JPL PUBLICATION 82-4

Image Processing Developments and Applications for Water Quality Monitoring and Trophic State Determination

Richard J. Blackwell

March 1, 1982

NASA
National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No. 82-4	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Image Processing Developments and Applications for Water Quality Monitoring and Trophic State Determination		5. Report Date 3-1-82 6. Performing Organization Code
7. Author(s) R. Blackwell		8. Performing Organization Report No.
9. Performing Organization Name and Address JET PROPULSION LABORATORY California Institute of Technology 4800 Oak Grove Drive Pasadena, California 91109		10. Work Unit No. 11. Contract or Grant No. NAS 7-100 13. Type of Report and Period Covered JPL Publication
12. Sponsoring Agency Name and Address NATIONAL AERONAUTICS AND SPACE ADMINISTRATION Washington, D.C. 20546		14. Sponsoring Agency Code 146-40-15-01-00

15. Supplementary Notes

16. Abstract

This report concludes a multiphase program to demonstrate the utility of remote sensing data analysis to water quality monitoring. This phase demonstrates that data analysis and image processing techniques can be applied to Landsat remote sensing data to produce an effective operational tool for lake water quality surveying and monitoring. The major findings of this report are: (1) Digital image processing and analysis techniques have been designed, developed, tested and applied to Landsat Multispectral Scanner (MSS) data and conventional surface acquired data. Utilization of these techniques can facilitate the surveying and monitoring of large numbers of lakes in an operational manner. (2) Supervised multispectral classification, when used in conjunction with surface acquired water quality indicators, can be used to characterize water body trophic status. (3) Unsupervised multispectral classification, when interpreted by lake scientists familiar with a specific water body, can yield classifications of equal validity with supervised methods and in a more cost-effective manner. (4) Image data base technology can be used to great advantage in characterizing other contributing effects to water quality. These effects include drainage basin configuration, terrain slope, soil, precipitation and land cover characteristics.

17. Key Words (Selected by Author(s))	18. Distribution Statement
Environment Pollution Computer Programming and Software	Unclassified - Unlimited

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	116	

JPL 0184 11/80

JPL PUBLICATION 82-4

Image Processing Developments and Applications for Water Quality Monitoring and Trophic State Determination

Richard J. Blackwell

March 1, 1982

NASA
National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

CONTENTS

Figures

CONTENTS (Contd.)

CONTENTS (Contd.)

CONTENTS (Contd.)

Tables

CONTENTS (Contd.)

ACKNOWLEDGEMENTS

The contributions to this document of A. Y. Smith, J. D. Addington, K. J. Hussey and D. H. P. Boland are gratefully acknowledged.

ABSTRACT

This report concludes a multiphase program to demonstrate the utility of remote sensing data analysis to water quality monitoring. This phase demonstrates that data analysis and image processing techniques can be applied to Landsat remote sensing data to produce an effective operational tool for lake water quality surveying and monitoring. The major findings of this report are:

(1) Digital image processing and analysis techniques have been designed, developed, tested and applied to Landsat Multispectral Scanner (MSS) data and conventional surface acquired data. Utilization of these techniques can facilitate the surveying and monitoring of large numbers of lakes in an operational manner.

(2) Supervised multispectral classification, when used in conjunction with surface acquired water quality indicators, can be used to characterize water body trophic status.

(3) Unsupervised multispectral classification, when interpreted by lake scientists familiar with a specific water body, can yield classifications of equal validity with supervised methods and in a more cost-effective manner.

(4) Image data base technology can be used to great advantage in characterizing other contributing effects to water quality. These effects include drainage basin configuration, terrain slope, soil, precipitation and land cover characteristics.

INTRODUCTION

This publication presents the results of the final work completed through the joint efforts of JPL's Image Processing Laboratory (IPL) and U.S. Environmental Projection Agency, Las Vegas (EPALV) on the Lake Classification Task. The report has been divided into two major sections.

Section I, entitled Significant Technology, presents brief descriptions of four digital image processing software programs on systems (LAKELOC, STORET, STATS2, and IBIS). LAKELOC, STORET and STATS2 were developed specifically for use by the lake classification projects. The IBIS system was developed originally for land use planning activities but subsequently successfully applied to the lake classification effort. These four systems formed the technical core through which all Landsat data and much control sensed data were analyzed.

Section II, entitled Technology Applications, presents detailed discussions of three major lake classification studies which were conducted between 1977 and 1979. The studies are presented separately in chronological order beginning with the Lake Mead Intensive Area Study. This study was designed as a demonstration of the application of Landsat data to water quality monitoring of a single, large water body. Multispectral classification was performed and the results compared with conventionally acquired contact-sensed measurements gathered on corresponding Landsat flyover dates. The Lake Tahoe Study was conducted as a demonstration of the application of IBIS to a water quality monitoring situation. A data base was constructed consisting of data elements from Landsat, digital terrain data and Tahoe Regional Planning Agency statistics. Data base manipulation and some preliminary modelling was performed for demonstration purposes. The third study, involving lakes in Montana, Utah and Michigan, was the final phase of the Lake Classification Project designed as a demonstration of computer program capabilities. Multispectral classification was performed on some lakes selected for this study.

SECTION I

SIGNIFICANT TECHNOLOGY

Three software systems were designed specifically to handle the task of lake classification. LAKELOC and its associated follow-on programs FARINA and STATUS were developed to aid the analyst with identification and isolation of lakes selected for study. Either Landsat or digital aircraft imagery could be used as input to LAKELOC. The two follow-on programs could be used to extract information for each lake of interest from each spectral band and present this information in the form of a computer listing.

The STORET system was designed to present contact-sensed water quality measurements in the form of scatter diagrams. The information source for this program was data from the National Eutrophication Survey.

The computer program STATS2 was developed to be used interactively to locate and identify sample site locations on lakes selected for study. Multispectral information pertaining to each sample site could be printed out. In addition, supervised classification could also be performed using STATS2 by using sample site statistics as input to an automatic classifier.

The IBIS system was developed at JPL originally for land use studies requiring the development of information data bases to be constructed from varied sources. IBIS was used in the lake classification project to build a data base composed of information from sources such as Landsat, Defense Mapping Agency digital terrain data and maps provided by the Tahoe Regional Planning Agency.

A. LAKELOC

As described in previous project reports (Refs. I-1, 2, 3, 4), lake isolation performed at the IPL relied solely upon batch processing techniques. Analysts used only the IR2 band from Landsat imagery to define water by examining histograms and setting an average threshold (Ref. I-5). This method, while fairly accurate for defining water boundaries and aiding the lake extraction process, was inherently slow. The analyst was required to wait for processing of hardcopy photo products before continuing with each phase of lake extraction. In addition, small lakes were extremely difficult to separate, especially if these lakes were directly associated with river systems or closely situated to other water bodies.

To streamline and increase the accuracy of lake extraction, analysts associated with this project developed an interactive system of lake isolation. The system consisted of three interactive programs which reduced the time needed to locate and isolate a lake, increased the accuracy and sensitivity of the water detection scheme and output a statistical and surface area listing for each lake to be surveyed. The system was called LAKELOC. The following section describes the hardware configuration, program operation, water detection algorithm and hardcopy output associated with the LAKELOC system.

1. Hardware

The host computer is currently an IBM 370/158. The display controller used is a Ramtek G100B, a versatile video display device that can be used to display gray level images and graphics data. The Ramtek is a solid state refresh memory system with a display format of 512 lines by 640 elements. Readback from the refresh memory is available under software control. It is possible to display 6-bit gray level images along with two graphics planes, and the user may selectively write or erase the displays point by point. Manipulation of the graphics data can be accomplished with the aid of a trackball cursor. Figure I-1 illustrates the configuration of the interactive hardware as it is arranged for the operation of the LAKELOC program.

2. Operation of LAKELOC

For the purposes of illustration, we will follow the operation of LAKELOC as it would be applied to a scene in southern Illinois. Although any number of lakes could be extracted from a scene, we will limit the operation to one lake, in this case, Crab Orchard Lake.

For a given digital data scene, such as Landsat, the user may selectively display 512 x 640 element subareas until he locates the water body of interest. Automatic linear contrast stretching of the displayed scene can be performed during this operation to aid in the location of the lake.

