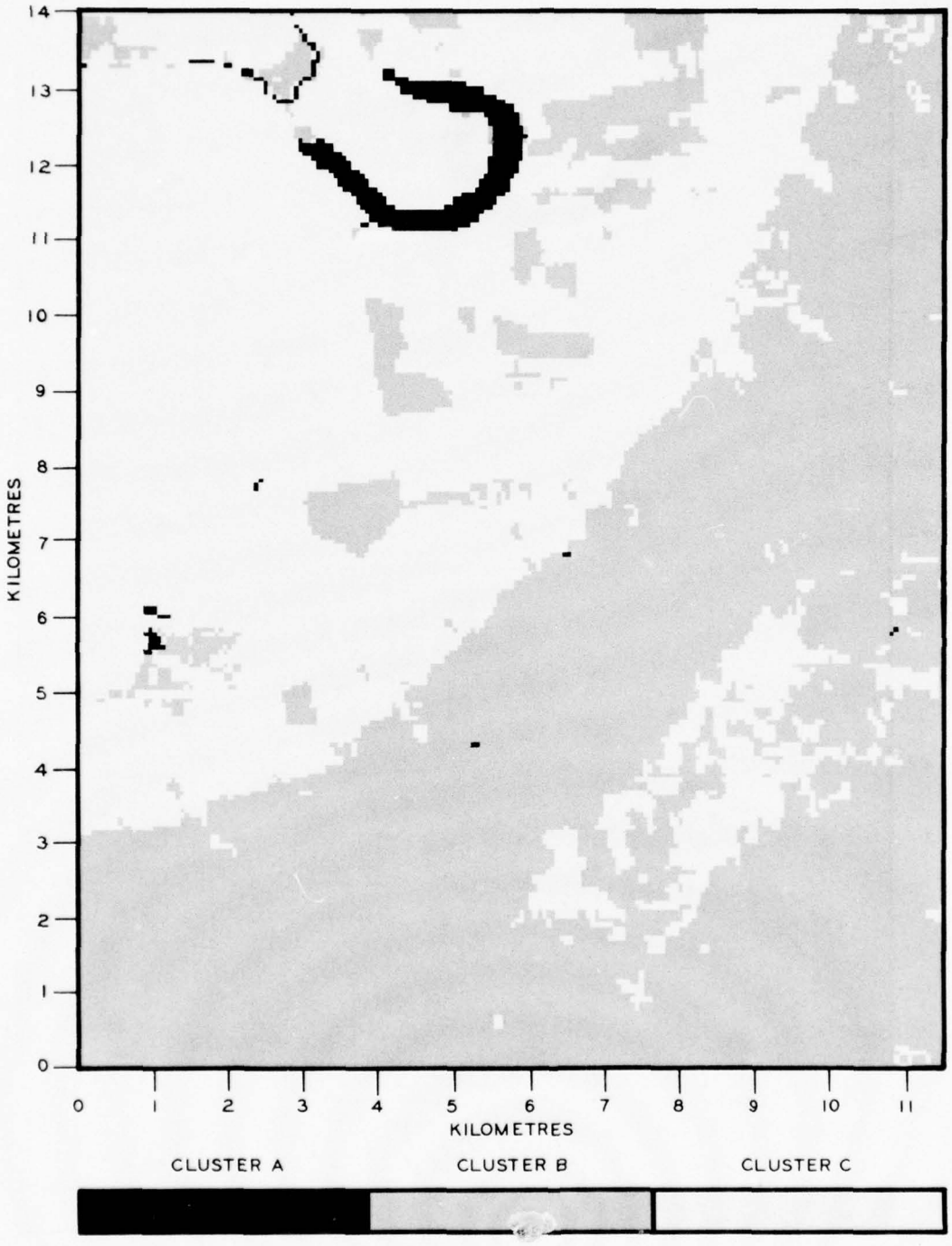


Figure 48. Study area as interpreted by cluster class set 2 (bands 5 and 7, 13 October 1975 data)

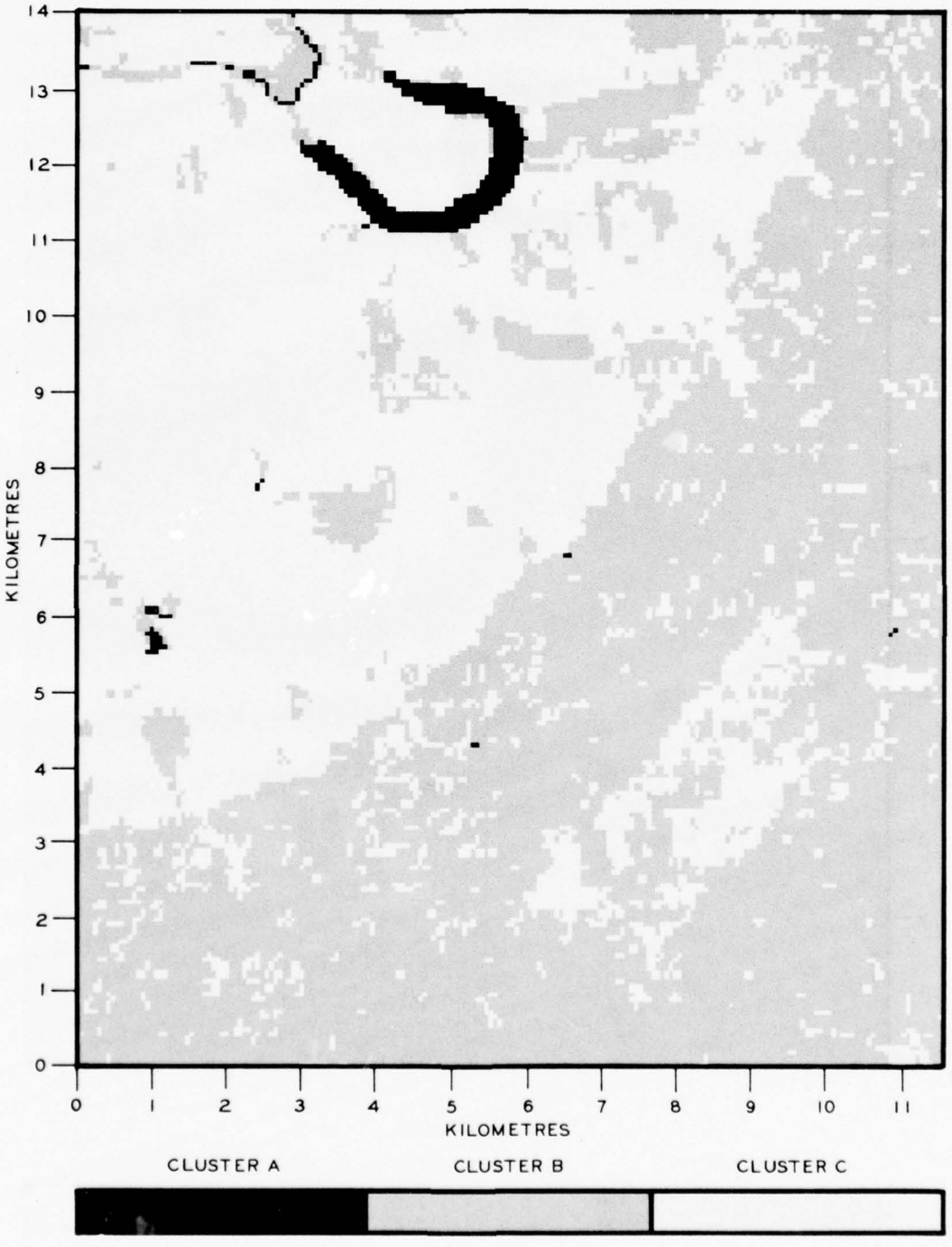


Figure 49. Study area as interpreted by cluster class set 3 (bands 6 and 7, 13 October 1975 data)

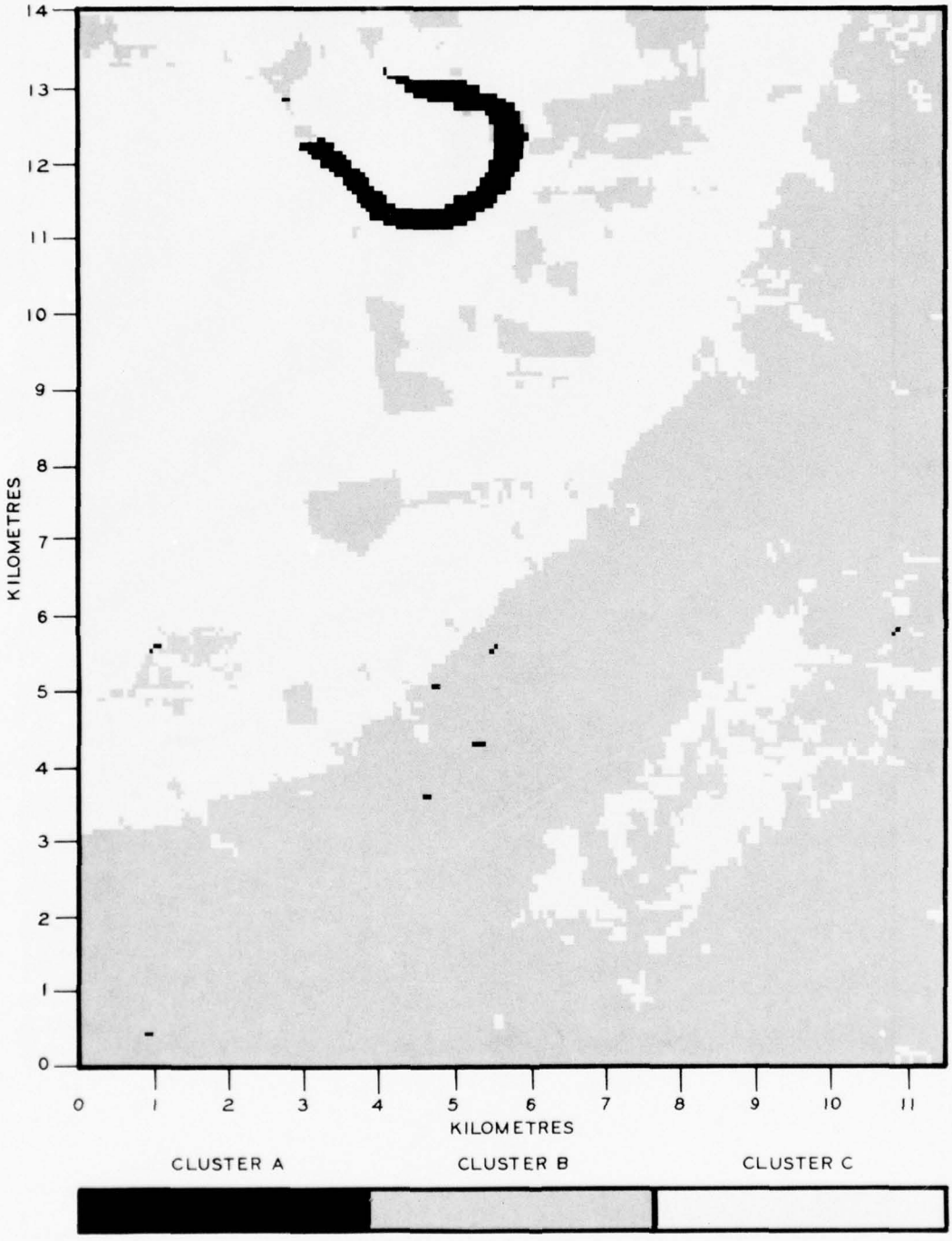


Figure 50. Study area as interpreted by cluster class set 4 (bands 5 and 6, 13 October 1975 data)

the "stretched" CCT class sets given in Table 14.

99. A comparison between the maps generated by the four cluster class sets and an air photo of the same region (Figure 14) indicates that cluster class set 2 (Figure 48), which employs bands-5 and -7 radiance data, provides the best interpretation of the terrain conditions (Block 12.16, Plate 2). The maps produced by cluster class sets 1 and 3 (Figures 47 and 49) depict numerous patches of nonforest (cluster C) within the large forest areas. Also, set 1 does not detect the presence of the river in the upper-left part of the study area. The map produced by cluster class set 4 (Figure 50) also does not detect some of the prominent water bodies (the river and a small lake in the west-central part of the study area). On the other hand, cluster class set 4 did give an excellent interpretation of the distribution of forest and nonforest in the study area.

100. The subjective impressions described above were then checked by comparing the area of each class, as determined by automatic classification, with the areas as carefully mapped by air-photo interpretation by a skilled interpreter with excellent ground truth data (Figure 14). The areas in Figure 14 were carefully measured using a dot-count technique. Table 15 gives the resultant comparisons. The results are somewhat surprising. Note that the errors from set 3 are somewhat less than those obtained from set 2, despite the fact that the subjective impression is that the map produced by set 2 is superior to that produced by set 3. The question is, why does not the subjective impression match the objective analysis? The apparent reason is subtle. The areas presented in Table 15 are meaningful only as totals; there is no necessary relation between the total area covered by a category and the location of individual pixels of that category. Thus, at least in principle, every pixel could be classified wrongly, and yet the totaled areas could agree perfectly. The obvious conclusion is that total areas are not meaningful as a basis for a judgment as to which of several classification schemes is superior. The only valid basis is to make the comparison on a pixel-by-pixel basis. When this is done (Figures 14 and 47-50), it will be noted that cluster class sets 1 and 3 have many

obviously misclassified pixels scattered throughout both the forest and nonforest areas. The remarkable agreement between total areas is fortuitous. Cluster class set 4 has far fewer misclassified pixels in the forest and nonforest areas and is deficient only in the fact that it does not appear to map water surfaces as effectively as set 2. However, on balance, cluster class set 2 was chosen as the best of the classification schemes for the October 1975 data.

101. In the event that the classification system is an exclusive type (paragrapn 93), the procedure is to go directly from Block 12.17 to Block 13 (Plate 1). In this case, of course, the cluster class sets will include not only class limits for the desired terrain types but also an "unclassified" category, which will accumulate all pixels not meeting the criteria of the cluster class sets.

102. Block 13. Having made all of the decisions required to classify every pixel in the scene or area of concern, the procedure was to do exactly that, using the original geometrically uncorrected CCT's. The program that performed this operation accepts the limiting CCT class values for each cluster (which represents a terrain type) in each wavelength band. The four CCT values defining each pixel are then compared with the CCT class specifications (Table 14), and that pixel is then coded with a number that specifies the terrain type that correlates with the set of CCT classes into which it fits. The result is a new CCT containing only the category codes instead of the original sets of CCT values.

103. Block 14. The next procedural step was to bring the new CCT containing only the category codes into geometric accordance with the air photos. The Landsat pixel array could have been brought into geometric accordance prior to classification. There are, however, compelling reasons for not doing so. The most important one revolves around the matter of pixel shape. The Landsat pixel is rectangular, approximately 58.1 by 79.2 m; the pixel produced by the Optronics film reader/writer (see Appendix A) is a perfect square. The discrepancy can be approximately corrected by compiling a new pixel array in which every 3rd and 20th scan line of the original CCT is duplicated. This

procedure increases the number of scan lines in a Landsat scene from 2340 to 3237. Thus, there are approximately 1.36 times as many pixels in the reformatted CCT as in the original Landsat CCT. The new tape can now be used on the film reader/writer to produce an image that is approximately the same scale both parallel to and at right angles to the orbital path; whereas had the original tape been used, the film reader/writer would have produced an image in which the scales differed by a factor of approximately 1:1.36.

104. The important aspect of this, with respect to the classification process, is that all of the available data are in the original CCT; reformatting adds pixels but not information. Thus, if all manipulations required in the classification procedure are done prior to reformatting, substantial savings in computer time are effected, simply because fewer pixels need be handled.

105. The second major reason for deferring shape correction stems from the image skew produced by the rotating earth. Approximately 24 sec are required for Landsat to move 185 km (the length of a Landsat scene) along its orbit, and during this time, the rotation of the earth carries the surface eastward about 11.49 km at the latitude of Vicksburg, Mississippi. This means that the surface moves eastward with respect to the satellite orbital path one pixel width (57 m) in about 0.124 sec. Therefore, in the time that it takes the surface to move eastward the width of one pixel, the satellite moves southward along the orbit about 11.6 scan lines. This systematic error can be approximately corrected (see Reference 8 for a more accurate correction) by offsetting each successive group of 12 scan lines to the westward one pixel width. For an area the size of a Landsat scene, a total of 260 "false pixels" are inserted on the tape, divided appropriately between the ends of each scan line, to bring the digital array back into a rectangular array. The result of this process is only barely perceptible at a scale of 1:1,000,000 on the resulting film because of the small pixel size used to expose the film, but it is nevertheless effective for approximately correcting for skew. The effect, however, is to add 841,620 pixels to the data array. If these pixels are included in the data array prior

to classification, special arrangements must be made so that they are not included among those pixels to be classified. Regardless of how this is done, the process will require computer time. Thus, conservation of computer time is achieved by deferring geometric correction until after the classification procedure is completed.

106. Block 15. The final step is to write a map showing the distribution of terrain types, as determined by the procedure described above. The point of departure is the reformatted CCT in which each pixel is identified by a "terrain type" code (see paragraph 102) and in which each pixel is located in such a way that it will be located in the proper position when fitted over a base map.

107. The light-emitting diode (LED) in the Optronics film reader/writer (see Appendix A) is capable of producing 256 light-intensity levels, which results in as many "gray shades" on the exposed film. Since the human eye has difficulty distinguishing more than about five to seven shades of gray, it is usually wise to restrict the number of categories on any one map to no more than seven different categories. In the present instance, since only three categories were required, this constraint offered no difficulties. The gray shades that were selected for best portrayal of the three categories were as follows:

Mapping Cluster	Terrain Type	LED Intensity Level	Gray Shade
A	Water	255	Black
B	Forest	70	Gray
C	Nonforest	0	White

These values were used to produce a map of the study area (Figure 51), and also a map of the entire October 1975 Landsat scene (Figure 52), each classifying the areas with the three terrain types.

Supplemental tests of procedure

108. Since only one test of a procedure constitutes a somewhat uncertain basis for evaluating its utility, two additional Landsat scenes (Table 2) covering essentially the same region were processed. The differences in season (note that one scene was obtained in the winter

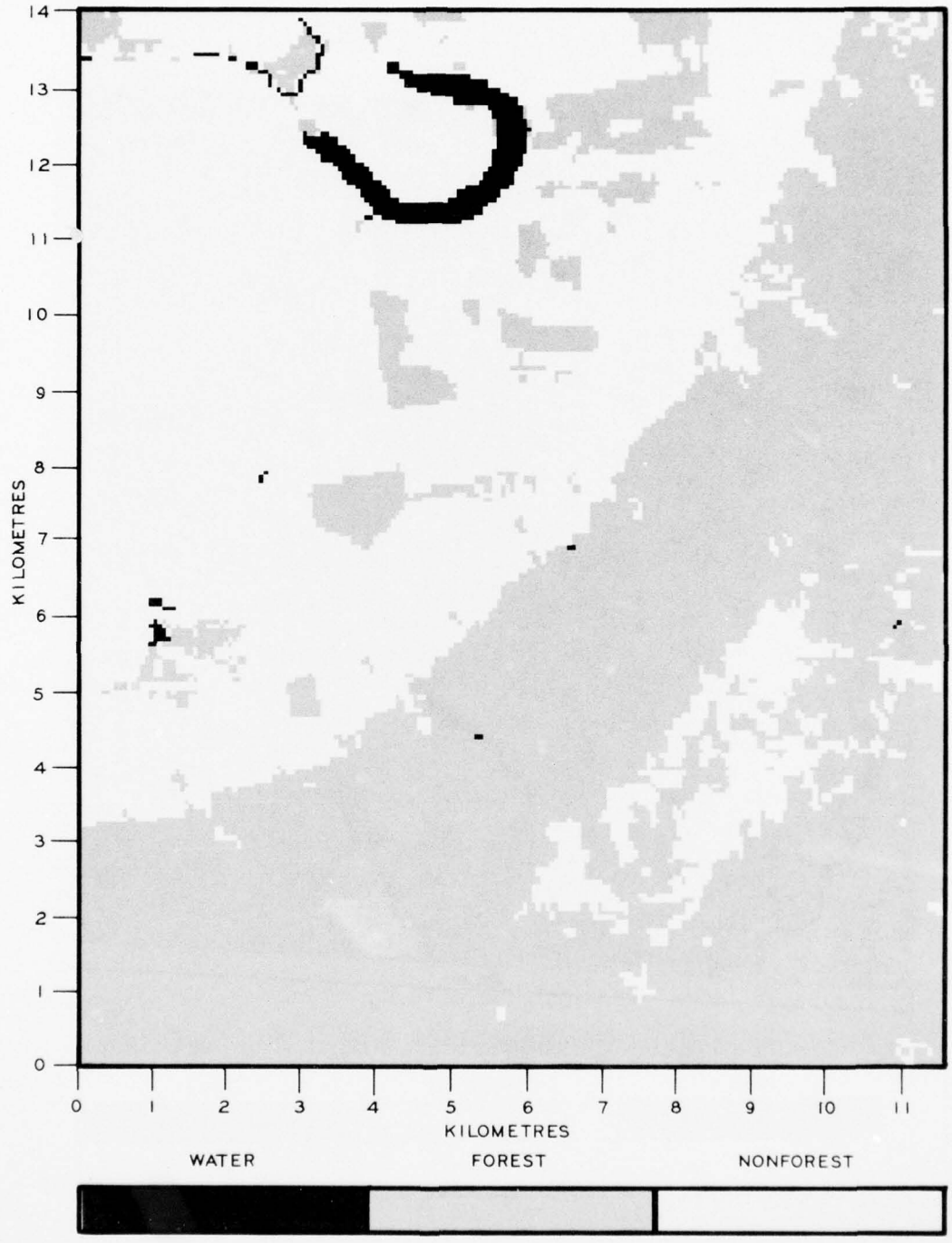


Figure 51. Study area showing interpretation of water, forest, and nonforest from 13 October 1975 data

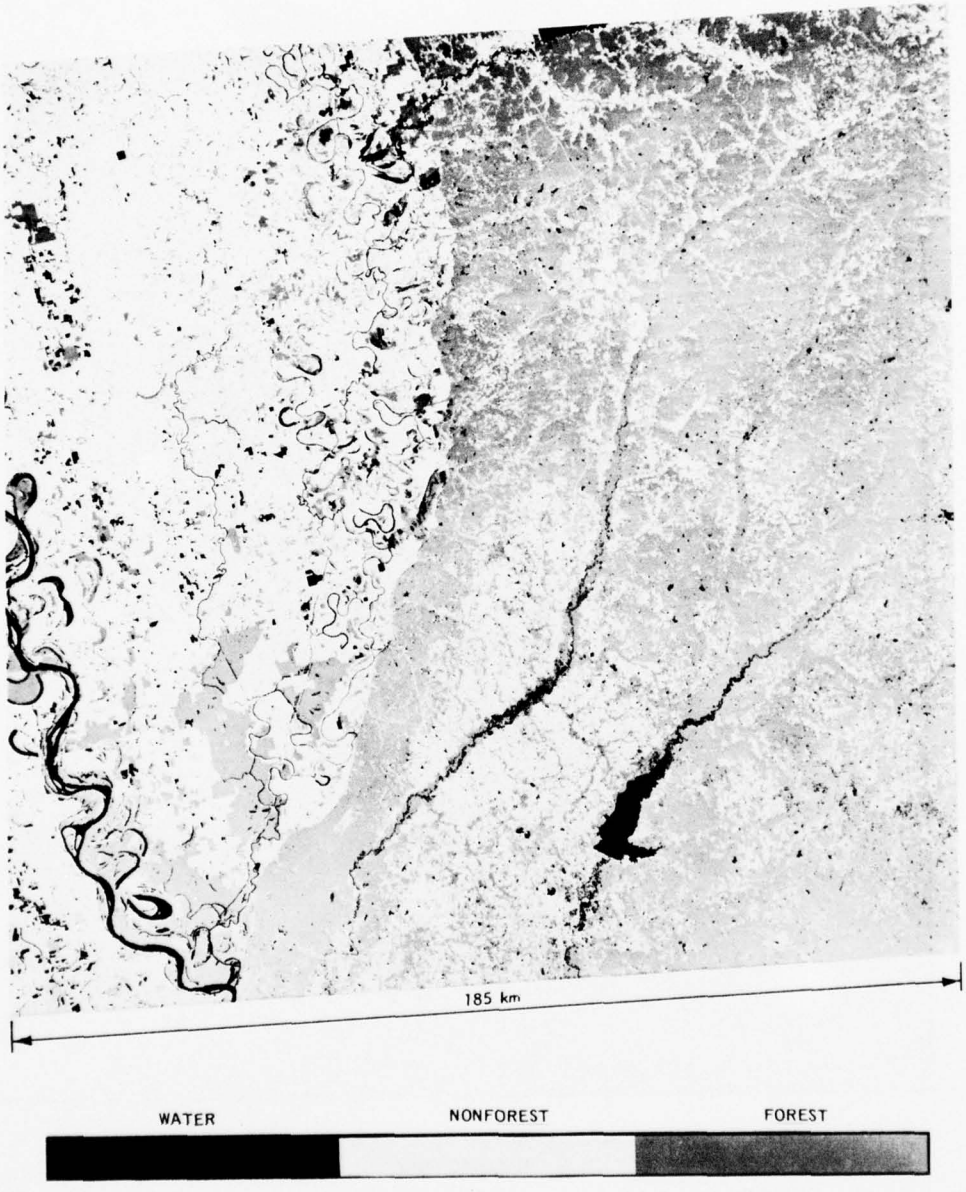


Figure 52. Entire 185-km Landsat scene for 13 October 1975

(21 February 1975), one in summer (11 July 1974), and another in the autumn (13 October 1975)) were expected to produce significant changes in terrain reflectance properties and, therefore, in spectral signatures. The sun position also changes significantly and thus causes changes in spectral signatures because of differences in reflectance geometry. Finally, normal agricultural practices could be expected to change field characteristics in the cultivated areas. These factors all conspire to make certain that the cluster class sets derived for the October 1975 scene would not be suitable for the other two scenes. In effect, the analytical procedure must start from the beginning for each scene.

109. However, the same test area (Figures 4 and 5), as well as the same sample sites (Figure 33) was used in all cases. Further, a modest amount of ground truth data had been collected (see Part III) prior to the analysis of the imagery data, and thus the terrain type at each sample site was known (Table 4). With this somewhat tenuous ground truth data, the July 1974 and February 1975 CCT's were subjected to exactly the same procedure as the October 1975 CCT.

110. Analysis of 11 July 1974 data. Table 16 presents the result of the normalcy analysis on the 3- by 3-pixel arrays around the 30 sample sites. Table 17 shows the effects of changes in the size of the screening window on the number of sample sites. For instance, only 10 sample sites survive with a screening window of $\pm$1.5 CCT units, and only 15 survive with a screening window of $\pm$1.8 CCT units. As a general principle, the degree of representativeness can be expected to be directly related to the number of sample sites. The initial cluster matrix was constructed with those sample sites derived from a screening window of $\pm$1.8 CCT units and a variance window (Table 18) of $\pm$1.5 CCT units. After diagonalization, the cluster matrix is arranged, as in Figure 53. At this point, the procedure is at Block 12.8, Plate 2, and it is evident that clean clusters cannot be formed. There is a possibility that a smaller variance window will eliminate the cluster overlap (Block 12.18, Plate 2).

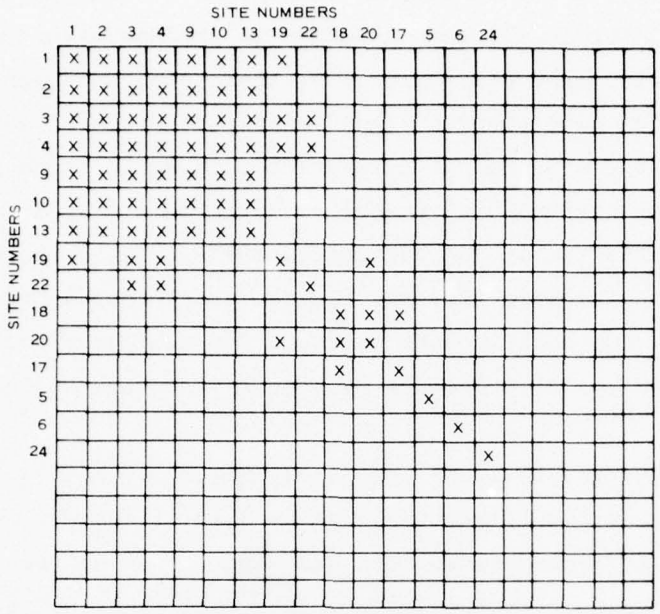


Figure 53. Cluster matrix, 11 July 1974 data, with screening window of ±1.8 CCT units and variance window of ±1.5 CCT units

111. Reducing the variance window to ±1.3 CCT units (the smallest acceptable window) results in the cluster matrix in Figure 54. It is clear that clean clusters cannot be formed. One alternative remains, namely that of reducing the screening window (Block 12.19, Plate 2), in the hope that clean clusters will form.

112. A screening window of ±1.5 CCT units was then selected, primarily because it was successful for the October 1975 data. Figure 55 shows the cluster matrix that results from using the variance window of ±1.5. Since the clusters have not yet cleanly formed, the obvious thing to do would be to reduce the variance window. The use of a smaller variance window (±1.3 CCT units) results in the CCT classes given in Table 19 and the cluster matrix in Figure 56. Figure 57 is a graphic presentation of the resulting cluster class limits.

113. An examination of Figure 57 indicates that, just as with the October 1975 data, separation of water is very easy with either band 6 or band 7. Although previous experience had indicated that

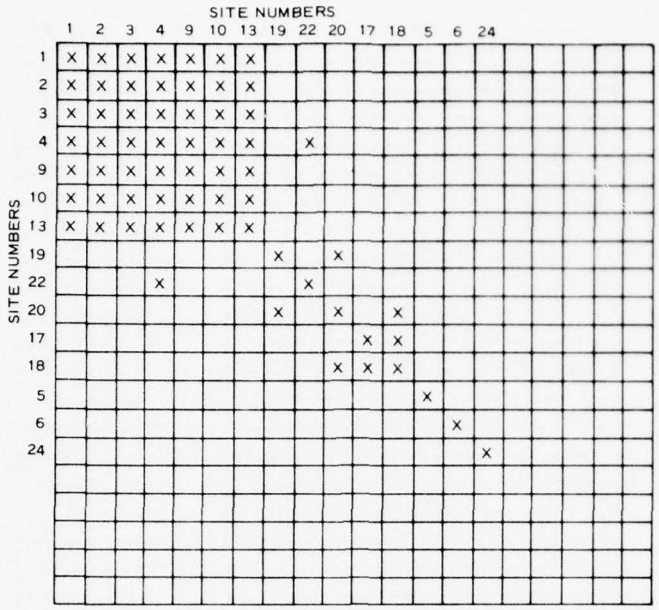


Figure 54. Cluster matrix, 11 July 1974 data, with screening window of ±1.8 CCT units and variance window of ±1.3 CCT units.