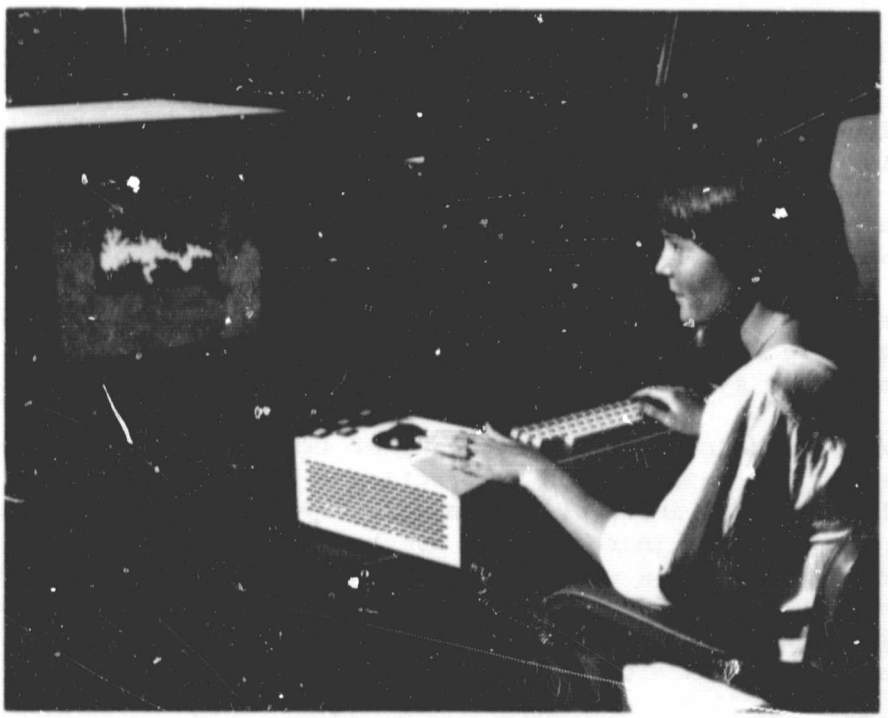

Figure I-1. User console with video display device and track ball

Once a lake has been located, the trackball cursor is then set on the desired lake and a default 50 x 50 element box is drawn on the graphics plane about the cursor position. Figure I-2 illustrates the default box drawn about the cursor position on Crab Orchard Lake. Since only the area within the box will be acted on by the water detector, the user must correct the size and position of the box relative to the lake so that the lake is contained within the box boundaries. The size is changed by a simple command to the program, which allows the manipulation of the trackball cursor to control the box dimensions. The position of the box is also controlled in the same manner by the trackball. An automatic mode can also be used in which the user places the cursor on the lake of interest and gives the command "window". A box is automatically drawn about the lake boundaries with no need for trackball manipulation. If a lake encompasses a major portion of the video display, a command can also be given to the program to "window" the entire screen in an automatic mode. This allows the user to forego manipulation of the cursor if the entire screen dimension is desired. Figure I-3 illustrates Crab Orchard Lake completely enclosed by the box after manipulation by the cursor.

Once the box has been satisfactorily positioned about the lake boundaries the water detector can be used to isolate the water body in a binary form. (A detailed description of the water detector follows in the section entitled "Water Detection Algorithm".) In the binary form, the water bodies appear as white and non-water features as black. The user can also magnify the area within the box boundaries by issuing a "zoom" command with the appropriate magnification factor.

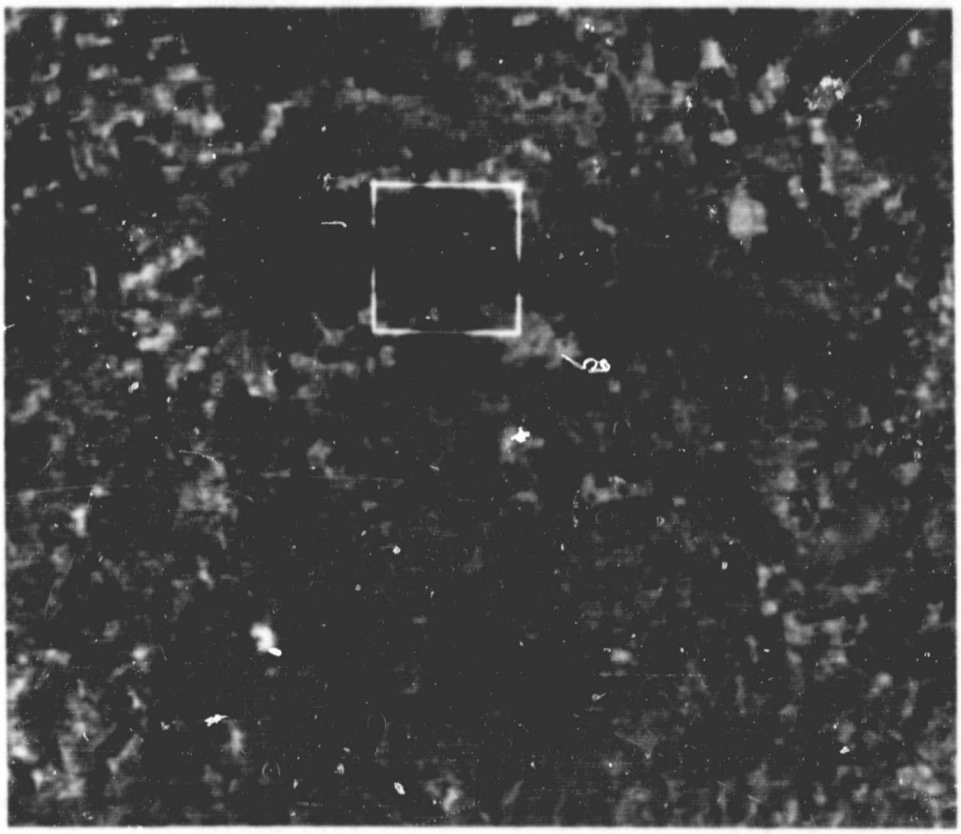

Figure I-2. Crab Orchard Lake with default 50 x 50 element box

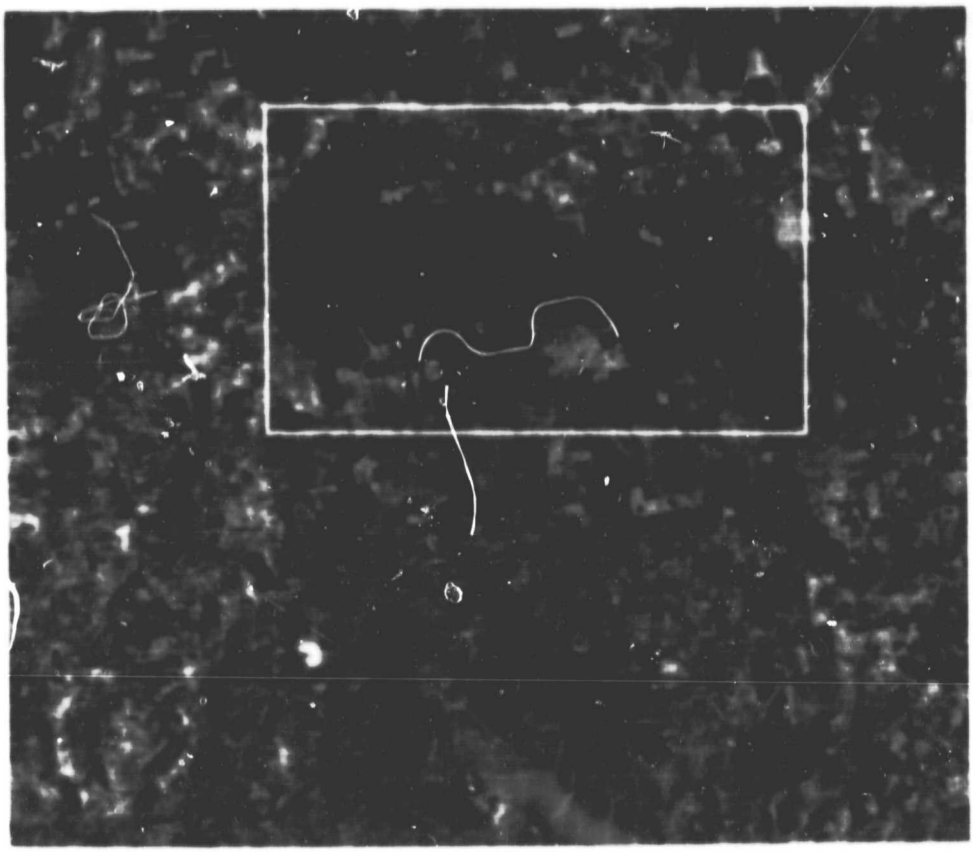

Figure I-3. Crab Orchard Lake with correctly sized and positioned box

The zoom command thus redisplays a magnified picture of the boxed area only, directly over the existing image. At the conclusion of the editing session the magnified image is erased from the screen, leaving the original image. This allows the user to continue uninterrupted with other lakes contained in the existing scene.