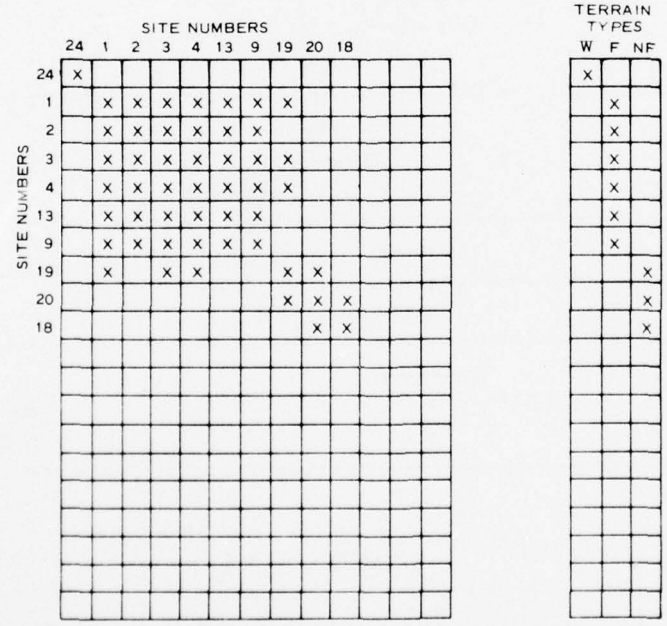


Figure 55. Cluster matrix, 11 July 1974 data, with screening window and variance window of ±1.5 CCT units

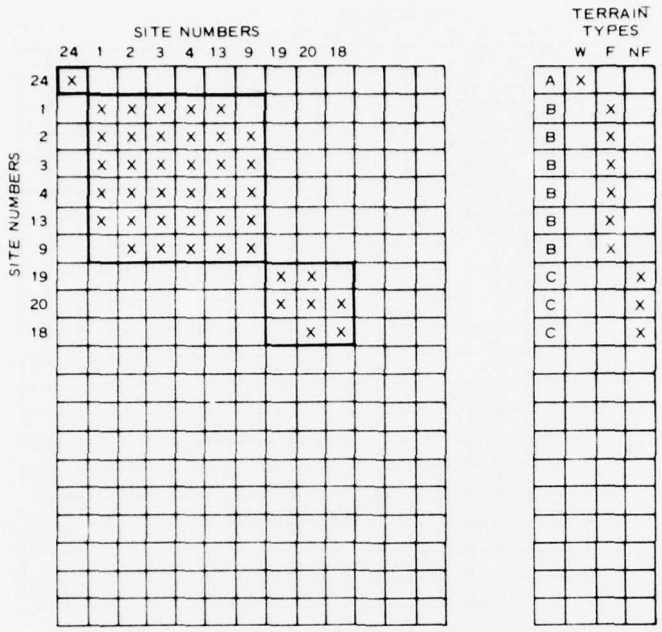


Figure 56. Cluster matrix, 11 July 1974 data, with screening window of ±1.5 CCT units and variance window of ±1.3 CCT units

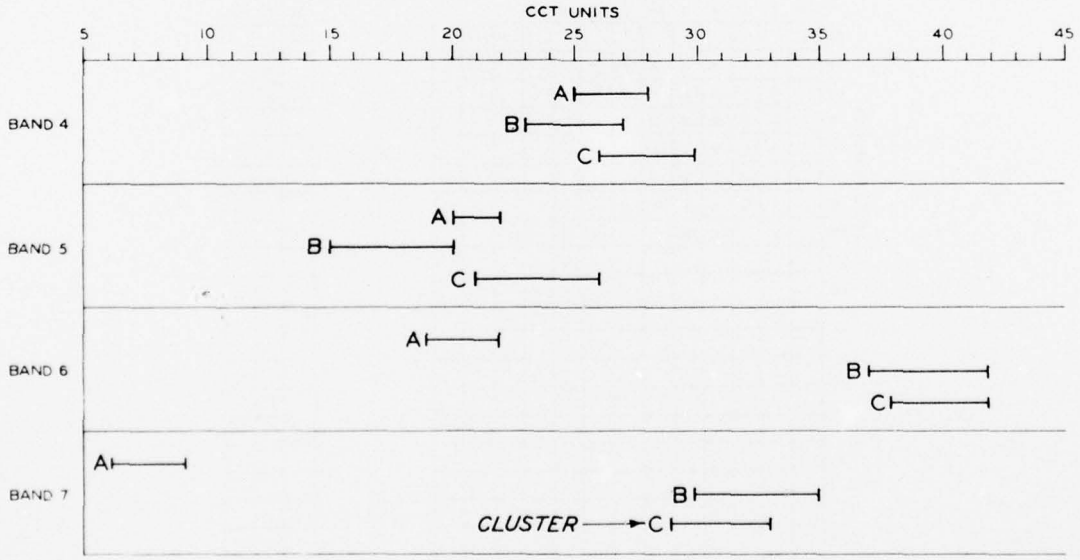


Figure 57. Graphic display of cluster class limits, 11 July 1974 data, with screening window of ±1.5 CCT units and variance window of ±1.3 CCT units

band 7 would yield superior results, it was decided to test both bands. Clusters B and C (forest and nonforest) can obviously only be cleanly separated by the use of band-5 data. Table 20 presents the resulting stretched cluster class sets as eventually used. Figure 58 shows the final map product of the use of cluster class set 1.

114. Analysis of 21 February 1975 data. The same sample sites (Figure 33) were used for the analysis of the February 1975 as for the July 1974 and October 1975 data sets. Table 21 gives the mean and standard deviations of the signatures taken from the sample 3- by 3-pixel arrays, and Table 17 lists the sites that survived normalcy evaluation at various screening windows. Table 22 shows the effect of changes in the variance window on the CCT class limits of the 17 sites that survived normalcy analysis with a screening window of ±1.8 CCT units.

115. Figure 59 is the cluster matrix that results from the use of a variance window of ±1.5 CCT units, and Figure 60 is the matrix in diagonal form. It is evident that only two clusters can be formed, one consisting of only one site (site 29). Reducing the variance window to ±1.3 CCT units yields a cluster matrix, as shown in Figure 61. It is clear that site 13 is the anomalous site; its removal results in a clean separation between forest and nonforest, as indicated by the site characteristics (Table 4) along the right margin of the figure. If it could be removed from the set of site numbers, all would be well.

116. The procedure for removing site 13 from contention is to reduce the screening window. Table 17 indicates that site 13 will be eliminated if the screening window is reduced to ±1.5 CCT units. Figure 62 shows the cluster matrix resulting from the use of a screening window and a variance window of ±1.5 CCT units. Though the forest sites form two discrete clusters and are somewhat underrepresented (2 out of 12 sites), at least the nonforest sites now form a discrete cluster. Figure 63 is a graphic representation of the resulting cluster class limits (Table 23). The water site, site 24, has been arbitrarily added, even though it did not meet the normalcy criterion.

117. As noted in Figure 63, band 7 can be used to separate clusters A and B_1 from clusters B_2 and C; band 5 offers an opportunity

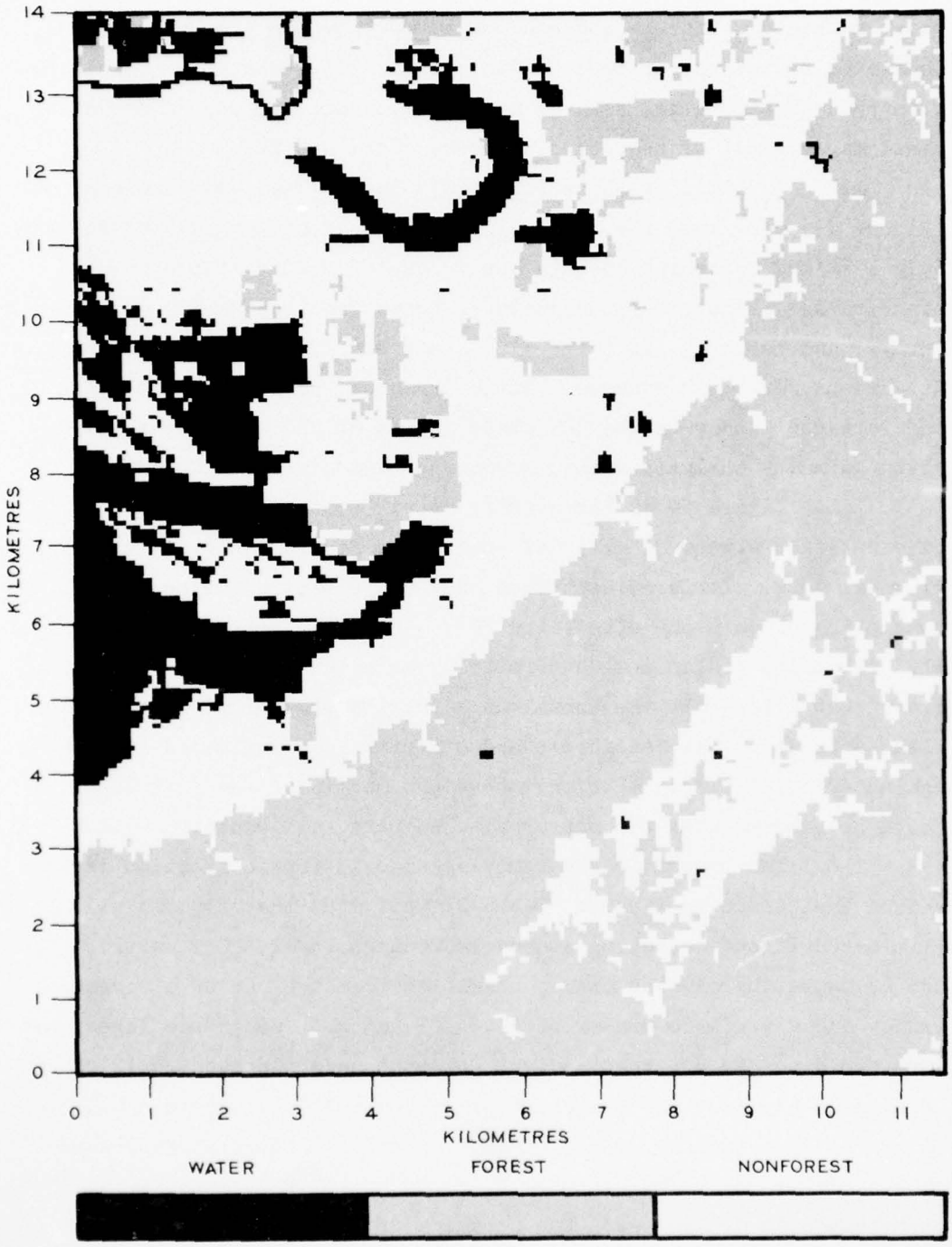


Figure 58. Study area showing interpretation of water, forest, and nonforest from 11 July 1974 data

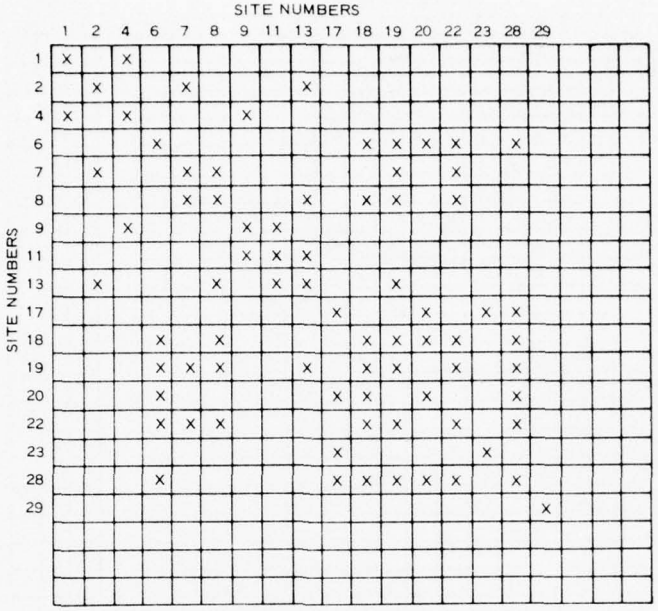


Figure 59. Cluster matrix, 21 February 1975 data, with screening window of ±1.8 CCT units and variance window of ±1.5 CCT units

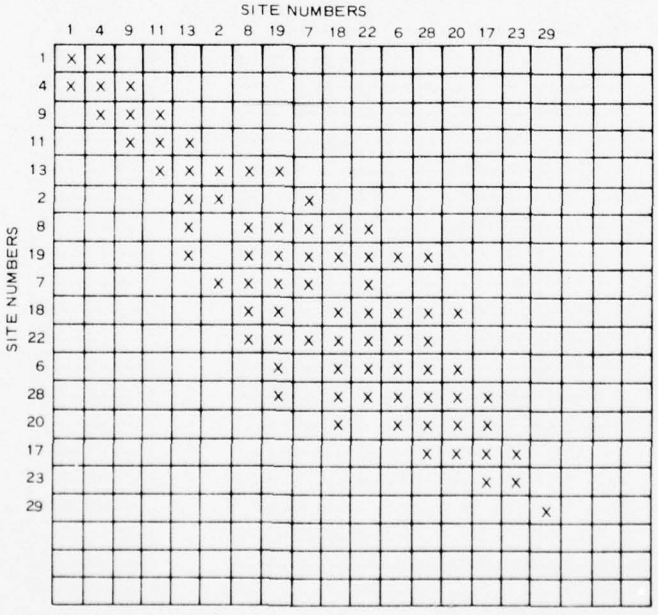


Figure 60. Diagonalized cluster matrix, 21 February 1975 data, with screening window of ±1.8 CCT units and variance window of ±1.5 CCT units

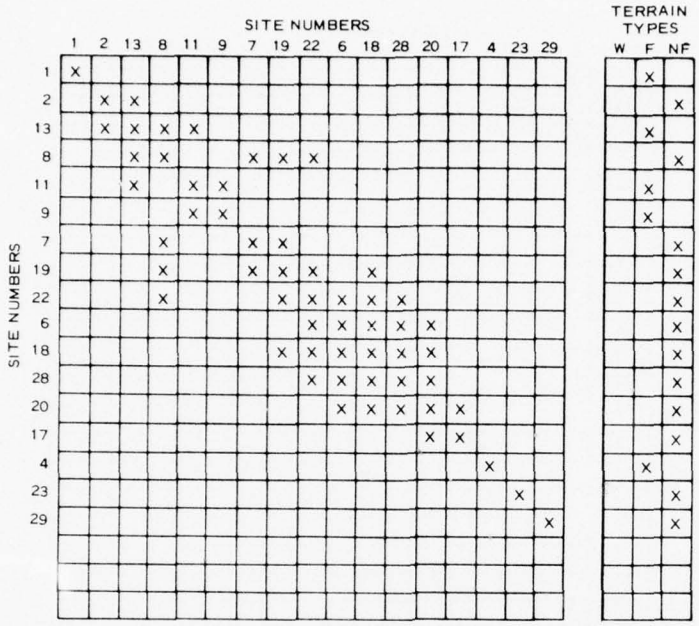


Figure 61. Cluster matrix, 21 February 1975 data, with screening window of ±1.8 CCT units and variance window of ±1.3 CCT units

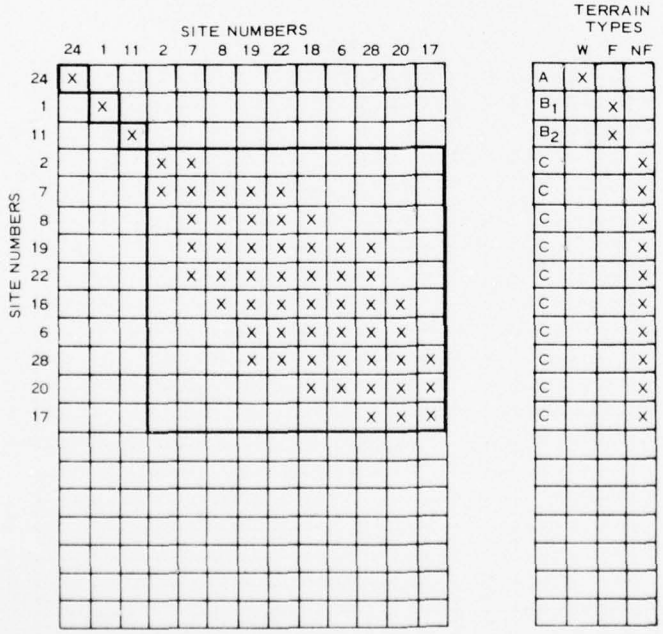


Figure 62. Cluster matrix, 21 February 1975 data, with screening window and variance of ±1.5 CCT units

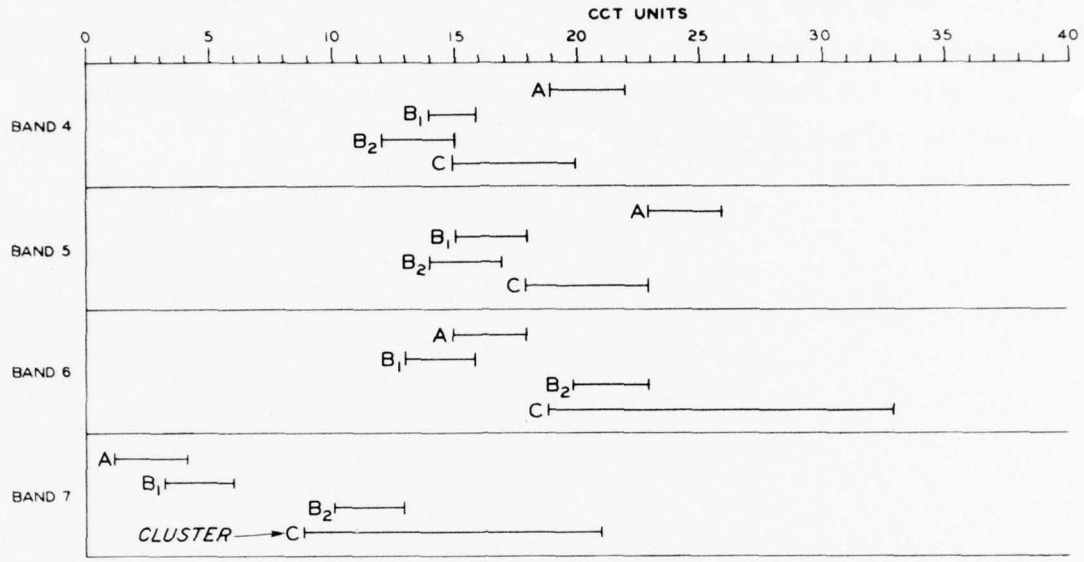


Figure 63. Graphic presentation of cluster class limits, 21 February 1975 data, with screening window and variance window of ±1.5 CCT units

of sharply separating cluster A from cluster B_1 and cluster B_2 from cluster C; and band 6 can accomplish the same separations as band 7. Thus, two possible cluster class sets exist that will separate the four clusters.

118. With the cluster class distributions (Figure 63) given, two stretched cluster class sets were determined (Table 24). A test indicated that cluster class set 2 was superior. The map of the study area produced with these cluster class limits (Figure 64) shows the two forest clusters B_1 and B_2 grouped together and displayed as one shade of gray.

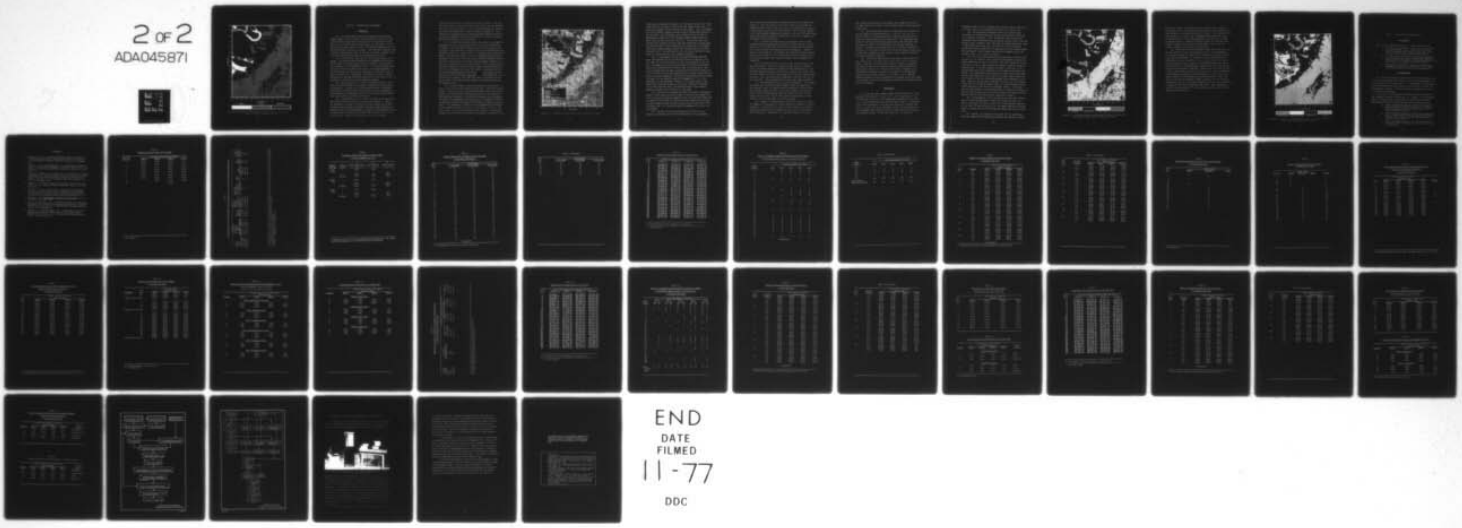
119. The formation of the two forest clusters, B_1 and B_2, rather than just one discrete cluster can be understood by referring to the cluster class limits in Table 24. It is evident from the table that signatures from Landsat pixels belonging to the forest cluster B_1 cannot be grouped with the "normal" forest cluster B_2, because their band-7 CCT values are too low and very near the values of water. On the other hand, they cannot be grouped with water either, since their band-5 CCT values

are much lower than those of water. The final interpretation of this situation was that the pixels that fit the B_1 cluster class set were a mixture of water and forest terrain types and thus formed naturally into a separate and discrete cluster. This is discussed further in greater detail in paragraphs 134-136.

AD-A045 871 ARMY ENGINEER WATERWAYS EXPERIMENT STATION VICKSBURG MISS F/G 15/4
ACQUISITION OF TERRAIN INFORMATION USING LANDSAT MULTISPECTRAL --ETC(U)
SEP 77 H STRUVE, W E GRABAU, H W WEST

UNCLASSIFIED WES-TR-M-77-2-2 NL

2 OF 2
ADA045871



END
DATE FILMED
11-77
DDC

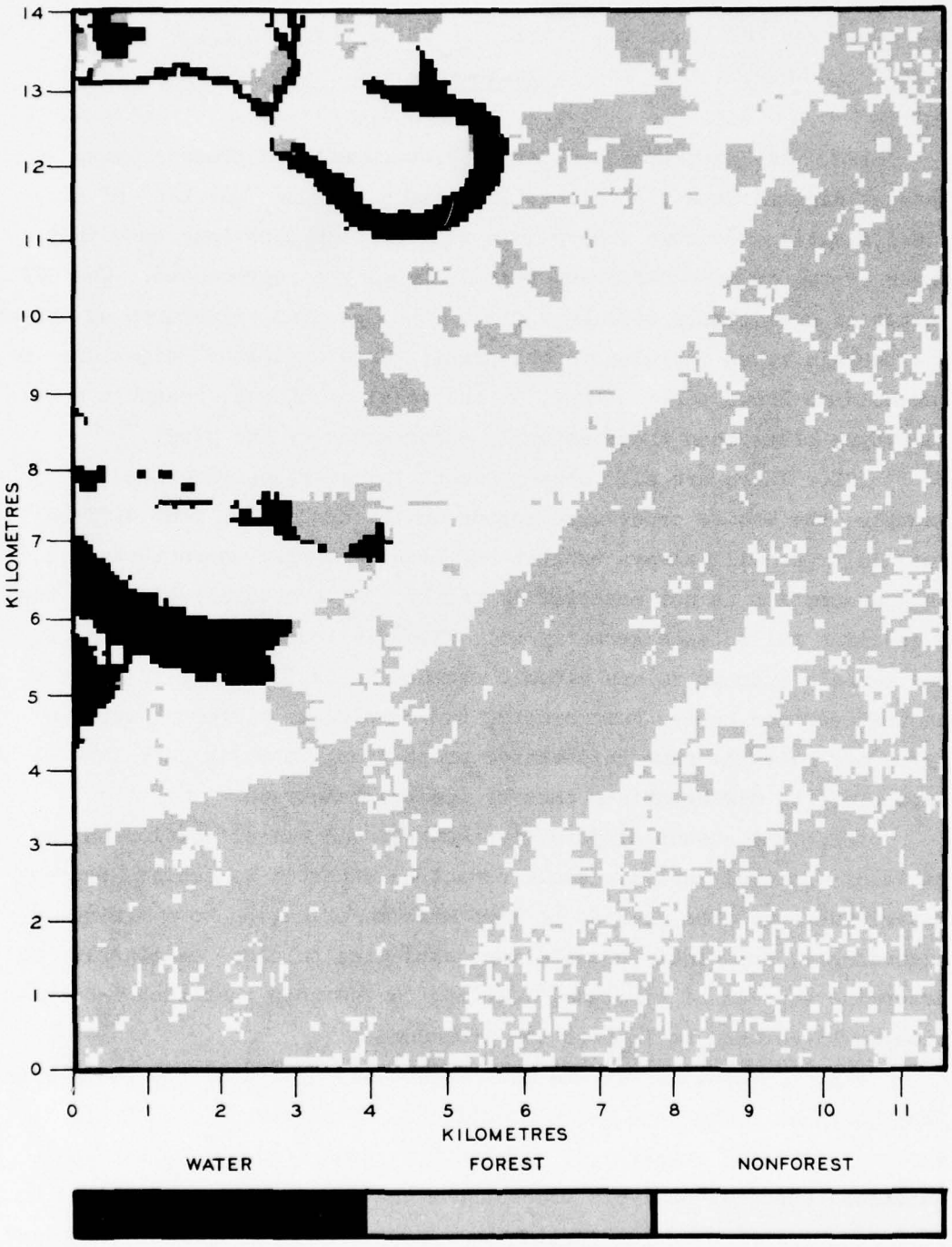


Figure 64. Study area showing interpretation of water, forest, and nonforest from 21 February 1975 data

PART III: EVALUATION AND ADJUSTMENTS

Evaluation

120. Any semiautomated terrain interpretation process using Landsat digital data will almost inevitably produce "patches" of misclassification, because some pixels will fall in locations such that parts of two or more different terrain types are represented. The CCT values of such pixels will be a mix of the spectral signatures of all of the terrain types included in the pixel. Such "spurious" signatures may resemble no known terrain type, or they may, by chance, resemble a terrain type other than those actually encompassed by the pixel.[4]

121. There are also other reasons for misclassification. For example, the entire procedure depends on the assumption that different terrain types will always exhibit different spectral characteristics. This assumption is not necessarily valid. It is entirely possible that a specific reflectance geometry will shift the spectral signature of one terrain type to values closely resembling or identical to those of another terrain type. Furthermore, two completely different terrain types may have the same reflectance properties, in which case it will be impossible to differentiate them by spectral analysis.

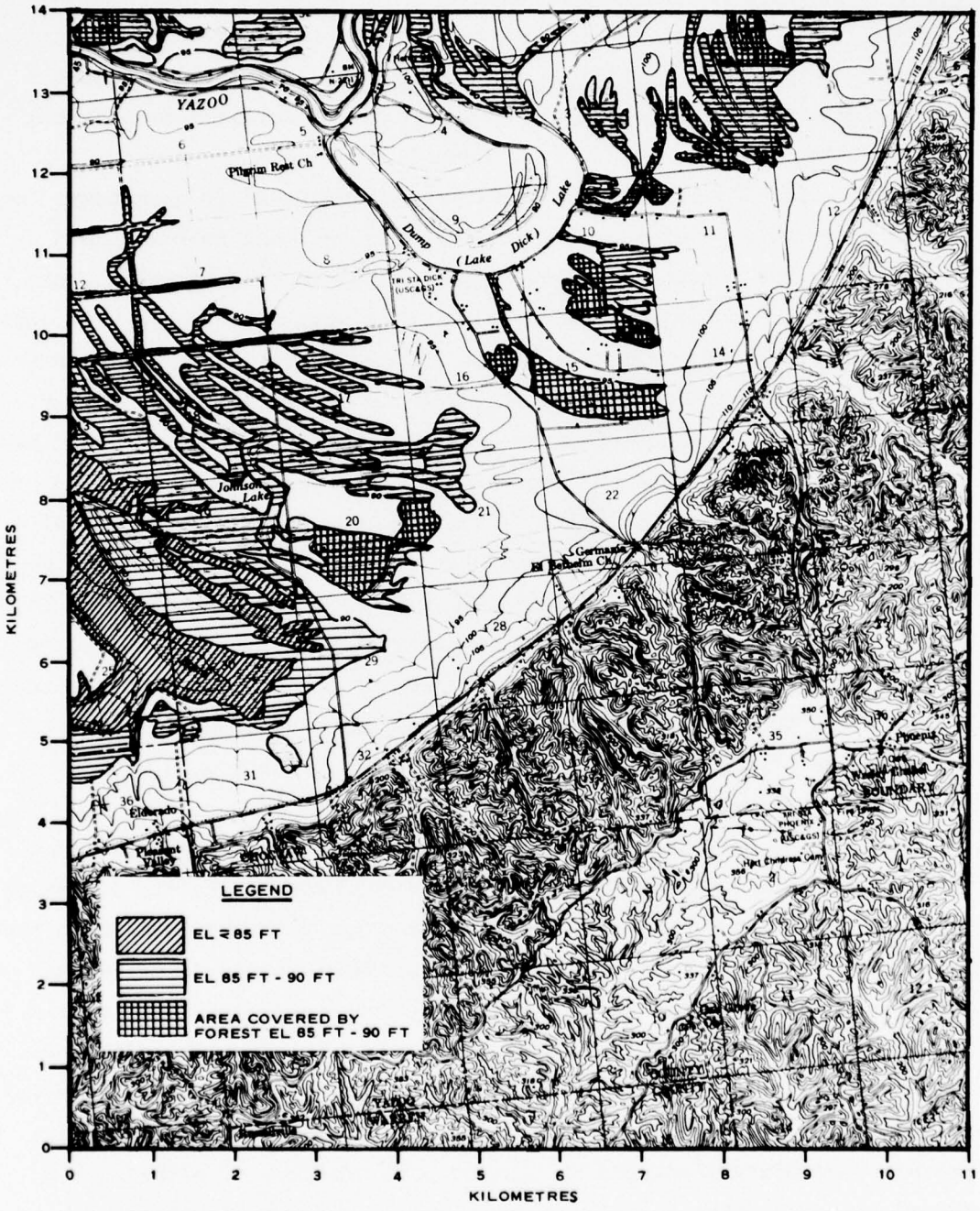
122. The essential point is that terrain classification on the basis of spectral analysis alone cannot be expected to achieve perfect discrimination. The goal is to keep misclassification to a minimum. Accordingly, the final products were subjected to close examination to determine sources of misclassification, in the hope that such information could be used to improve the procedure.

123. Unfortunately, the most recent air photos of the study area that could be found were taken in mid-July 1973 (Figure 5). Thus, it was impossible to compare the terrain type maps produced by the semiautomated spectral analysis procedure with actual ground conditions as revealed by air-photo interpretation. However, it was possible to visit the study area on the day of the October 1975 overpass and to record the conditions. Figures 16-20 illustrate some of the critical

conditions, and Figure 21 indicates their ground locations. The area was visited again on 8 May 1976, and critical points were photographed from the air at low altitude. A set of these photographs are included as Figures 22-30, and their locations indicated in Figure 31. The combination of ground truth data and low-altitude examination from the air provided detailed control on the locations of critical boundaries and the identification of critical areas.

124. In addition, local people were interviewed in an attempt to establish conditions at the times of the July 1974 and February 1975 overpasses. The conditions in every patch could not be recorded, but the crops in and conditions of certain fields were established with considerable reliability. Figures 12 and 13 illustrate much of this information. The major difficulty is that the floodplain was still partly inundated by the water of the great floods of 1974-75 (Figure 7) during the time of the July 1974 and February 1975 overpasses, and this is, of course, not visible on the July 1973 air photo. In an attempt to anticipate what might show on the two Landsat images, two water levels, one at 85 ft to correspond to the July 1974 overpass, and another at 90 ft to correspond approximately to the February [illegible] overflight, were projected onto the best available topographic [illegible] of the study area. Figure 65 shows the result. Note that large areas of floodplain on the western edge of the study area were subject to inundation at both the 85- and 90-ft levels.

125. Too much credence should not be placed in the precise locations of the limits of the projected flooded areas. There are two major reasons. First, a very small error in measuring elevations can move the position of a contour line a substantial distance in such flatland as the Mississippi floodplain. The existing topographic maps are largely the product of photogrammetric analysis, and it is extremely difficult to determine elevations accurately by that method in regions of very low relief. As a consequence, contour lines in the Mississippi floodplain are not uncommonly misplaced by tens or even hundreds of metres. The second reason is that the elevations plotted in Figure 7 reflect gage readings on the Yazoo River at Satartia. The gage is very close to the



Figures 65. Potentially flooded areas for two Yazoo River stages

study area, as indicated in Figure 4, but the water levels on the floodplain itself do not march to exactly the same rhythm as the river. Instead, they normally lag well behind the river gages. Thus, for example, the July 74 Landsat overpass occurred during a falling stage (Figure 7), and the water on the floodplain could be expected to be at some indeterminately higher level than the gage reading, simply because the water would not have had time to drain off. On the other hand, the February 1975 overpass occurred on what was essentially a rising stage, and thus the levels on the floodplain could be expected to be somewhat lower than the gage reading, because the low areas would not yet have had time to fill. The difficulty is that in neither case can the exact amount of lag be predicted with complete reliability.

126. As anticipated, the most striking differences among the three final products of the classification system, as applied to the test area, are in the delineation of water areas (Figures 51, 58, and 64). For example, the July 1974 map (Figure 58) shows very extensive water areas, especially at the western edge of the study area. When this is compared with the map of potentially flooded areas (Figure 65), it will be seen that the areas mapped as water coincide closely with the areas of possible inundation. The absolute distributions will not match, of course, for the reasons discussed in paragraph 125. Very close examination of the photomosaic (Figure 5) will reveal that the tiny water areas scattered through the forest area covering the southeast half of the study area and in the cultivated area in the southeast quadrant represent small reservoirs or farm ponds.

127. The map of the February 1975 data also shows a considerable amount of water along the western margin of the test area (Figure 64), but it is much less than that mapped from the July 1974 data (Figure 58). A comparison with the map of potential flooding (Figure 65) shows that the areas mapped as water coincide closely with the areas below the 85-ft contour line. For reasons discussed in paragraph 125, this is not unreasonable.