Magnification of the image allows the user to easily determine the exact boundaries of the lake as opposed to any extraneous water information which may also be displayed in the scene. The task of editing out extraneous pixels has been in the past the most time-consuming chore in the water analysis project. Aided only infrequently by a map, the user had to decide what constituted the boundaries of the lake in question. Previously the user relied on pixel listings and hardcopy photographs to locate the lake boundaries. In the case of a large lake, one was often hampered by cumbersome pixel listings which had to be carefully pieced together to recreate a lake image. With LAKELOC, the magnification factor in conjunction with the easily manipulated trackball allowed the user to perform the editing task in a matter of minutes as opposed to a duration of several days. Figure I-4 illustrates Crab Orchard Lake in binary format. The detached white areas represent extraneous water information which is not associated with the lake.

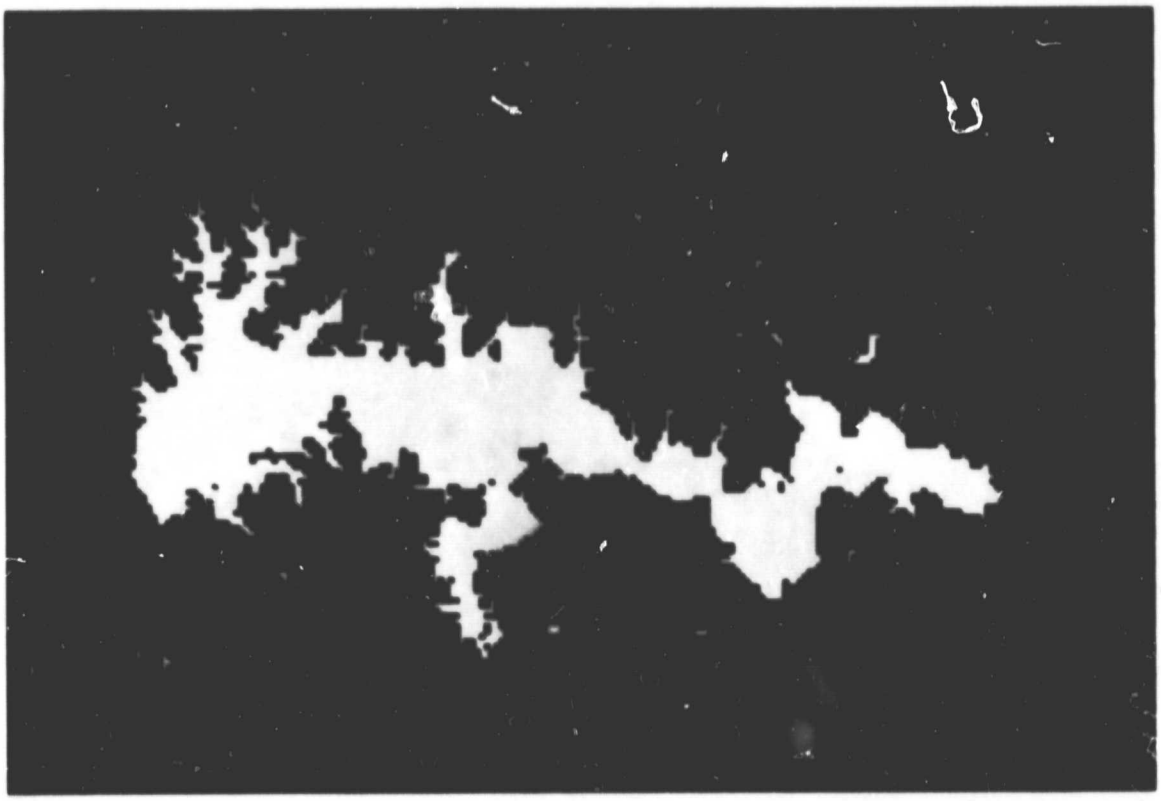

Figure I-4. Crab Orchard Lake in binary form with magnification factor of 2, before editing of extraneous water features

The removal of water bodies not associated with the lake of interest can be performed in three ways. In the first method the trackball-controlled cursor is set point-by-point on the areas to be removed. The default size of the area removed is one pixel; however, the user can specify the number of surrounding pixels to be removed for each erase operation. This method is most useful when working in close proximity to the boundaries of the lake of interest, when it is most imperative not to remove too large a section of pixels. The second method is used for removing closely spaced extraneous pixels which are not in close proximity to the lake. This method utilizes continuous erasures as the trackball cursor is moved across the screen. The size of the area about the cursor position to be erased can also be controlled by the user in this mode. The third method allows the user to place the cursor on the lake of interest and isolate it automatically. The user must first, if necessary, "detach" the lake from any other water bodies by one of the two previous methods. A command is then entered which quickly erases all water information contained within the boxed area except for the lake on which the cursor is situated. Figure I-5 depicts Crab Orchard Lake still in binary format after all extraneous information has been removed during the editing phase.

Once the user is satisfied that he has isolated the lake of interest, he then assigns the lake a name and commands the program to save the binary image of the

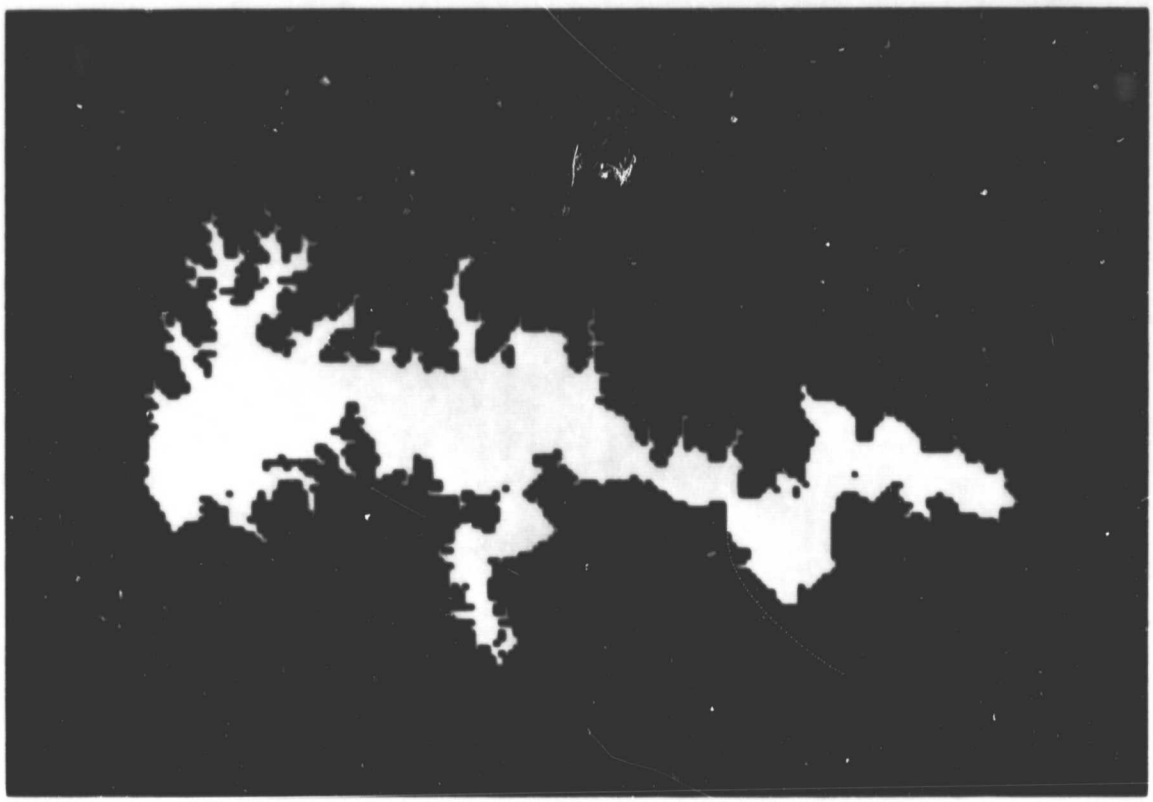

Figure I-5. Crab Orchard Lake in final edited form

lake on a disk data set. The lake's position in the disk data set is exactly the same as it is in the original Landsat scene. LAKELOC returns to the user the exact position and size of the extracted lake image as it appears on the disk data set, and a parameter data set is also created which contains this positional information. This information is necessary to the operation of the follow-on programs for LAKELOC. At this time the user is able to continue processing any number of lakes or, if finished, he can fetch the follow-on programs which will process the statistical data.

3. Follow-on Programs

The output from LAKELOC consists of a binary mask disk data set containing the extracted lakes and a parameter data set containing the positional information and lake names. The output size of the binary disk data set is exactly the same as the size of the original Landsat image used as input to Lakeloc. In the next step, the binary data set is used by the program FARINA to make out of each corresponding spectral channel in the original Landsat frame each water feature which has been processed through LAKELOC. The output is four data sets containing the original DN values for each lake in each of the spectral bands. This output can in turn be used as input to the program STATUS and as input to follow-on MSS classification programs.