128. The map of the study area based on October 1975 data (Figure 51) indicates that the water had drained off the floodplain,

just as would be anticipated by considering the Yazoo River gage levels (Figure 7). Note also that the Yazoo River, which loops through the extreme northwest corner of the test area, is so narrow that it shows as a discontinuous line of pixels (see Reference 4 for a detailed explanation of the phenomena responsible); whereas, in the July 1974 and February 1975 scenes (Figures 58 and 64, respectively), the river is mapped as a continuous line of water pixels.

129. It may be concluded that the procedure provides a highly reliable method of mapping water surfaces, even when the water is quite turbid (as it is in the Yazoo River in time of flood) and very shallow (as it must have been over large areas of the floodplain). However, there are some anomalies. For example, the large area centered approximately at coordinates 3.5-7.0 is classified as forest in Figures 58 and 64, although it ought to be flooded according to the topographic map (Figure 65).

130. The difficulty lies in the fact that the two terrain types are not exclusive. The area is actually a large patch of floodplain forest, and ground examination revealed that the area beneath the trees had been inundated during the flood period. However, the Landsat sensor received radiation from the forest canopy and not from the water beneath the trees. Thus, the classification algorithm correctly classified the area as forest. Close examination will reveal other similar examples.

131. Some anomalies are not so readily explainable. For instance, the February and October 1975 maps (Figures 64 and 51, respectively) both indicate a patch of forest at about coordinates 6.0-10.0, but the July 1974 map (Figure 58) shows none at that location. The 1973 air photo (Figure 5) indicates that there was a patch of forest at that location in July 1973. Why did the algorithm misclassify it as non-forest in the July 1974 data? It is not yet known.

132. However, the Landsat maps do reveal the removal of forest. For example, the large patch of forest centered on coordinates 7.6-12.0 (Figure 51) is indicated on the 1973 air photo (Figure 5) as having a long tongue extending to the eastward, but none of the Landsat scenes show its presence. Examination on the ground revealed that it had

been cleared and converted to agricultural land (Figures 14 and 15b). The same situation is evident at the southeast corner of the large patch (Figure 15a).

133. The flooded forest region at coordinates 3.7-7.0 also contains an example of misclassification. In the July 1974 map (Figure 58), there is a discontinuous line of nonforest pixels along the western and southern edges of the forest patch. The topographic map (Figure 9) indicates no elevation changes that could bring this small zone above the level of the floodwaters; the area must have been flooded. The anomaly is probably the result of spurious signatures.[4] The pixels in this region fell partly on forest and partly on water, and by chance the integration of the radiance values produced a signature that met the nonforest criteria.

134. Why did the same phenomenon not occur in the February 1975 data (Figure 64)? A reference to Figure 63 will reveal the probable reason. Note that one of the forest clusters (cluster B_1) exhibits CCT values in band 7 so low that they overlap the CCT values of the water cluster (cluster C). Further, even in band 5 (Figure 64 was constructed with cluster class set 2; see Table 24), the separation between water and forest clusters is not large. Thus, a spurious signature caused by a pixel falling across a boundary is far more likely to be classed as one of the two elements comprising the combination than some third category.

Adjustments

135. The fact that a forest cluster looked so much like a water cluster in the February 1975 data led to a close examination of the CCT value data (Table 21). This examination revealed that four sites, 1, 3, 4, and 24, exhibited very low CCT values in band 7. An examination of the photomosaic of the study area (Figure 33) indicates that sites 1, 3, and 4 are all in floodplain forest, and an examination of the topographic maps (Figures 9 and 65) reveals that those three sites were probably flooded. On the other hand, site 2 is also in a

floodplain forest, but the topographic map indicates that it may be on a low ridge so that it was probably above water on 21 February 1975.

136. The Mississippi floodplain forests are composed almost exclusively of broadleaf deciduous species, and they would be expected to be leafless in February. Thus, the radiation from the surface in floodplain forest areas would come only partly from the trees; some, and perhaps a major component, would come from the surface beneath the trees. If that surface were water, then the band-7 CCT values of flooded forest areas would be expected to be abnormally low, but somewhat higher than open water. This is exactly the condition suggested by the CCT values of sites 1, 3, 4, and 24 in band 7.

137. This led to the possibility that open water, flooded forest, and dry (i.e. not flooded) forest could be discriminated by a more detailed subdivision of band-7 CCT values. With this in mind, the cluster class sets in Table 25 were selected for testing. Figure 66 shows the resulting map. It is quite apparent that many, if not most, of the forested areas that could be expected to be flooded are correctly classified. The scattering of small patches of flooded forest mixed among the dry forest almost all occur close to the curved border between the hilly country to the east and the Mississippi River floodplain to the west. Since the sun was only 55 deg above the horizon to the southeast at the time of the Landsat overpass (Table 2), these obviously misclassified areas may be the result of topographic effects. The terrain tends to slope steeply to the northeast along the valley wall (Figure 9), and this would produce very low CCT values, which would create the same general effect on CCT values as water, especially in band 7.

138. An examination of the site data for the July 1974 overpass does not reveal any depression of band-7 CCT values for sites 1, 3, and 4. The reason, of course, is that in July the deciduous trees are in full leaf, and thus virtually all of the energy reaching the Landsat sensor is coming from the forest canopy and almost none from the forest floor.

139. However, the experience with band-7 data in the February 1975 data suggested that anything that modified the open-water surface

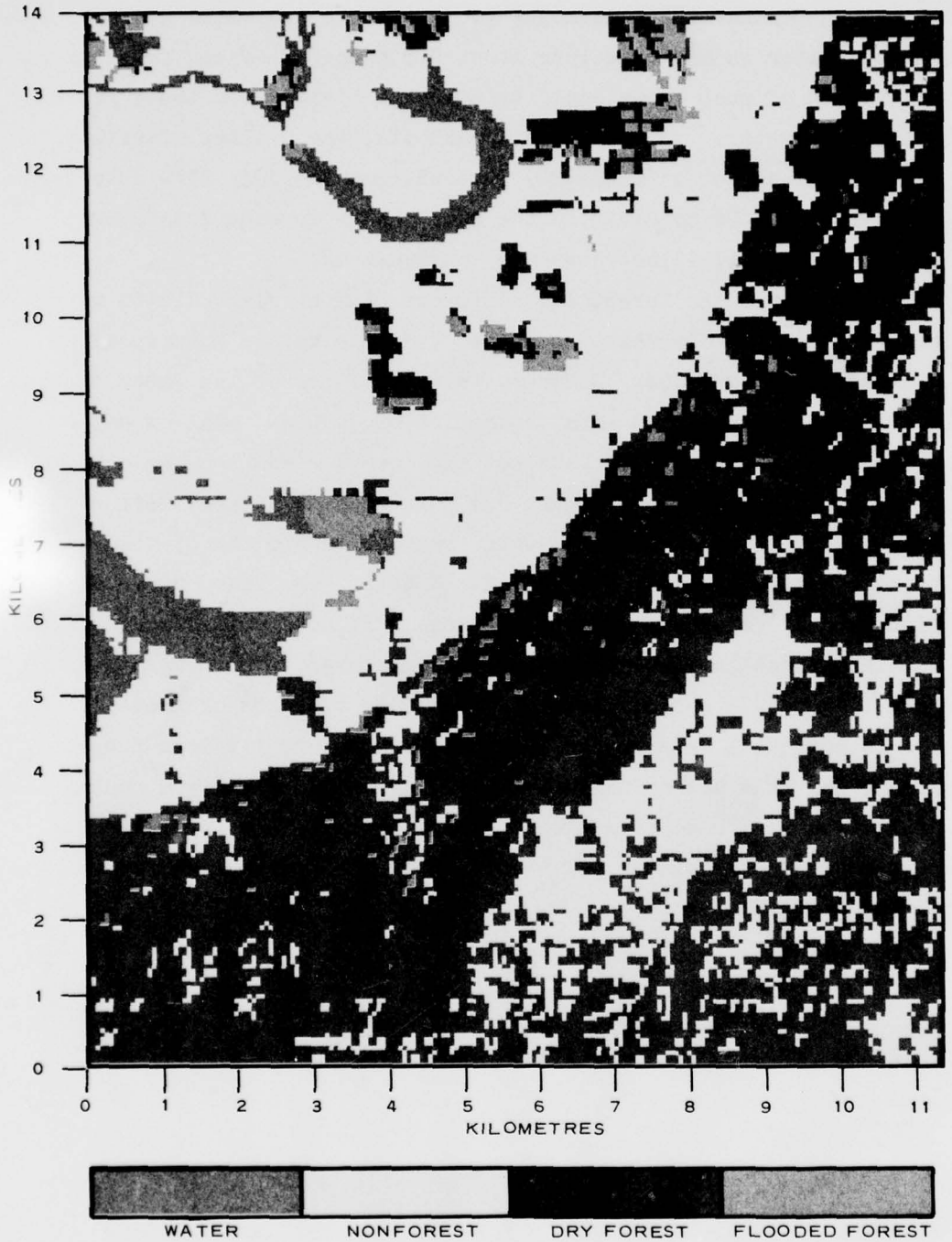


Figure 66. Study area showing interpretation of water, nonforest, and two types of forest from 21 February 1975 data

might be detectable. For example, by July, weeds and some crops are tall enough so that the tops might project above the water surface, provided the water is not more than about 0.5 m deep. If so, then the band-7 CCT values of such areas would be somewhat higher than those representing deep water. To test this hypothesis, the cluster class limits of cluster class set 1 (Table 20) derived from the July 1974 data were altered in Table 26 by dividing the water range in band 7 into two classes. Figure 67 illustrates the consequences.

140. Since the forest and nonforest cluster class limits were not altered, these two classes were mapped the same as the corresponding two classes in Figure 58. However, the alteration of the water cluster class limits has resulted in a separation of deep and shallow water. The deepwater regions, Dump Lake and the lowest elevation areas (below the 85-ft contour lines in Figure 9), have been delineated very effectively, and most of the shallow-water regions (above the 85-ft contour) have also been accurately classified. However, the Yazoo River is classified as shallow water, which it obviously is not. This is undoubtedly a misclassification caused by spurious signatures. The river is quite narrow (Figure 25) and bordered on both sides by bands of trees. The river is not quite wide enough to permit a pixel to fit into the water surface, and thus all of the pixels falling on the river are contaminated by radiation reflected from the bordering trees. This is apparently enough to raise the band-7 CCT values above the limits used for shallow and deep water (Table 26).

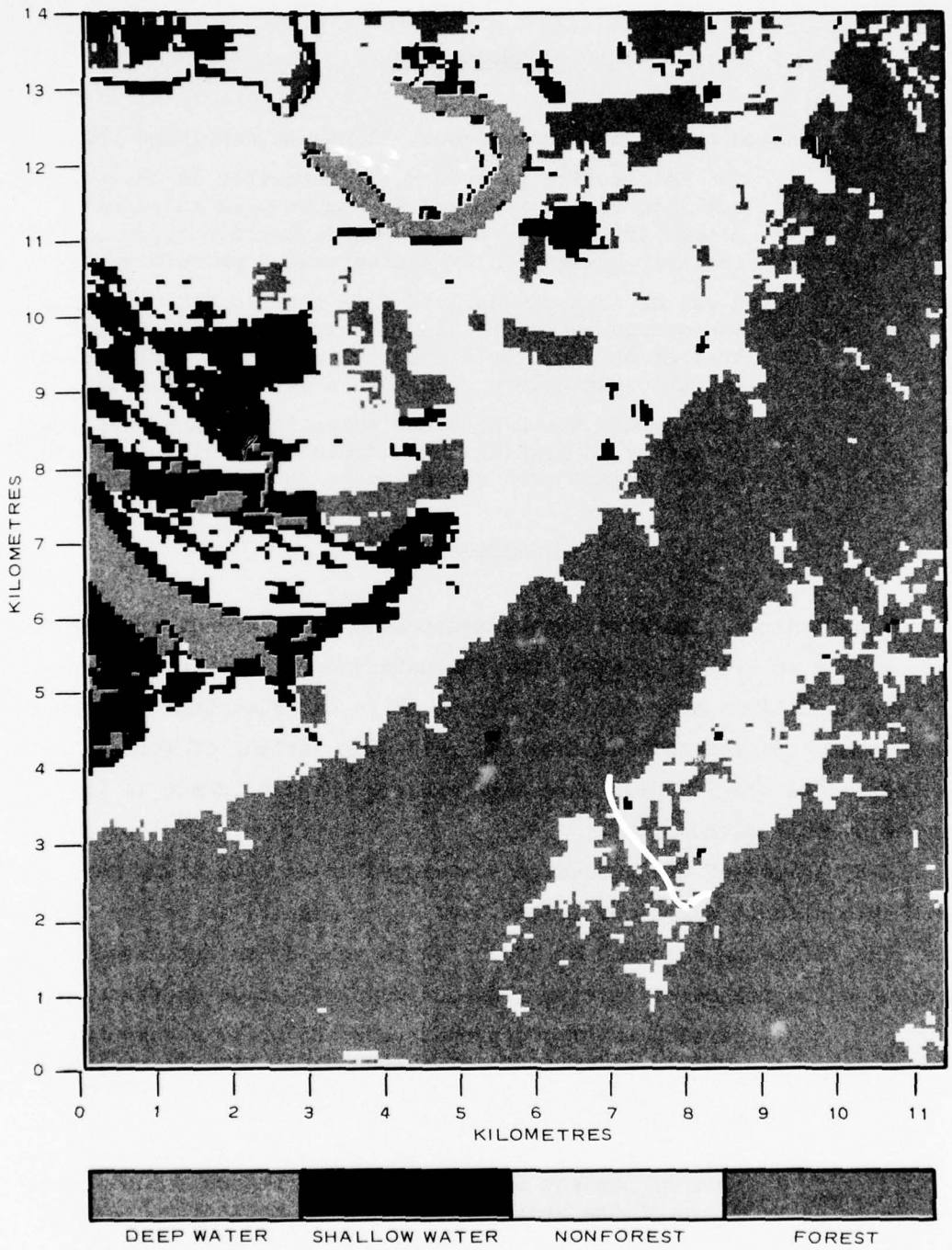


Figure 67. Study area showing interpretation of two types of water, nonforest and forest from 11 July 1974 data

PART IV: CONCLUSIONS AND RECOMMENDATIONS

Conclusions

141. Based on previous experience, it may be concluded that:

a. The interactive procedure, as described in the text and outlined in Plates 1 and 2, can be used to quickly obtain criteria by which certain terrain types can be reliably classified by semiautomated procedures.

b. A set of criteria derived from terrain data obtained at one season cannot be used to classify data obtained at another season. Both spectral reflectance characteristics and absolute colors change on a seasonal basis.

c. Some ground truth data are essential, since they are necessary in judging the acceptability of the clusters of CCT values used as the basis for classification.

Recommendations

142. Since this study has demonstrated that Landsat spectral data can be used to determine areal terrain data and related information on in situ conditions over large areas, it is recommended that the procedure be further demonstrated and assessed in a wide variety of terrain and environmental conditions. This will enhance the confidence in the use of the overall methodology.

143. Some of the procedures developed under this study have not been automated to permit efficient use. As a result, it is recommended that the following research be conducted to provide an automated or semiautomated method of interpreting and mapping worldwide terrain:

a. Develop automated procedures, and write a computer program to permit more efficient extraction of radiance data on preselected terrain features by the use of multipass Landsat data sets.

b. Design computer procedures to analytically compare two sets of Landsat data for the same region for different times of the year and to output a map portraying the changes in the conditions.

c. Develop automated procedures for selecting gray-tone densities for better map portrayal of terrain and land-use conditions.

REFERENCES

1. Friesz, R. R. et al., "European Waterways Study, A Procedure for Describing Tactical Gaps," Technical Report M-70-12, Jul 1970, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss.

2. Link, L. E., Jr., and Shamburger, J. H., "Application of Remote Sensors to Army Facility Management," Technical Report M-74-2, Feb 1974, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss.

3. Struve, H., Grabau, W. E., and West, H. W., "Acquisition of Terrain Information Using Landsat Multispectral Data; Correction of Landsat Spectral Data for Extrinsic Effects," Technical Report M-77-2, Report 1, Jun 1977, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss.

4. Grabau, W. E., "Pixel Problems," Miscellaneous Paper M-76-9, May 1976, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss.

5. Scott, F. T. et al., "Soil Survey of Yazoo County, Mississippi," Jun 1975, Soil Conservation Service, Department of Agriculture, in cooperation with the Mississippi Agricultural and Forestry Experiment Station, Government Printing Office, Washington, D. C.