STATUS, utilizing the parameter data set from LAKELOC or punched parameter cards, produces a statistical analysis in the form of a hardcopy printer listing of lake statistics in all four spectral channels. The lakes are listed by name and ranked according to size. Two tables are printed, the first consisting of lake statistics such as pixel count, surface area calculations and shoreline perimeter calculations. The second table lists lake MSS statistics, such as the mean DN level for each lake in each spectral channel and the corresponding standard deviations. Figure I-6 is a reproduction of the lake statistics table, including the statistics for Crab Orchard lake among other lakes processed from the same Landsat frame. Figure I-7 reproduces the lake MSS statistics for the same lakes listed in Figure I-6.

*** LAKE STATISTICS ***

** LAKE NAME **	** TOTAL PIXELS **	** SURFACE AREA **			** SHORELINE **	
		SQUARE FEET	ACRES	HECTARES	FEET	METERS
PINCKNEYVILLE 99	59	4064491.0	93.3	37.8	11650.1	3550.8
DUQUOIN 98	74	5097786.0	117.0	47.4	17379.6	5297.0
WASHNGTN CO 123	103	7095567.0	162.9	65.9	31260.7	9527.8
DEVILS KITCH 128	338	23284612.0	534.5	216.3	84010.2	25605.1
LITL GRASSY 119	547	37682283.0	865.1	350.1	101302.8	30875.6
CEDAR 39	895	58900095.0	1352.2	547.2	161456.4	49209.5
EGYPT 132	1121	77224569.0	1772.8	717.4	240158.9	73196.9
KINKAID 40	1267	87282363.0	2003.7	810.9	239724.2	73064.4
CRAB ORCHARD 127	4027	277416003.0	6368.6	2577.3	363440.0	110771.1
RENO 29	10694	736698966.0	16912.2	6844.1	583720.1	177909.3

Figure I-6. Lake statistics table listing dimension calculations

*** LAKE MSS STATISTICS ***

** LAKE NAME **	** MEAN **				*	** STANDARD DEVIATION **			
	GREEN	RED	IR1	IR2	*	GREEN	RED	IR1	IR2
PINCKNEYVILLE 99	45.03	29.71	23.71	11.36	*	2.75	2.25	5.86	6.65
DUQUOIN 98	35.04	18.28	19.22	9.88	*	2.33	2.03	5.92	6.44
WASHNGTN CO 123	40.63	27.75	25.39	13.42	*	5.93	8.53	8.98	7.81
DEVILS KITCH 128	33.12	16.12	17.38	9.89	*	1.78	1.86	6.80	7.89
LITL GRASSY 119	36.83	18.60	16.98	9.08	*	2.94	2.52	6.74	7.86
CEDAR 39	34.68	19.15	17.80	9.16	*	2.91	3.00	6.38	7.54
EGYPT 132	38.27	22.81	20.06	10.92	*	3.68	3.39	7.83	9.10
KINKAID 40	35.89	18.94	17.98	9.68	*	3.00	3.65	7.30	7.77
CRAB ORCHARD 127	38.84	23.69	20.82	8.13	*	3.11	3.47	6.37	6.40
RENO 29	40.50	26.73	19.87	7.03	*	5.46	6.51	7.49	6.11

Figure I-7. Lake statistics table listing MSS means and standard deviations

4. Water Detection Algorithm

In the past, the detection of pixels whose instantaneous field of view (IFOV) is that of water was accomplished by thresholding Landsat's infrared band 7 (0.8 - 1.1). The low reflectance of water in this spectral range conveniently produced a bimodal distribution of DN's: one peak for water, another peak for non-water. This technique was accurate except in the case where the IFOV of the scanner was viewing a combination of water and non-water areas such as at the shoreline of a lake. In this situation, the problem became one of trying to estimate the proportion of each material in the IFOV.

Horowitz (Ref. I-6) Work and Gilmer et al. (Ref. I-7) have investigated the proportion estimation problem and have encouraging results. Work and Gilmer have estimated the proportions of water, bare soils and green vegetation using Landsat bands 5 and 7. This technique requires an estimate of the spectral signature for pure water, pure bare soil, and pure vegetation. While the spectral signature of water is fairly easy to estimate, that for soil and vegetation becomes more difficult. The many variables involved, such as different soil types, vegetation cover types and thickness of the vegetative cover cause considerable error when estimation is attempted by a completely automatic processor.

An alternate approach, and the one chosen to implement in the LAKELOC system, considers the mixture classes to be only water and non-water. Bands 5 and 7 are used in the detection process, as it was found that bands 4 and 6 offer little in additional information. The estimation of the spectral signature for water and non-water is made over a region within and immediately surrounding the water body.

The spectral signatures (mean DN values) were estimated by an iterative procedure. First the 2-dimensional space (band 5 vs. band 7) is partitioned into two regions in which the populations of water and non-water typically cluster and the mean is then recomputed for those DN's which fall within a neighborhood of the initial mean. This process is continued until a convergent mean has been found for each region.

The proportion estimation implemented uses a technique proposed by McCloy (Ref. I-8). In Figure I-8, W is the mean for water, U is the mean for non-water and P is the DN for any given pixel. P' is the projection of P onto the line segment WU. If /WU/ is the length of the line segment WU and /WP'/ is the length on line segment WP", then the proportion estimate q for water is:

$$q = 1 - \frac{/WP'/}{/WU/}$$

where $0 \leq q \leq 1$. If P' does not fall between W and U, it is given the position of the closest point W or U. A decision threshold is set for q at which the pixel is defined to be water or non-water.

B. STORET

The STORET work was begun by JPL in late 1978 and left uncompleted due to termination of the project. The work was attempting to develop experimental

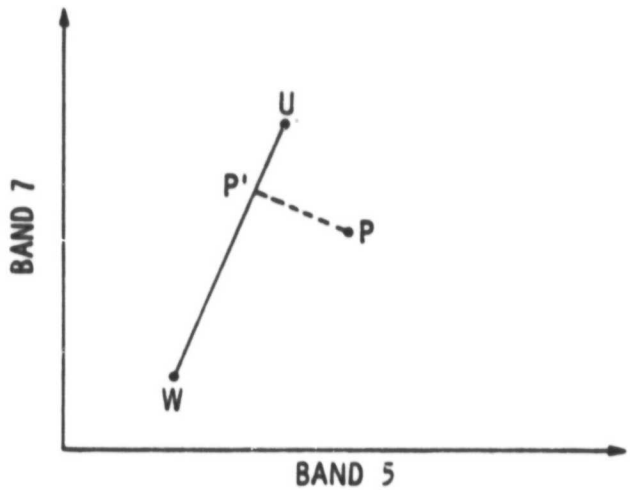

Figure I-8. Geometric interpretation of water detection algorithm

models that could relate remotely sensed multispectral data to actual water quality measurements.

EPA produced for JPL a set of STORET data that contained the chemical composition of water samples along with Secchi depths and turbidity measures for many of the fresh water lakes in the U.S. for the years 1973, 1974 and 1975. Most sample sites had samples taken during summer, fall and spring months. Computer software was developed that would allow the analyst to extract from the STORET files specific water quality measurements for a given year and season. Due to the high volume of data, an analyst would typically pick some subset of the 21 possible water quality measures for analysis.

In order to find relationships between remotely sensed data and water quality measurements, some basic software tools had to be developed. A data format structure was defined that would allow both water quality measures and remotely sensed data to be stored concurrently in one data file. Each measure was considered as a separate variate. The name of each variate was stored in a header record for the file. This scheme allows much flexibility in both the data format and the software that manipulates the data.

1. Basic Software

Some basic software developed for analysis of this multivariate data is now described.

(1) Plots. A program that could plot, on CRT graphics plane, scattergrams of selected pairs of variates. For example, chlorophyl-A might be plotted against total organic nitrogen; Secchi depth vs Landsat MSS-5, etc. These plots could be made in linear or log domain, with manual or automatic scaling of axes. A linear regression line could also be plotted along with the correlation coefficient of the plotted points. Provisions existed for easily producing hard copy prints of scattergrams after they were plotted on the CRT (see Figures I-9, 10, 11, 12).