6. Fenneman, N. M., Physiography of Eastern United States, 1st ed., McGraw-Hill, New York, 1938.

7. West, H. W. and Floyd, H. M., "An Automated System for Collecting, Processing, and Displaying Environmental Baseline Data," Technical Report M-76-11, Nov 1976, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss.

8. Kennedy, J. G. and Williamson, A. N., "A Technique for Achieving Geometric Accordance of Landsat Digital Data," Miscellaneous Paper M-76-16, Jul 1976, U. S. Army Engineer Waterways Experiment Station, CE, Vicksburg, Miss.

Table 1

Subjectively Chosen CCT Value Classes*

CCT Value Class No.	CCT Value Ranges/Spectral Band			
	Band 4	Band 5	Band 6	Band 7
1	22-26	15-18	17-23	5-11
2	26-28	18-21	23-26	11-14
3	28-31	21-23	26-28	14-18
4	31-33	23-26	28-33	18-23
5	33-35	26-30	33-35	23-25
6	35-37	30-33	35-40	25-31
7	37-40	33-36	40-42	31-39
8		36-41	42-46	39-48
9			46-50	

* See Figure 2.

Table 2

Characteristics of the Three Landsat Spectral Data Sets for Area Near Vicksburg, Mississippi

Landsat Scene No.*	Coordinates Of Center Of Scanned Area		Date of Satellite Overpass	Time of Satellite Overpass (Local Times) hr,min,sec	Image and Data Quality (0 = inferior, 9 = excellent)	Cloud Cover %	Haze Conditions	Sun Azimuth deg N	Sun Elevation deg from Zenith
	Latitude	Longitude							
1718-15590**	33°11' N	90°19' W	11 Jul 74	10,59,0(CDT)	8	0	Haze-free	103	31
2030-15555†	33°06' N	90°06' W	21 Feb 75	9,55,50(CST)	8	0	Haze-free	136	55
2264-15540†	33°02' N	90°08' W	13 Oct 75	10,54,00(CDT)	8	0	Haze-free	141	49

* The three digits preceeding the hyphen represent the number of days since launch; the number following the hyphen is Greenwich time in hours, minutes, and tens of seconds.

** Landsat 1 (launched on 23 July 1973).

† Landsat 2 (launched on 22 January 1975).

Table 3

Calculated Flight Path Angles and Pixel Sizes for Three Landsat Data Sets

Landsat Data Set	Solution Set*	θ_f (Flight Path Angle) deg	D_X (Pixel Width) m	D_Y (Pixel Length) m
11 Jul 1974	1	9.03	58.3	
	2	9.13	____	80.9
	Average	9.08	58.3	80.9
21 Feb 1975	1	11.00	57.5	
	2	9.13	____	80.7
	Average	10.07	57.5	80.7
13 Oct 1975	1	9.03	58.3	
	2	8.59	____	78.9
	Average	8.81	58.3	78.9

* Solution sets 1 and 2 to the simultaneous equations exist since there were three unknowns to be determined from four equations.

Table 4

Terrain Conditions* at Sample Sites at the Time of the Three Overflights

Site No.	Overflights		
	11 Jul 1974	21 Feb 1975	13 Oct 1975
1	F_1	F_1	F_1
2	F_1	A_6	A_1
3	F_1	F_1	F_1
4	F_1	F_1	F_1
5	A_1	A_6	A_2
6	A_1	A_6	A_2
7	A_1	A_6	A_2
8	A_1	A_6	A_2
9	F_3	F_3	F_3
10	F_3	F_3	F_3
11	F_3	F_3	F_3
12	F_2	F_2	F_2
13	F_2	F_2	F_2
14	F_2	F_2	F_2
15	P_2	P_2	P_2
16	P_2	P_2	P_2
17	P_1	P_1	P_1
18	A_7	A_6	A_3
19	A_7	A_6	A_3
20	A_7	A_6	A_1
21	A_3	A_6	A_7
22	A_3	A_6	A_3
23	A_3	A_4	A_7
24	W	W	W
25	A_3	A_5	A_3

(Continued)

* See paragraph 56 for key to terrain conditions.

Table 4 (Concluded)

Site No.	Overflights		
	11 Jul 1974	21 Feb 1975	13 Oct 1975
26	A_1	A_6	A_1
27	A_1	A_6	A_1
28	A_3	A_6	A_3
29	A_1	A_6	A_7
30	A_3	A_6	A_3

Table 5

Signatures from Sample Sites, 13 October 1975

Site No.*	Mean and Standard Deviations,** CCT Units			
	Band 4	Band 5	Band 6	Band 7
1	15.8(0.7)	14.8(1.0)	25.6(1.0)	14.0(0.9)
2	20.2(0.7)	22.6(0.7)	35.1(1.5)	23.1(1.5)
3	15.4(0.7)	14.4(0.5)	27.4(1.9)	16.6(1.7)
4	16.1(0.6)	15.8(0.4)	24.7(0.9)	13.2(0.8)
5	21.8(1.2)	25.9(1.8)	29.8(0.7)	16.4(0.7)
6	18.9(1.1)	21.6(0.9)	32.2(2.1)	20.9(2.5)
7	20.4(1.7)	25.0(1.9)	29.9(0.9)	17.6(1.5)
8	25.8(0.8)	32.2(0.5)	33.0(1.2)	16.9(0.6)
9	14.7(0.9)	14.7(0.5)	26.8(1.5)	17.2(1.6)
10	15.0(0.9)	14.3(1.0)	23.4(1.4)	13.9(1.2)
11	13.9(0.6)	14.0(0.9)	24.6(1.9)	15.4(2.5)
12	15.6(1.0)	14.3(0.7)	24.9(0.8)	14.8(1.0)
13	16.3(0.5)	16.0(1.5)	26.1(1.6)	15.7(0.9)
14	17.6(2.3)	20.7(3.9)	28.3(2.4)	16.7(1.7)
15	19.0(0.6)	22.2(1.0)	29.8(0.8)	19.0(0.6)
16	17.8(1.6)	19.7(2.4)	29.2(1.2)	18.0(1.1)
17	21.2(1.2)	24.3(1.3)	32.0(2.2)	19.6(1.8)
18	21.3(1.1)	24.2(1.1)	29.4(0.5)	20.7(0.5)
19	19.7(0.7)	23.1(0.8)	31.4(1.1)	16.9(0.9)
20	23.4(0.9)	28.4(0.7)	30.9(0.6)	16.0(0.7)
21	23.6(0.9)	28.7(0.9)	31.7(0.7)	17.3(1.0)
22	21.1(1.1)	26.6(1.0)	31.4(0.7)	19.6(0.5)
23	25.6(0.5)	33.6(0.7)	35.2(0.7)	17.8(0.8)
24	18.1(1.2)	16.8(1.2)	12.6(1.9)	2.2(1.6)
25	25.2(0.4)	32.1(0.7)	34.2(1.2)	18.4(0.7)
26	19.7(0.7)	24.6(2.1)	34.2(1.0)	22.3(0.9)
27	19.0(0.9)	20.8(0.8)	31.8(0.7)	21.3(0.7)
28	21.4(0.9)	24.8(0.7)	35.0(0.7)	21.9(0.8)
29	24.0(1.3)	28.6(2.6)	32.1(1.5)	17.7(1.7)
30	20.7(1.0)	24.6(0.5)	31.8(1.1)	19.3(0.5)

* See Figure 33 for the locations of these sites.

** Mean and standard (in parentheses) deviations are for a 3- by 3-pixel array.

Table 6

Effect of Changing Screening Window Criterion on Number of Acceptable Sample Sites, 13 October 1975 Data

Sample Site	Screening Window, CCT Units				
	±1.3	±1.5	±1.6	±1.7	±1.8
1	X	X	X	X	X
2		X	X	X	X
3					
4	X	X	X	X	X
5					X
6					
7					
8	X	X	X	X	X
9			X	X	X
10		X	X	X	X
11					
12	X	X	X	X	X
13			X	X	X
14					
15	X	X	X	X	X
16					
17					
18	X	X	X	X	X
19	X	X	X	X	X
20	X	X	X	X	X
21	X	X	X	X	X
22	X	X	X	X	X
23	X	X	X	X	X
24					
25	X	X	X	X	X

(Continued)

Table 6 (Concluded)

Sample Site	Screening Window, CCT Units				
	±1.3	±1.5	±1.6	±1.7	±1.8
26					
27	X	X	X	X	X
28	X	X	X	X	X
29					
30	X	X	X	X	X
Total number meeting criterion	15	17	19	19	20

Table 7

Effects of Variance Window on CCT Class Limits

13 October 1975 Data

No.*	Window	Class Limits,* CCT Units Spectral Band			
		Band 4	Band 5	Band 6	Band 7
1	1.3	15-17	14-16	24-27	13-15
	1.5	14-17	13-16	24-27	13-15
	1.8	14-18	13-17	24-27	12-16
2	1.3	19-21	21-24	34-36	22-24
	1.5	19-22	21-24	34-37	22-25
	1.8	18-22	21-24	33-37	21-25
4	1.3	15-17	15-17	23-26	12-14
	1.5	15-18	14-17	23-26	12-15
	1.8	14-18	14-18	23-26	11-15
5	1.3	21-23	25-27	29-31	15-18
	1.5	20-23	24-27	28-31	15-18
	1.8	20-24	24-28	28-32	15-18
8	1.3	25-27	31-33	32-34	16-18
	1.5	24-27	31-34	32-34	15-18
	1.8	24-28	30-34	31-35	15-19
9	1.3	13-16	13-16	26-28	16-18
	1.5	13-16	13-16	25-28	16-19
	1.8	13-16	13-16	25-29	15-19
10	1.3	14-16	13-16	22-25	13-15
	1.5	14-16	13-16	22-25	12-15
	1.8	13-17	13-16	22-25	12-16
12	1.3	14-17	13-16	24-26	14-16
	1.5	14-17	13-16	23-26	13-16
	1.8	14-17	13-16	23-27	13-17
13	1.3	15-18	15-17	25-27	14-17
	1.5	15-18	15-17	25-28	14-17
	1.8	15-18	14-18	24-28	14-17
15	1.3	18-20	21-23	29-31	18-20
	1.5	18-20	21-24	28-31	18-20
	1.8	17-21	20-24	28-32	17-21
18	1.3	20-23	23-25	28-31	19-22
	1.5	20-23	23-26	28-31	19-22
	1.8	20-23	22-26	28-31	19-22

(Continued)

* Sites selected using a screening window of ±1.8 CCT units.

Table 7 (Concluded)

Site No.	Variance Window	Class Limits, CCT Units Spectral Band			
		Band 4	Band 5	Band 6	Band 7
19	1.3	18-21	22-24	30-33	16-18
	1.5	18-21	22-25	30-33	15-18
	1.8	18-21	21-25	30-33	15-19
20	1.3	22-25	27-30	30-32	15-17
	1.5	22-25	27-30	29-32	15-17
	1.8	22-25	27-30	29-33	14-18
21	1.3	22-25	27-30	30-33	16-19
	1.5	22-25	27-30	30-33	16-19
	1.8	22-25	27-30	30-33	16-19
22	1.3	20-22	25-28	30-33	18-21
	1.5	20-23	25-28	30-33	18-21
	1.8	19-23	25-28	30-33	18-21
23	1.3	24-27	32-35	34-36	17-19
	1.5	24-27	32-35	34-37	16-19
	1.8	24-27	32-35	33-37	16-20
25	1.3	24-26	31-33	33-35	17-20
	1.5	24-27	31-34	33-36	17-20
	1.8	23-27	30-34	32-36	17-20
27	1.3	18-20	20-22	31-33	20-23
	1.5	18-20	19-22	30-33	20-23
	1.8	17-21	19-23	30-34	20-23
28	1.3	20-23	24-26	34-36	21-23
	1.5	20-23	23-26	34-36	20-23
	1.8	20-23	23-27	33-37	20-24
30	1.3	19-22	23-26	31-33	18-21
	1.5	19-22	23-26	30-33	18-21
	1.8	19-22	23-26	30-34	18-21

Table 8

Terrain Type Assignment of Sites in Cluster Matrix*

13 October 1975 Data

Site No.	Terrain Types		
	Forest	Nonforest	Water
1	X		
2		X	
4	X		
8		X	
10	X		
12	X		
15		X	
18		X	
19		X	
20		X	
21		X	
22		X	
23		X	
25		X	
27		X	
28		X	
30		X	

* See Figure 42.

Table 9

Terrain Types Related to Site Numbers,
13 October 1975 Data

Site No.	Terrain Types			Cluster
	Forest	Nonforest	Water	
1	X			I
4	X			I
10	X			I
12	X			I
2		X		II
28		X		II
8		X		III
23		X		III
25		X		III
15		X		IV
18		X		IV
27		X		IV
30		X		IV
22		X		IV
19		X		IV
20		X		IV
21		X		IV

Table 10

CCT Classes Defining Spectral Signatures of Sites

Used in Cluster Matrix, Figure 42

13 October 1975 Data

No.	Class Limits,* CCT Units Spectral Band				Cluster
	Band 4	Band 5	Band 6	Band 7	
1	14-17	13-16	24-27	13-15	I
4	15-18	14-17	23-26	12-15	
10	14-16	13-16	22-25	12-15	
12	14-17	13-16	23-26	13-16	
2	19-22	21-24	34-37	22-25	II
28	20-23	23-26	34-36	20-23	
8	24-27	31-34	32-34	15-18	III
23	24-27	32-35	34-37	16-19	
25	24-27	31-34	33-36	17-20	
15	18-20	21-24	28-31	18-20	IV
18	20-23	23-26	28-31	19-22	
27	18-20	19-22	30-33	20-23	
30	19-22	23-26	30-33	18-21	
22	20-23	25-28	30-33	18-21	
19	18-21	22-25	30-33	15-18	
20	22-25	27-30	29-32	15-17	
21	22-25	27-30	30-33	16-19	

* Screening window = $\pm$1.5 CCT units; variance window = $\pm$1.5 CCT units.

Table 11

CCT Classes Defining Spectral Signatures of Sites

Used in Cluster Matrix, Figure 41

13 October 1975 Data

Site No.	Class Limits,* CCT Units Spectral Band			
	Band 4	Band 5	Band 6	Band 7
1	14-17	13-16	24-27	12-16
2	19-22	21-24	33-37	21-25
4	14-18	14-17	23-26	12-15
8	24-27	31-34	31-35	15-19
10	13-17	13-16	22-25	12-16
12	14-17	13-16	23-27	13-16
15	17-21	21-24	28-31	17-21
18	20-23	23-26	28-31	19-22
19	18-21	21-25	30-33	15-19
20	22-25	27-30	29-33	14-18
21	22-25	27-30	30-33	16-19
22	19-23	25-28	30-33	18-21
23	24-27	32-35	34-37	16-19
25	24-27	30-34	33-36	17-20
27	17-21	19-22	30-33	20-23
28	20-23	23-26	33-37	20-24
30	19-22	23-26	30-33	18-21

* Screening window = $\pm$1.5 CCT units; variance window = $\pm$1.7 CCT units.

Table 12

Clusters and Associated CCT Class Ranges*

13 October 1975 Data

Cluster**	Site No.	CCT Value Ranges Spectral Band			
		Band 4	Band 5	Band 6	Band 7
A	24	16-20	15-18	11-14	1-4
Cluster Class Limits		16-20	15-18	11-14	1-4
B	1	14-17	13-16	24-27	12-16
	4	14-18	14-17	23-26	12-15
	10	13-17	13-16	22-25	12-16
	12	14-17	13-16	23-27	13-16
Cluster Class Limits		13-18	13-17	22-27	12-16
C	2	19-22	21-24	33-37	21-25
	8	24-27	31-34	31-35	15-19
	15	17-21	21-24	28-31	17-21
	18	20-23	23-26	28-31	19-22
	19	18-21	21-25	30-33	15-19
	20	22-25	27-30	29-33	14-18
	21	22-25	27-30	30-33	16-19
	22	19-23	25-28	30-33	18-21
	23	24-27	32-35	34-37	16-19
	25	24-27	30-34	33-36	17-20
	27	17-21	19-22	30-33	20-23
	28	20-23	23-26	33-37	20-24
	30	19-22	23-26	30-33	18-21
Cluster Class Limits		17-27	19-35	28-37	14-25

* Variance window = $\pm$1.7 CCT units.

** See Figure 44.

Table 13

Application of the Cluster Group Stretching Rules to the Cluster Separation Logic (A)|(B|C) (7 5)

Cluster	Cluster Class Limits, CCT Units, Spectral Band			
	Band 4	Band 5	Band 6	Band 7
	Before Application of Rules			
A	16-20	15-18	11-14	1-4
B	13-18	13-17	22-27	12-16
C	17-27	19-35	28-37	14-25
	After Omnigroup Rules			
A	0-63	15-18	0-63	1-4
B	0-63	13-17	0-63	12-16
C	0-63	19-35	0-63	14-25
	After Intragroup Rules			
A	0-63	0-63	0-63	1-4
B	0-63	0-18	0-63	12-16
C	0-63	19-63	0-63	14-25
	After Intergroup Rules 1 and 2			
A	0-63	0-63	0-63	0-4
B	0-63	0-18	0-63	12-16
C	0-63	19-63	0-63	14-63
	After Intergroup Rule 3			
A	0-63	0-63	0-63	0-4
B	0-63	0-18	0-63	12-63
C	0-63	19-63	0-63	12-63
	After Intergroup Rule 4			
A	0-63	0-63	0-63	0-8
B	0-63	0-18	0-63	9-63
C	0-63	19-63	0-63	9-63

Table 14

Stretched Cluster Class Sets, 13 October 1975

Cluster	Cluster Class Limits, CCT Units Spectral Band			
	Band 4	Band 5	Band 6	Band 7
		Cluster Class Set 1		
A	0-63	0-63	0-18	0-63
B	0-63	0-63	19-27	0-63
C	0-63	0-63	28-63	0-63
		Cluster Class Set 2		
A	0-63	0-63	0-63	0-8
B	0-63	0-18	0-63	9-63
C	0-63	19-63	0-63	9-63
		Cluster Class Set 3		
A	0-63	0-63	0-63	0-8
B	0-63	0-63	0-27	9-63
C	0-63	0-63	28-63	9-63
		Cluster Class Set 4		
A	0-63	0-63	0-18	0-63
B	0-63	0-18	19-63	0-63
C	0-63	19-63	19-63	0-63

Table 15

Comparison of Areas of Classes, Automated Interpretations Versus Air-Photo Interpretation

Terrain Type	Air-Photo Interpretation, %	Automated Interpretations (%) by Cluster Class Sets			
		1	2	3	4
Water	1.6	1.2* (25)**	1.3 (19)	1.4 (13)	1.1 (31)
Forest	50.1	50.9 (2)	52.4 (5)	50.7 (1)	52.4 (5)
Nonforest	48.3	47.9 (1)	46.3 (4)	47.9 (1)	46.5 (4)

* Values indicate percentage of total area of the study area.
** Values in parentheses indicate percent "error."

Table 16

Signatures from Sample Sites, 11 July 1974

Site No.*	Mean and Standard Deviations,** CCT Units			
	Band 4	Band 5	Band 6	Band 7
1	25.0(0.)	19.1(1.0)	38.7(0.5)	31.4(1.4)
2	25.4(0.9)	19.0(0.9)	38.7(0.5)	31.8(1.3)
3	25.0(0.)	19.1(0.9)	39.2(0.7)	33.3(1.0)
4	26.0(1.1)	19.2(1.0)	40.4(1.3)	33.8(1.2)
5	32.0(1.7)	30.6(1.3)	41.0(0.9)	29.3(1.3)
6	36.0(1.1)	35.4(1.8)	44.0(0.7)	32.3(1.3)
7	30.8(1.5)	29.0(1.9)	40.7(1.0)	28.4(1.8)
8	34.8(1.5)	35.1(1.9)	39.2(1.3)	24.9(1.1)
9	24.6(0.7)	16.7(0.9)	38.8(0.8)	33.6(0.7)
10	25.0(0.)	17.1(0.8)	38.8(0.8)	33.0(1.7)
11	24.7(0.5)	16.3(0.5)	37.2(0.8)	31.1(2.2)
12	24.7(0.5)	17.4(1.6)	37.8(1.6)	31.2(2.9)
13	24.9(0.9)	17.2(1.1)	38.4(0.5)	31.8(1.4)
14	27.2(2.5)	22.9(4.1)	37.7(0.9)	29.1(1.3)
15	27.8(1.7)	24.3(2.5)	36.2(1.6)	26.6(1.9)
16	27.2(1.3)	22.0(2.7)	38.6(1.5)	30.1(1.6)
17	30.1(0.6)	26.1(1.2)	38.7(1.0)	28.9(1.6)
18	29.2(0.7)	25.0(1.5)	40.3(0.5)	30.7(1.2)
19	27.7(0.9)	22.0(0.9)	39.3(1.4)	31.2(1.4)
20	28.6(0.9)	24.4(0.7)	41.2(0.4)	32.2(1.4)
21	30.8(0.7)	27.7(1.8)	41.9(1.8)	32.8(3.4)
22	26.8(0.7)	20.1(0.6)	42.3(0.9)	35.9(1.8)
23	31.0(2.6)	28.9(4.0)	45.8(1.0)	39.2(3.7)
24	26.3(1.0)	21.2(0.7)	20.3(1.0)	7.7(0.5)
25	26.7(0.9)	19.9(0.6)	48.4(0.7)	49.9(2.1)
26	31.6(2.1)	29.6(2.4)	42.4(1.6)	32.3(1.5)
27	31.2(0.8)	27.2(1.4)	47.7(0.5)	41.2(2.2)
28	29.5(0.7)	27.9(4.5)	43.4(1.1)	34.7(3.5)
29	29.3(1.4)	26.2(2.1)	42.7(0.9)	33.0(2.5)
30	31.3(2.3)	30.1(3.2)	43.6(1.1)	33.7(1.4)

* See Figure 33 for the locations of these sites.

** Mean and standard (in parentheses) deviations are for a 3- by 3-pixel array.

Table 17

Effect of Changing Screening Window Criterion on Number of Acceptable Sample Sites, 11 July 1974 and 21 February 1975 Data

Sample Site	Screening Window, CCT Units ±1.1		±1.3		±1.5		±1.7		±1.8	
	74	75	74	75	74	75	74	75	74	75
1				X	X	X	X	X	X	X
2		X	X	X	X	X	X	X	X	X
3	X		X		X		X		X	
4			X		X		X	X	X	X
5							X		X	
6				X		X		X	X	X
7						X		X		X
8		X				X		X		X
9	X		X		X		X	X	X	X
10							X		X	
11				X		X		X		X
12										
13					X		X	X	X	X
14										
15										
16										
17		X		X		X	X	X	X	X
18		X		X	X	X	X	X	X	X
19				X	X	X	X	X	X	X
20		X		X	X	X	X	X	X	X
21										
22		X		X		X		X	X	X
23								X		X
24	X		X		X		X		X	
25										
26										
27										
28				X		X		X		X
29								X		X
30										
Total Acceptable	3	6	5	10	10	12	13	17	15	17

Table 18

Effect of Variance Window on CCT Class Limits

11 July 1974 Data

Site No.*	Variance Window	Class Limits, CCT Units Spectral Band			
		Band 4	Band 5	Band 6	Band 7
1	1.3	24-26	18-20	37-40	30-33
	1.5	24-26	18-21	37-40	30-33
	1.7	23-27	17-21	37-40	30-33
	1.8	23-27	17-21	37-40	30-33
2	1.3	24-27	18-20	37-40	31-33
	1.5	24-27	18-20	37-40	30-33
	1.7	24-27	17-21	37-40	30-33
	1.8	24-27	17-21	37-40	30-34
3	1.3	24-26	18-20	38-40	32-35
	1.5	24-26	18-21	38-41	32-35
	1.7	23-27	17-21	38-41	32-35
	1.8	23-27	17-21	37-41	32-35
4	1.3	25-27	18-20	39-42	33-35
	1.5	25-27	18-21	39-42	32-35
	1.7	24-28	18-21	39-42	32-35
	1.8	24-28	17-21	39-42	32-36
5	1.3	31-33	29-32	40-42	28-31
	1.5	31-33	29-32	40-42	28-31
	1.7	30-34	29-32	39-43	28-31
	1.8	30-34	29-32	39-43	28-31
6	1.3	35-37	34-37	43-45	31-34
	1.5	35-37	34-37	43-45	31-34
	1.7	34-38	34-37	42-46	31-34
	1.8	34-38	34-37	42-46	31-34
9	1.3	23-26	15-18	38-40	32-35
	1.5	23-26	15-18	37-40	32-35
	1.7	23-26	15-18	37-40	32-35
	1.8	23-26	15-18	37-41	32-35
10	1.3	24-26	16-18	38-40	32-34
	1.5	24-26	16-19	37-40	32-34
	1.7	23-27	15-19	37-40	31-35
	1.8	23-27	15-19	37-41	31-35

(Continued)

* Sites selected using a screening window of ± 1.8 CCT units.

Table 18 (Concluded)

Site No.	Variance Window	Class Limits, CCT Units Spectral Band			
		Band 4	Band 5	Band 6	Band 7
13	1.3	24-26	16-18	37-40	31-33
	1.5	23-26	16-19	37-40	30-33
	1.7	23-27	16-19	37-40	30-33
	1.8	23-27	15-19	37-40	30-34
17	1.3	29-31	25-27	37-40	28-30
	1.5	29-32	25-28	37-40	27-30
	1.7	28-32	24-28	37-40	27-31
	1.8	28-32	24-28	37-40	27-31
18	1.3	38-30	24-26	39-42	29-32
	1.5	28-31	24-26	39-42	29-32
	1.7	28-31	23-27	39-42	29-32
	1.8	27-31	23-27	39-42	29-32
19	1.3	26-29	21-23	38-41	30-32
	1.5	26-29	21-23	38-41	30-33
	1.7	26-29	20-24	38-41	30-33
	1.8	26-29	20-24	38-41	29-33
20	1.3	27-30	23-26	40-42	31-33
	1.5	27-30	23-26	40-43	31-34
	1.7	27-30	23-26	40-43	31-34
	1.8	27-30	23-26	39-43	30-34
22	1.3	26-28	19-21	41-44	35-37
	1.5	25-28	19-22	41-44	34-37
	1.7	25-28	18-22	41-44	34-38
	1.8	25-29	18-22	41-44	34-38
24	1.3	25-28	20-22	19-22	6-9
	1.5	25-28	20-23	19-22	6-9
	1.6	25-28	20-23	19-22	6-9
	1.8	25-28	19-23	19-22	6-9

Table 19

CCT Classes, 11 July 1974 Data, with a Variance Window of ±1.3 CCT Units

Site No.	Class Limits, CCT Units Spectral Band			
	Band 4	Band 5	Band 6	Band 7
24	25-28	20-22	19-22	6-9
1	24-26	18-20	37-40	30-33
2	24-27	18-20	37-40	31-33
3	24-26	18-20	38-40	32-35
4	25-27	18-20	39-42	33-35
13	24-26	16-18	37-40	31-33
9	23-26	15-18	38-40	32-35
19	26-29	21-23	38-41	30-32
20	27-30	23-26	40-42	31-33
18	28-30	24-26	39-42	29-32

Table 20

Stretched Cluster Class Sets,* 11 July 1974 Data

Cluster	CCT Cluster Class Limits Spectral Band				Terrain Type
	Band 4	Band 5	Band 6	Band 7	
			Cluster Class Set 1		
A	0-63	0-63	0-63	0-19	Water
B	0-63	0-20	0-63	20-63	Forest
C	0-63	21-63	0-63	20-63	Nonforest
			Cluster Class Set 2		
A	0-63	0-63	0-29	0-63	Water
B	0-63	0-20	30-63	0-63	Forest
C	0-63	21-63	30-63	0-63	Nonforest

* As eventually used.

Table 21

Signatures from Sample Sites, 21 February 1975

Site No.*	Mean and Standard Deviations,** CCT Units			
	Band 4	Band 5	Band 6	Band 7
1	15.0(0.7)	16.1(0.3)	14.1(1.3)	4.4(0.5)
2	16.2(0.4)	20.3(0.5)	20.8(1.1)	10.6(1.1)
3	14.6(1.4)	16.0(1.5)	16.4(1.5)	8.6(2.5)
4	15.3(1.2)	16.9(1.6)	17.4(1.6)	7.9(1.4)
5	16.3(0.5)	17.1(1.1)	38.2(1.2)	28.1(2.5)
6	16.6(0.5)	19.9(0.6)	28.4(0.5)	16.0(1.2)
7	17.3(1.2)	21.7(1.4)	23.8(1.0)	11.0(0.9)
8	16.9(0.6)	19.6(0.5)	25.1(0.8)	12.7(0.9)
9	13.4(1.1)	15.3(0.9)	19.6(1.7)	10.9(1.5)
10	13.1(0.8)	14.6(1.1)	21.6(1.6)	10.6(2.2)
11	13.3(0.5)	15.8(1.2)	21.3(1.0)	11.7(0.9)
12	13.3(0.9)	14.3(1.0)	22.6(3.2)	12.0(2.0)
13	14.9(0.8)	17.7(1.3)	23.7(1.6)	13.0(0.7)
14	15.6(1.1)	17.0(1.9)	21.7(2.9)	14.9(4.0)
15	16.3(2.5)	18.1(3.0)	28.2(4.3)	21.2(1.3)
16	17.2(1.1)	21.4(1.7)	30.8(1.7)	21.0(2.4)
17	17.1(0.6)	19.1(0.6)	31.1(0.9)	19.3(0.7)
18	17.2(0.4)	20.0(0.)	28.2(0.8)	15.3(0.5)
19	17.4(0.7)	20.8(0.8)	25.8(1.2)	13.2(1.0)
20	17.7(0.7)	21.3(1.0)	28.9(0.8)	17.3(1.1)
21	22.7(1.0)	28.8(1.6)	32.4(3.4)	20.6(2.1)
22	17.0(0.7)	21.6(0.7)	26.6(0.7)	13.8(1.1)
23	19.5(1.0)	22.4(1.7)	33.8(1.4)	20.0(0.9)
24	20.3(1.8)	24.4(0.5)	16.9(1.2)	2.2(2.0)
25	17.4(0.7)	20.1(0.6)	32.7(3.4)	22.9(1.5)
26	21.0(0.9)	24.9(1.3)	32.4(2.5)	18.2(1.0)
27	22.2(0.7)	26.3(1.7)	29.9(2.2)	17.3(0.7)
28	18.2(0.7)	22.0(0.9)	28.8(0.8)	16.6(1.2)
29	22.5(1.2)	28.8(1.6)	31.0(1.3)	18.2(1.1)
30	21.4(0.7)	24.2(2.2)	32.1(2.6)	20.2(1.2)

* See Figure 33 for the locations of these sites.

** Mean and standard (in parentheses) deviations are for a 3- by 3-pixel array.

Table 22

Effect of Variance Window on CCT Class Limits

21 February 1975 Data

Site No.*	Variance Window	Class Limits,* CCT Units Spectral Band			
		Band 4	Band 5	Band 6	Band 7
1	1.3	14-16	15-17	13-15	3-6
	1.5	14-16	15-18	13-16	3-6
	1.8	13-17	14-18	12-16	3-6
2	1.3	15-17	19-22	20-22	9-12
	1.5	15-18	19-22	19-22	9-12
	1.8	14-18	19-22	19-23	9-12
4	1.3	14-17	16-18	16-19	7-9
	1.5	14-17	15-18	16-19	6-9
	1.8	14-17	15-19	16-19	6-10
6	1.3	15-18	19-21	27-30	15-17
	1.5	15-18	18-21	27-30	15-17
	1.8	15-18	18-22	27-30	14-18
7	1.3	16-19	20-23	23-25	10-12
	1.5	16-19	20-23	22-25	10-12
	1.8	16-19	20-23	22-26	9-13
8	1.3	16-18	18-21	24-26	11-14
	1.5	15-18	18-21	24-27	11-14
	1.8	15-19	18-21	23-27	11-14
9	1.3	12-15	14-17	18-21	10-12
	1.5	12-15	14-17	18-21	9-12
	1.8	12-15	14-17	18-21	9-13
11	1.3	12-15	15-17	20-23	10-13
	1.5	12-15	14-17	20-23	10-13
	1.8	12-15	14-18	20-23	10-13
13	1.3	14-16	16-19	22-25	12-14
	1.5	13-16	16-19	22-25	12-14
	1.8	13-17	16-19	22-25	11-15
17	1.3	16-18	18-20	30-32	18-21
	1.5	16-19	18-21	30-33	18-21
	1.8	15-19	17-21	29-33	18-21

(Continued)

* Sites selected using a screening window of ±1.8 CCT units.