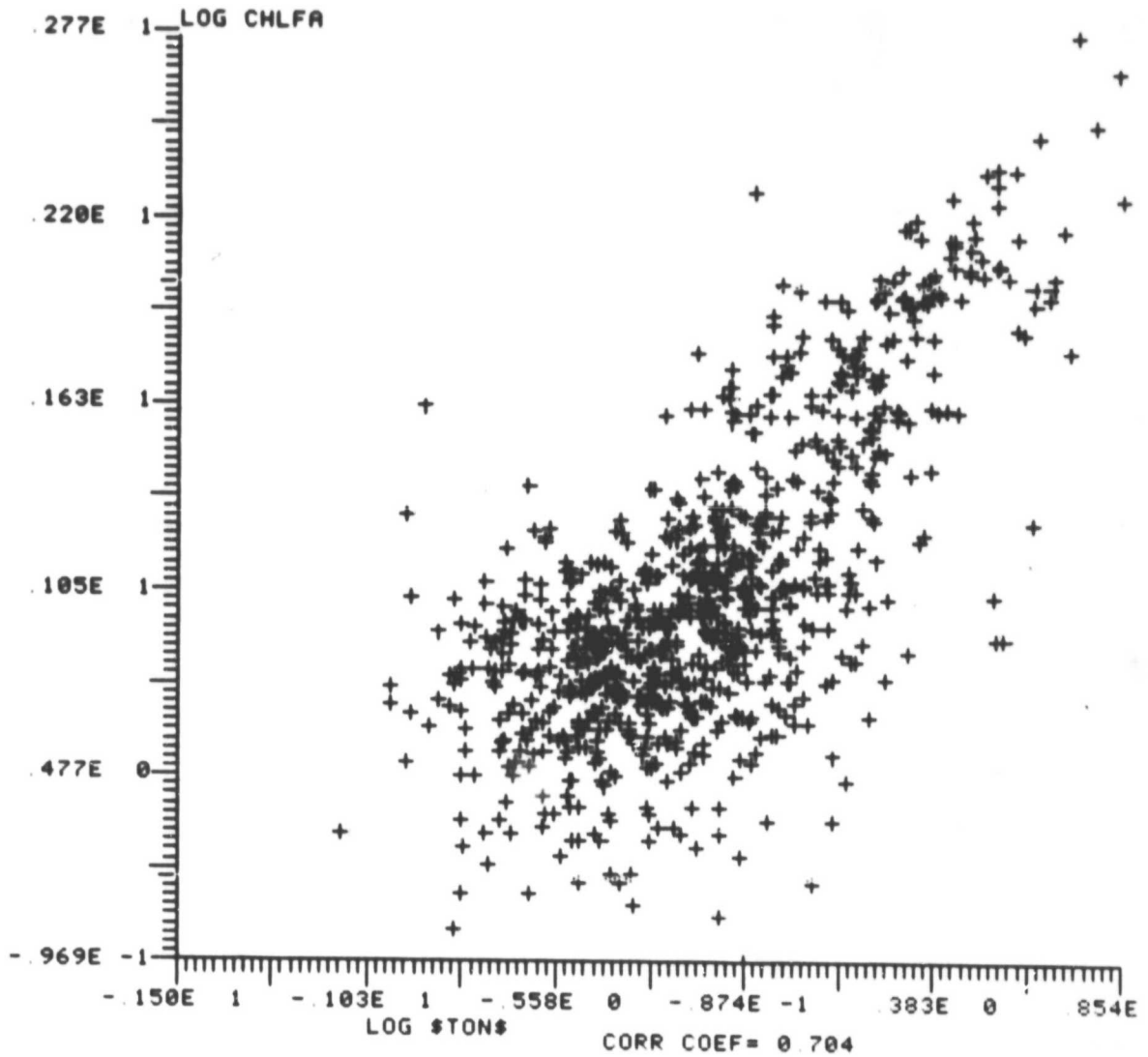

Figure I-9. A scattergram of the log of total organic nitrogen (TON) on the X-axis versus the log of chlorophyll a (CHLFA) on the Y-axis. Data is from STORET for the summer of 1973. The correlation coefficient is also given

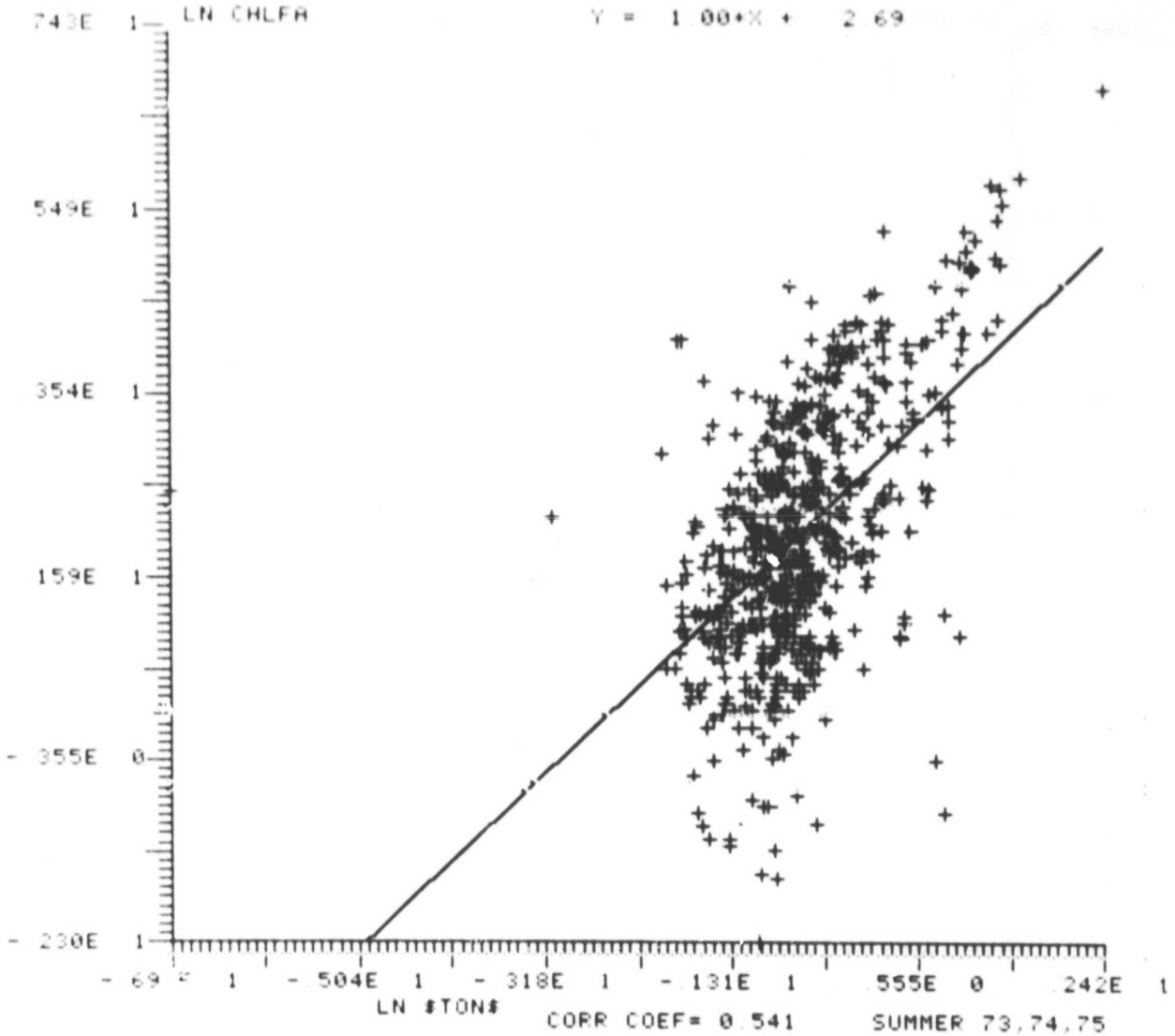

Figure I-10. A scattergram of the log of total organic nitrogen (TON) versus chlorophyll a (CHLFA). Included with the plot are the linear regression line and the slope. Data is for the summers of 1973, 1974, 1975

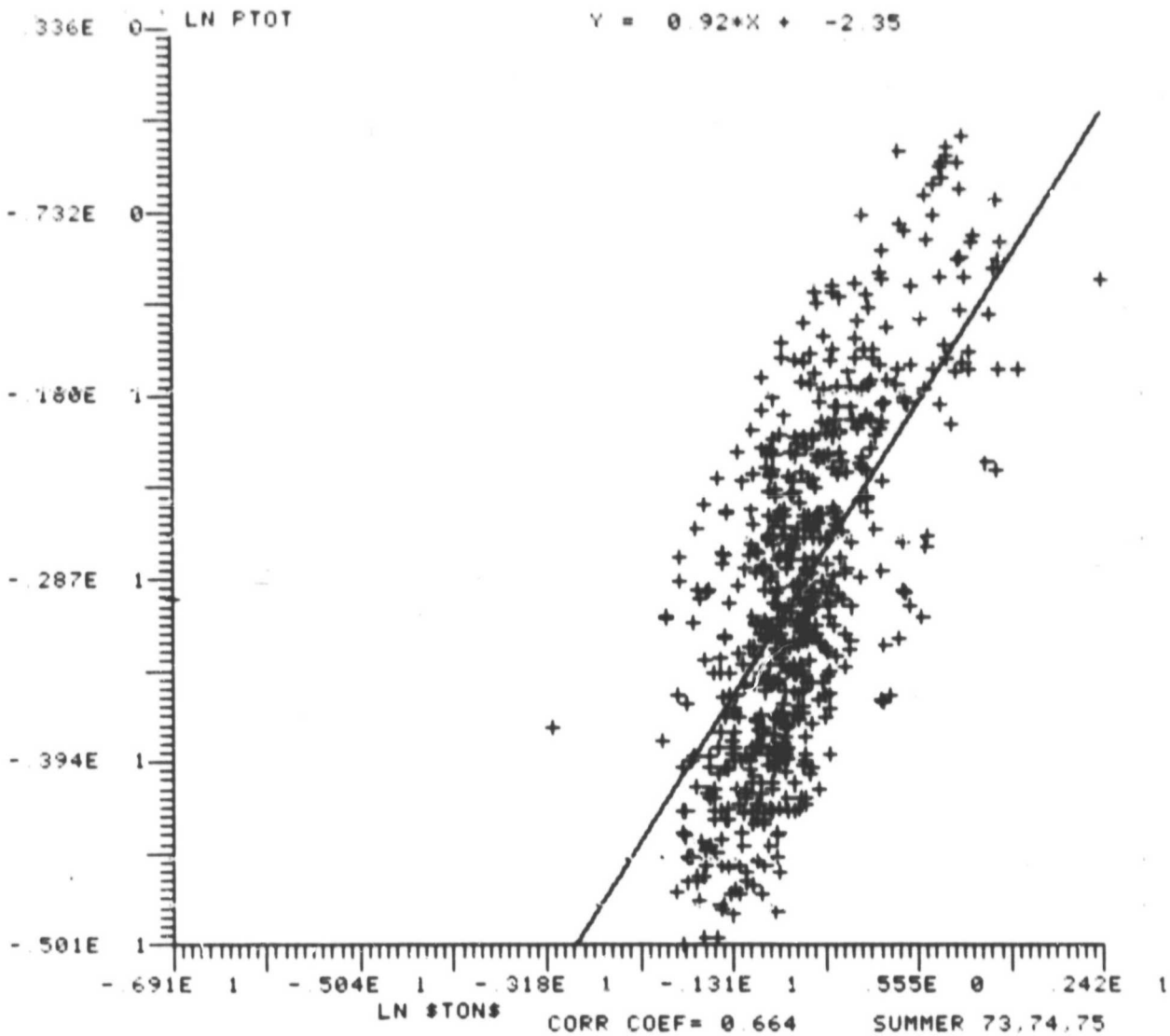

Figure I-11. A scattergram of the log of total organic nitrogen (TON) versus total phosphorus (PTOT) for the summers of 1973, 1974, 1975

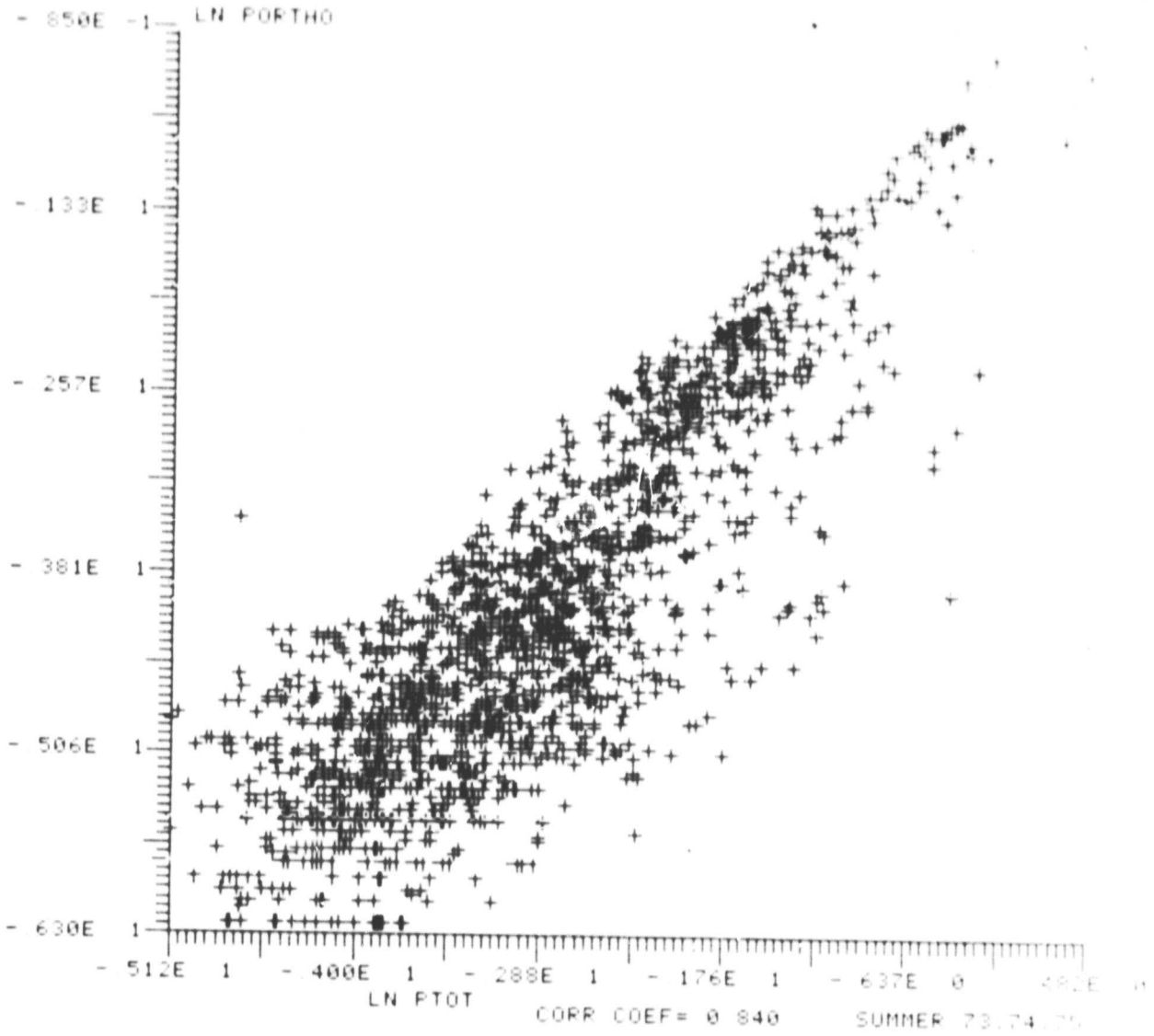

Figure I-12. A scattergram of the log of total phosphorus (PTOT) versus ortho-phosphorus (PORTHO). Note the expectedly high correlation coefficient

(2) Thresholding. It was often desirable to disregard samples which had a certain variate above or below some threshold value. Since turbid water has greatly different reflectance characteristics than nonturbid water, it was common practice to discard those samples with turbidity above some level.

(3) Normalization. Most water quality measures are in different units, e.g., mg/l, meters, etc.; correlation between them is difficult to ascertain. Normalization of each variate based upon its mean value and standard deviation allowed relationships between variates to be observed more readily.

(4) Ratioing. In many cases it is the ratio between water quality measures or reflectances that is important. A program was developed that allowed new variates to be computed that were arithmetic functions of existing variates. An example of a useful ratio would be the ratio between total organic nitrogen and total phosphorus.

(5) Principal Component Transformation. Correlation between many variates could be analyzed by performing a principal component transformation on selected sets of variates. This technique was also being investigated as a possible method of removing noise in the multivariate data.

2. Work Plan

At the time that the STORET work stopped, JPL had progressed to a point of trying to understand the interrelationships of the water quality measurements themselves. The aforementioned software had just been completed and an analyst was being trained in its use. The work plan for the future was as follows:

(1) Develop a reliable trophic state index from the water quality measurements. This implies the ability to compute a number whose value describes the trophic state of the water body from which the sample was taken. The STORET data gave us an excellent data base to work from and it is imagined would allow development of a trophic state index that could be applied over a larger regional area.

(2) Predict the trophic state index from remotely sensed data. Once the trophic state index was computed, it could be placed into our data base as an additional variate; then the existing software could be used to correlate multispectral remotely sensed data to the trophic state index.

(3) If the trophic state index could be reliably predicted, the next step was to predict the water quality measures from the remotely sensed data. More realistically, ratios or other relationships of the water quality measures are probably the best that could be done.

C. STATS2

During the course of the water quality task, it was frequently necessary to compute training area statistics of the multispectral imagery at precise locations

within the scene. These locations were typically those at which EPA personnel had taken water samples during the satellite or aircraft pass.