Table 22 (Concluded)

Site No.	Variance Window	Class Limits, CCT Units Spectral Band			
		Band 4	Band 5	Band 6	Band 7
18	1.3	16-18	19-21	27-29	14-17
	1.5	16-19	19-21	27-30	14-17
	1.8	15-19	18-22	26-30	14-17
19	1.3	16-19	20-22	25-27	12-14
	1.5	16-19	19-22	24-27	12-15
	1.8	16-19	19-23	24-28	11-15
20	1.3	16-19	20-23	28-30	16-19
	1.5	16-19	20-23	27-30	16-19
	1.8	16-19	20-23	27-31	16-19
22	1.3	16-18	20-23	25-28	13-15
	1.5	16-18	20-23	25-28	12-15
	1.8	15-19	20-23	25-28	12-16
23	1.3	18-21	21-24	33-35	19-21
	1.5	18-21	21-24	32-35	19-21
	1.8	18-21	21-24	32-36	18-22
28	1.3	17-19	21-23	28-30	15-18
	1.5	17-20	21-23	27-30	15-18
	1.8	16-20	20-24	27-31	15-18
29	1.3	21-24	28-30	30-32	17-19
	1.5	21-24	27-30	30-32	17-20
	1.8	21-24	27-31	29-33	16-20

Table 23

CCT Classes, 21 February 1975 Data, with a Variance Window of ±1.5 CCT Units

Site No.	Class Limits, CCT Units Spectral Band			
	Band 4	Band 5	Band 6	Band 7
24	19-22	23-26	15-18	1-4
1	14-16	15-18	13-16	3-6
11	12-15	14-17	20-23	10-13
2	15-18	19-22	19-22	9-12
7	16-19	20-23	22-25	10-12
8	15-18	18-21	24-27	11-14
19	16-19	19-22	24-27	12-15
22	16-18	20-23	25-28	12-15
18	16-19	19-21	27-30	14-17
6	15-18	18-21	27-30	15-17
28	17-20	21-23	27-30	15-18
20	16-19	20-23	27-30	16-19
17	16-19	18-21	30-33	18-21

Table 24

Stretched Cluster Class Sets,* 21 February 1975 Data

Cluster	Cluster Class Limits, CCT Units Spectral Band			
	Band 4	Band 5	Band 6	Band 7
	Cluster Class Set 1			
A	0-63	21-63	0-18	0-63
B_1	0-63	0-20	0-18	0-63
B_2	0-63	0-17	19-63	0-63
C	0-63	18-63	19-63	0-63
	Cluster Class Set 2			
A	0-63	21-63	0-63	0-7
B_1	0-63	0-20	0-63	0-7
B_2	0-63	0-17	0-63	8-63
C	0-63	18-63	0-63	8-63

* With two types of forests, B_1 and B_2.

Table 25

Cluster Class Sets Selected to Discriminate Between Flooded and Unflooded Forest

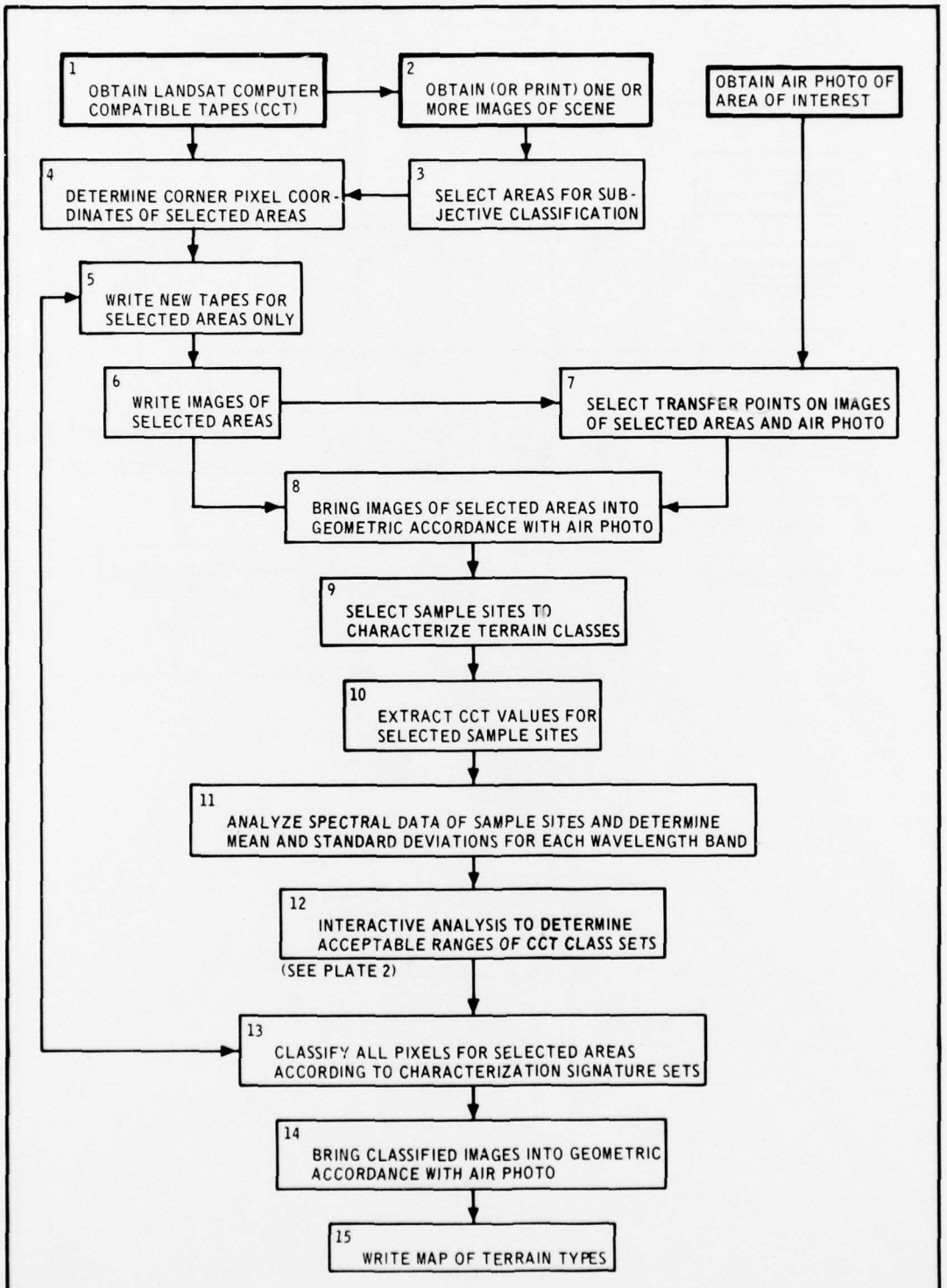
21 February 1975 Data

Cluster	Cluster Class Limits, CCT Units Spectral Band				Terrain Type
	Band 4	Band 5	Band 6	Band 7	
A	0-63	21-63	0-63	0-7	Open water
B_1	0-63	0-20	0-63	0-7	Flooded forest
B_2	0-63	0-17	0-63	8-63	Dry forest
C	0-63	18-63	0-63	8-63	Nonforest

Table 26

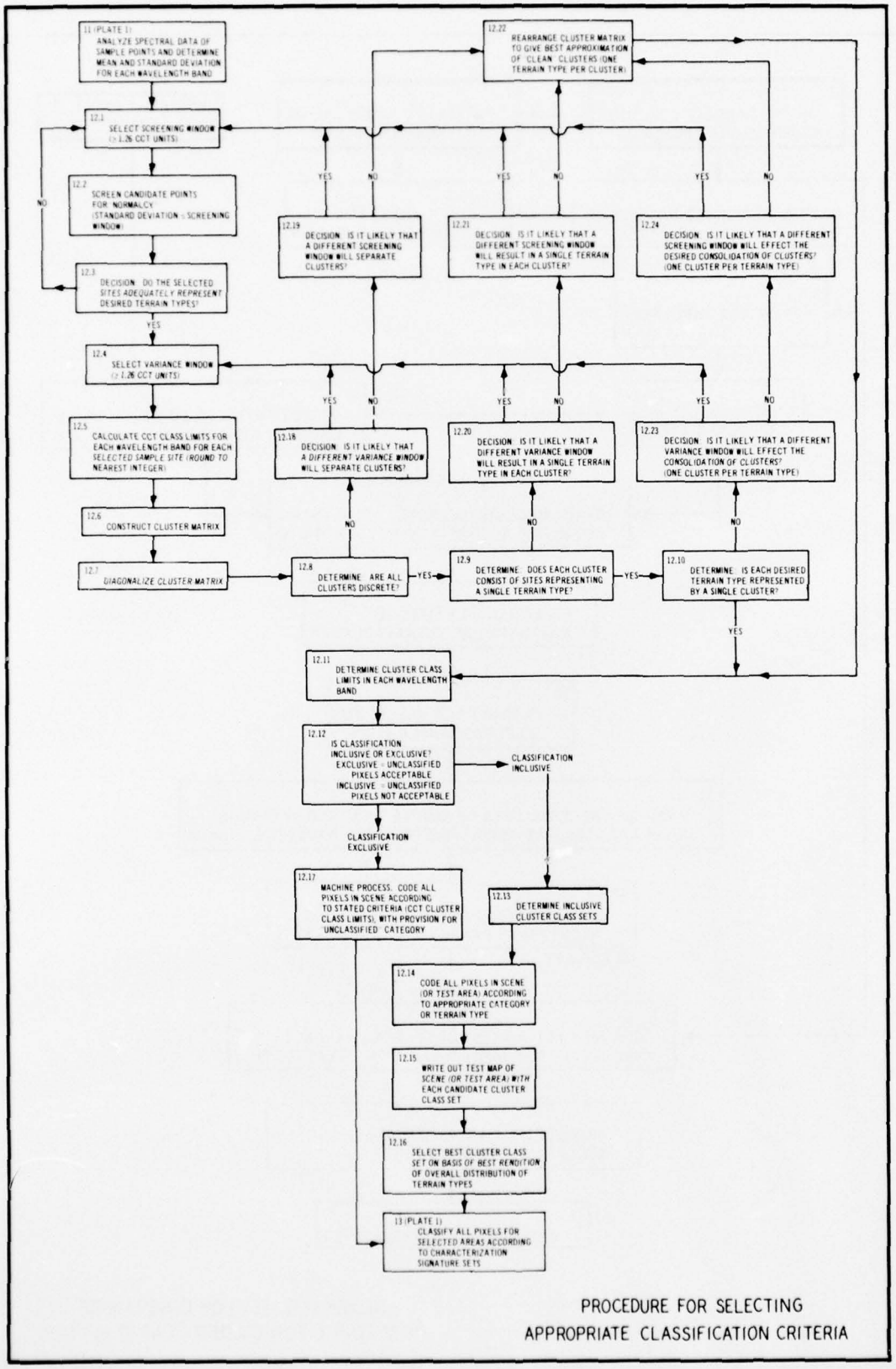
Revised Stretched Cluster Class Set 1, 11 July 1974 Data

Cluster	Cluster Class Limits, CCT Units Spectral Band				Terrain Type
	Band 4	Band 5	Band 6	Band 7	
A_1	0-63	0-63	0-63	0-8	Deep water
A_2	0-63	0-63	0-63	9-19	Shallow water
B	0-63	0-20	0-63	20-63	Forest
C	0-63	21-63	0-63	20-63	Nonforest



GENERALIZED FLOW DIAGRAM OF PROCEDURE FOR GUIDED CLASSIFICATION

PLATE 1



PROCEDURE FOR SELECTING APPROPRIATE CLASSIFICATION CRITERIA

PLATE 2

APPENDIX A: DESCRIPTION OF OPTRONICS FILM READER/WRITER

1. Maps depicting the interpreted terrain conditions are produced using an incremental film reader/writer (Figure A1). The film reader/writer is an electromechanical photograph-scanning and film-writing

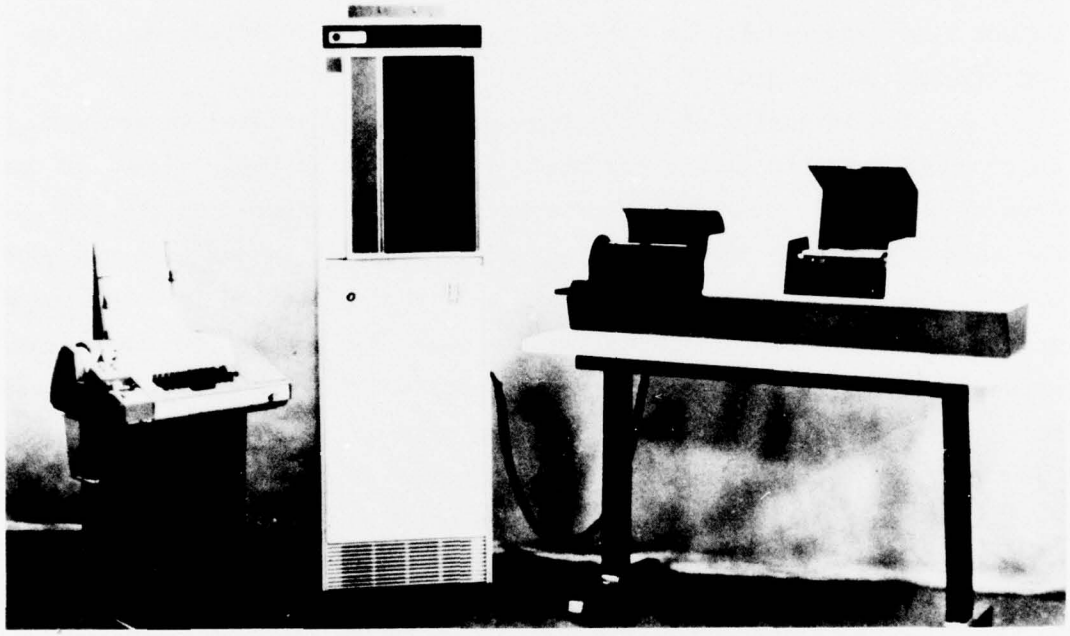


Figure A1. Film reader/writer

system designed to produce images up to 22.3 by 25.4 cm in size. The instrument can operate in either of two modes--an input or scanning mode, or an output or writing mode. The writing mode is used to convert CCT values to images in photographic form. For this mode of operation, the instrument is equipped with a rotating drum and an optical system consisting of a light-emitting diode (LED), a selectable aperture, and a lens system that focuses a spot of light from the LED onto the perimeter of the drum. The drum is housed in a light-tight enclosure, which is demountable and is removed to a photographic darkroom for film loading and unloading. A piece of film is clamped to the outside of the drum for exposure.

2. The film can be exposed using 12.5-, 25.0-, or 50.0-μm spots

of light from the LED. The raster interval and spot size, which are selectable, control the size of the negatives that will be produced. As the drum rotates, the carriage supporting the optical system is stepped one step per drum revolution in the axial direction at the selected raster interval until the total area of the film or the area of interest has been exposed. The use of high-speed film permits very short exposure times and results in a recording rate of up to 60,000 exposures (or pixels) per second.

3. The intensity of light from the LED is modulated incrementally in proportion to the values recorded on the magnetic tape. Thus, as the drum of the film reader/writer rotates, a spot is exposed on the film for each pixel value in a scan line (row). When an interrecord gap code occurs, the LED is extinguished until the drum revolution is completed and the carriage for the optical system has moved forward one increment. Exposure of the next row of spots then begins. This process is repeated until the end-of-file code occurs on the magnetic tape.

4. The film reader/writer is controlled by a minicomputer so that real-time manipulation of the CCT values can be used to produce a number of photographic effects. However, for this capability to be used successfully, very careful control over the photographic process and knowledge of the relation between the digital input from the magnetic tape and the photographic output of the film writer are required.

In accordance with letter from DAEN-RDC, DAEN-ASI dated 22 July 1977, Subject: Facsimile Catalog Cards for Laboratory Technical Publications, a facsimile catalog card in Library of Congress MARC format is reproduced below.

Struve, Horton

Acquisition of terrain information using Landsat multispectral data; Report 2: An interactive procedure for classifying terrain types by spectral characteristics / by Horton Struve, Warren E. Grabau, and Harold W. West. Vicksburg, Miss. : U. S. Waterways Experiment Station, 1977.

105, [28], 2 p., 2 leaves of plates : ill. ; 27 cm. (Technical report - U. S. Army Engineer Waterways Experiment Station ; M-77-2, Report 2)

Prepared for Assistant Secretary of the Army (R&D), Department of the Army, Washington, D. C., under Project 4A161101A91D, Task 02, Work Unit 095.

References: p. 105.

1. Data acquisition. 2. Landsat (Satellite). 3. Multispectral data. 4. Spectrum analysis. 5. Terrain classification. I. Grabau, Warren E., joint author. II. West, Harold W., joint author. III. United States. Assistant Secretary of the Army (Research and Development). II. Series: United States. Waterways Experiment Station, Vicksburg, Miss. Technical report ; M-77-2, Report 2.

TA7.W34 no.M-77-2 Report 2