To facilitate the training area location process, an interactive statistical training area program was developed called STATS2. With STATS2, the analyst was able to view the scene on a CRT, interactively choose the spectral channel that gave the best rendition and contrast-enhance the scene. Through the use of a trackball-controlled cursor, the analyst could then outline the precise location for which multispectral statistics were desired. The program would then compute the statistics and save them on a disk file.

STATS2 also contained capabilities for editing, updating and merging of statistics. In addition, the analyst could perform a "first-cut" multispectral classification so as to decide whether or not the training area was well chosen and would perform well in a later classification procedure.

D. IBIS

Although the Image Based Information System (IBIS) was not designed specifically for the water quality monitoring tasks, one of its major initial applications was directly connected with watershed monitoring in the Lake Tahoe Basin. The IBIS system was originally designed as a computer-based approach to the analysis of geographical situations such as those associated with land use planning activities.

1. System Description

IBIS is composed of general purpose and specialized computer programs which can be grouped into logical steps to build an information data base. Functionally, IBIS represents a selection of programs which operate under the IPL's VICAR Image Processing System. The IBIS system is raster (image) based. Most data entered into the system are in raster format. However, the system is also designed to allow integration of graphical and tabular data as well (Ref. I-13).

Logical and mathematical interfaces are provided in the system to link all data files in an IBIS data base superstructure. Figure I-13 is a configuration diagram depicting the IBIS data base concept. By manipulating the interfaces, information can be derived from simple associations to or comparisons between two or more data files stored in the data base. More complex procedures, such as polygon overlay and cross-tabulations can be performed to acquire other significant data.

The image formatted data plane is the primary data type utilized in IBIS processing. IBIS data planes are obtained directly in image form, as from Landsat imagery, or the planes can be derived from data compiled by sources such as the U.S. Department of Agriculture, U.S. Bureau of the Census and the Defense Mapping Agency.

Most image data sets entered into the data base are derived from Landsat or other multispectral scanner sources. Since image data planes can be derived from different sources, often no common spatial alignment exists between them. The system thus provides a means of registering these data sets to obtain a unified spatial surface.

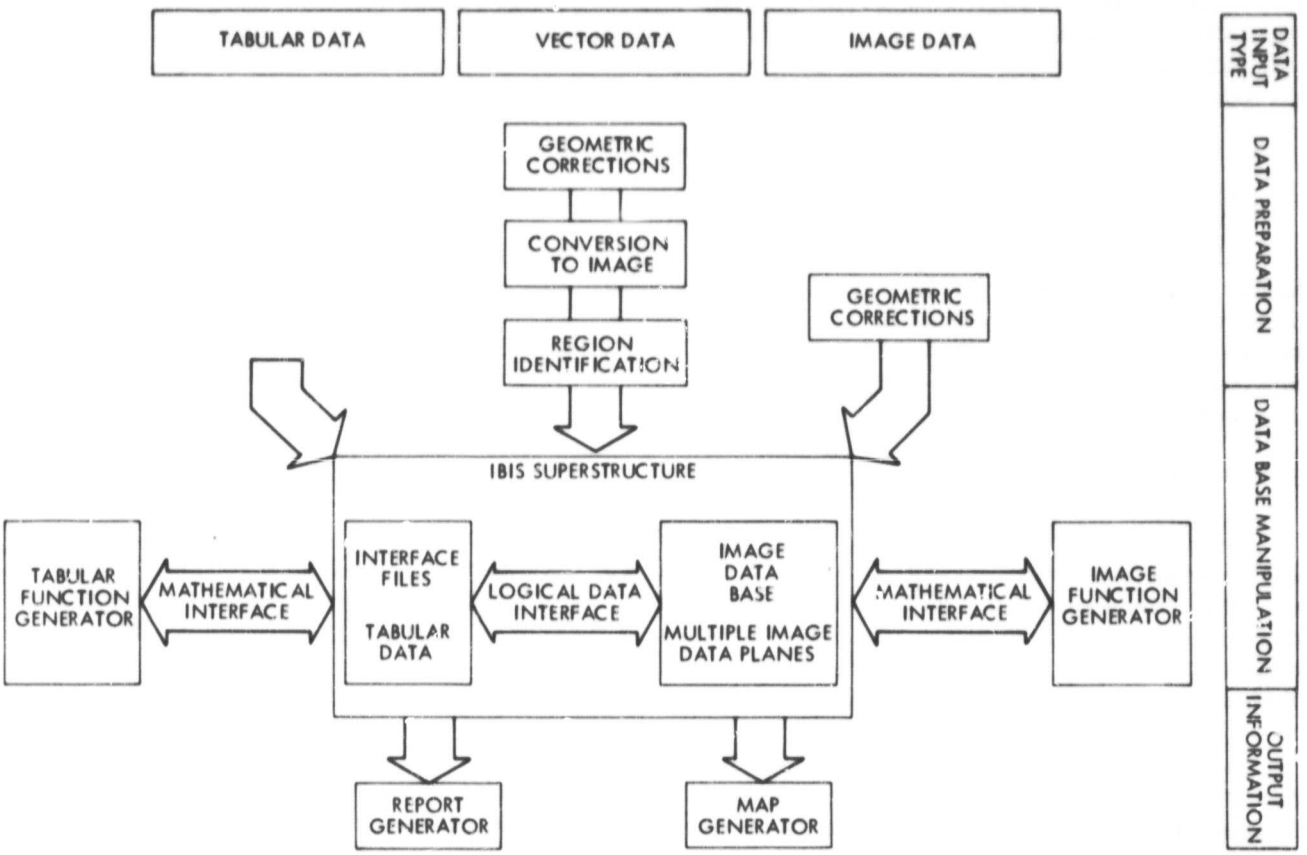

Figure I-13. A configuration diagram (from Ref. 1-9)

Graphical or vector data can also be entered into the IBIS data base. Graphical data, such as those obtained from conventional maps, are electronically digitized on a coordinate digitizer. As with image data, the graphical data files must be in registry with the primary data base. Geometric corrections are performed before the graphical files are transformed into image space using a two-step process. First, a general surface fit is achieved through the use of a least squares affine transformation. Then an exact geometric correspondence to the primary data base is obtained through a "rubber sheeting" procedure. Deformation of the original surface is controlled by selection of tiepoints which link geographical features identifiable on both the graphical data file and the primary data base.

Tabular data can be entered into IBIS through computer cards or digital tape. These data are stored in a tabular file which is linked to the data base through a logical interface.

The most important element of the IBIS concept is the georeference plane. The georeference plane is a map-based graphical representation of areas of interest, such as drainage basins in a watershed area. The georeference plane is constructed in the form of a polygon file which is registered to the primary data base and used in data aggregation and map generation procedures.

Once the georeference plane is transformed into image space, each polygon or region of interest is identified by assigning a unique data number (DN) to each region. After region identification, the georeference plane is used in higher order IBIS procedures such as polygon overlay or modeling. Each image plane is referenced to one or more georeference planes. All tabular files are also linked to at least one georeference plane. By its unique construction, the IBIS data base provides the user the ability to manipulate data from several sources, which despite their original disparity, are all referenced to a common base. Figure I-14 provides a conceptualization of the registered data plane format of the IBIS system.

2. Application of IBIS to Water Quality Studies

Major emphasis in past water quality studies conducted at the IPL had been placed on assessing the viability of Landsat as a monitoring device. A useable system for lake extraction and classification was developed as a result of these tasks. However, lakes were only examined through the comparison of certain water quality measurements with multispectral scanner data analysis. As a result, not much emphasis was placed on locating and defining any contributing factors, such as the impact of non-point source pollutants from the surrounding land mass. Possible sources for pollution were suggested only for the purposes of discussion in the analysis of classification results.

In order to conduct a more comprehensive study which attempted to define the causes and dynamics of lake pollution using remotely sensed data, the resources of IBIS were essential. Remotely sensed data, such as Landsat imagery, required augmentation by conventionally acquired data, such as hand-drawn maps, runoff measurements and precipitation data, if an accurate data base were to be produced. Similarly, conventionally acquired data were greatly enhanced when integrated with spectral information provided by satellite. The task of combining these data,

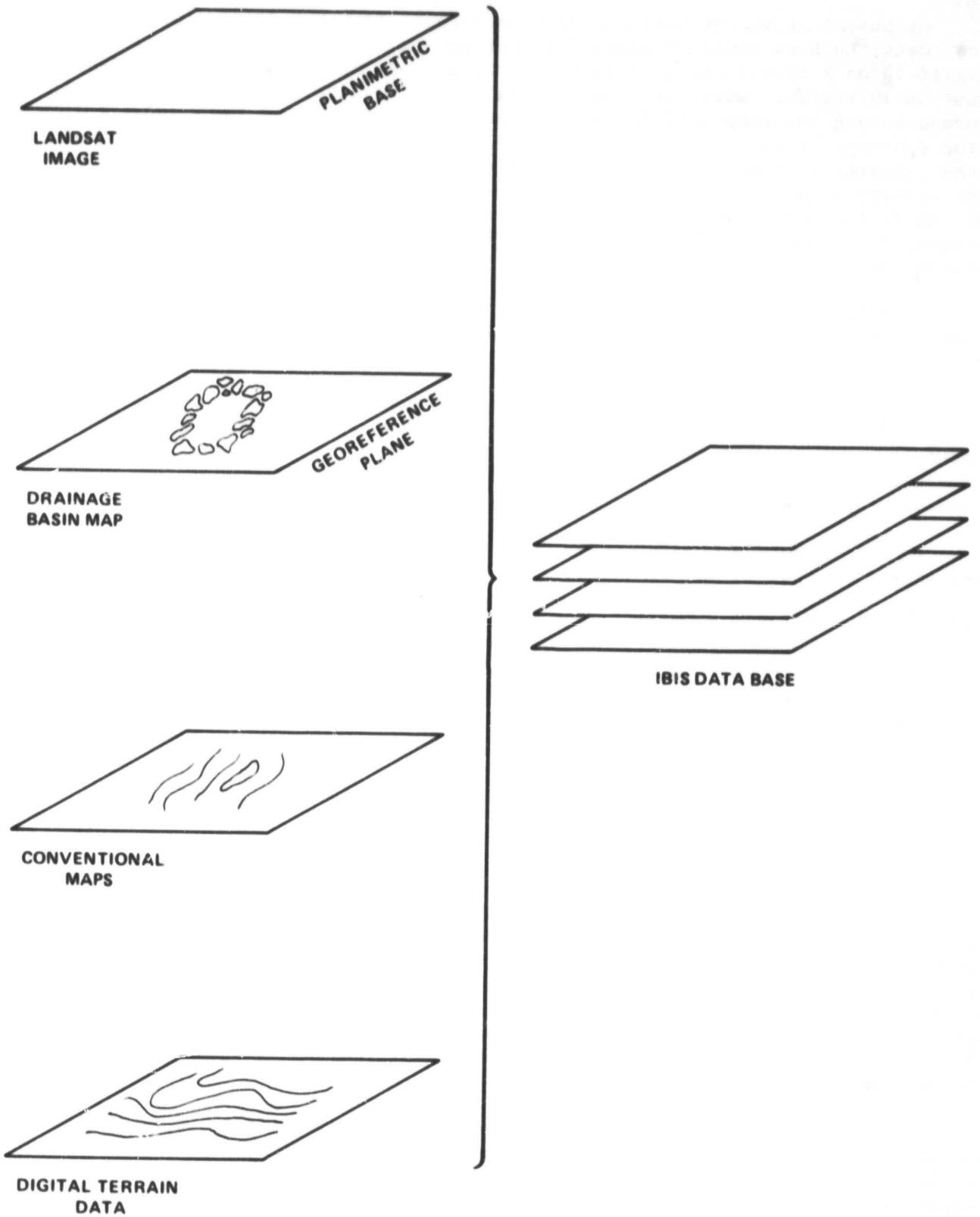

Figure I-14. Conceptual diagram depicting data planes in registration forming the IBIS data base

however, was enormous unless a system were available which could tie all data elements to a common base. Rather than viewing a precipitation map side by side with a Landsat scene, it was more desirable to examine the map as it appeared "overlayed" on the Landsat image. This technique could be easily accomplished using the IBIS resources.

Tremendous amounts of environmental data have been gathered for the watersheds of our nation's larger lakes. In the case of the Lake Tahoe Basin, the U.S. Department of Agriculture had conducted an intensive study of the basin's watershed in an effort to define controls on future land use. These data provided an excellent source of material which could be integrated with Landsat MSS and digital topographic data to not only test the IBIS concept but to add a new dimension to the water quality management projects.

SECTION II

TECHNOLOGY APPLICATIONS

This section details the techniques used and significant results of three demonstration projects conducted during the period of 1977 through 1979. The Lake Mead Intensive Area Study was completed in 1978. This study applied digital image processing techniques to Landsat data acquired over the Lake Mead area. Multispectral classification was performed and the results were compared with contact-sensed water quality measurements gathered by EPA personnel.

The Lake Tahoe study applied digital image processing techniques of the IBIS system to build an information data base for the Lake Tahoe watershed area. The data base was used to demonstrate the capabilities of the IBIS system and its applicability to a watershed monitoring situation. Landsat based images were created which combined conventional maps with Landsat data. Additionally some preliminary modelling of terrain data was attempted.

The study of selected lakes in Montana, Utah and Michigan was undertaken as a final demonstration of digital lake isolation and classification techniques using Landsat data. Multispectral classification was performed on some lakes selected for this study. Results are presented only in terms of digital image processing products as no analysis has yet been made by EPA personnel.

A. LAKE MEAD INTENSIVE AREA STUDY

1. Purpose of the Lake Mead Study

The study of Lake Mead was performed in order to establish the feasibility of monitoring the trophic condition of a water body over a specific period of time. This project was based upon the application of Landsat MSS data and airborne sensor-acquired data. These remotely sensed data were supported by ground truth measurements gathered by EPA limnologists. The period of Landsat coverage for the study was from June through September 1976 on Lake Mead, Nevada.

Trophic classification of Lake Mead proved to be most successful when performed by an unsupervised clustering routine. The results of unsupervised classification and a comparison with ground truth measurements are presented within this section.

The promising results obtained from the Colorado Lakes Study (Ref. I-2) and the Illinois Lakes Task (Ref. I-3) led to the definition of a project in which a single water body of considerable size would be more closely monitored. The intent of the project was to apply the skills and techniques which had been developed by the IPL during prior lake classification investigations to a new and unique situation. In the past the lakes surveyed had been extremely small, sometimes limited to only 30 pixels in a Landsat frame. Much effort had been expended in the development of the image processing skills necessary to perform the task of lake identification, extraction and analysis in a timely manner (Ref. I-5). The Lake Mead Intensive Area Study offered a test of the techniques which had evolved from the previous work. In addition, the study was oriented as a temporal investigation in which one relatively large lake would be characterized through the correlation of ground truth measurements and Landsat MSS analysis, and classified as to both trophic classes and turbidity levels on a space-time basis.

2. The Lake Mead Study Area

Lake Mead is the largest man-made lake in the United States. It was created by the interstate impoundment of the Colorado River by Hoover Dam in 1935. Lake Mead, the reservoir formed by this dam, was completely filled in by 1941 (Ref. II-1,2). Located in the Mojave Desert, as seen in Fig. II-1, the lake bed consists mainly of sand, silt and clay joined with some expansive areas of soluble beds of gypsum and rock salt (Ref. II-2). The multiple use of water from Lake Mead results in widely fluctuating water levels, which generally follow a pattern of slow drawdown from midsummer to the following spring when a rapid rise occurs (Refs. II-3,4). The water is hard, with total dissolved solids of about 800 ppm. The annual temperature cycle of Lake Mead can be classified as warm monomictic. This is characteristic of lakes of the warmer latitudes in which the temperature of the water never falls below 4°C at any depth, one circulation occurs each year in winter and the lake is directly stratified in summer (Ref. II-5). The reservoir provides municipal and industrial water, water for irrigation, hydroelectric power and a recreational area visited by over one million people each year (Ref. II-6).

Due to the proximity of Lake Mead to the USEPA's Environmental Monitoring and Support Laboratory at Las Vegas (EPA/LV), sampling surveys could be conveniently and cost-effectively conducted. Further, Lake Mead was relatively large when compared with lakes previously studied and meteorological conditions of the Lake Mead area seemed favorable in terms of cloud cover and ice conditions, which can severely limit the extent of Landsat scene availability. One of the most important factors determining the selection of Lake Mead was that the reservoir is situated within the 14% image overlap of the Landsat sensor scan path. This allows coverage on successive days, thereby providing not only more images to select from, but also an opportunity to test the validity of trophic classification from one day to the next.

3. Sampling Procedures

Following the selection of Lake Mead as the study site, the USEPA provided sampling crews to measure water quality parameters on the lake from early June through September 1976. Samples were collected by Bell-Huey pontoon-equipped helicopters, shown in Fig. II-2, which landed on the water at sampling stations ranging throughout the reservoir. Care was taken that the sampling crews conducted their surveys to coincide with the overpass of the Landsat satellites so as to insure the validity of the data comparison. Each sample site was identified and marked on maps by sighting on prominent land features so that the exact location could be resampled on the next Landsat flyover date. Samples were analyzed after acquisition at the Environmental Monitoring and Support Laboratory at Las Vegas and data recorded by the STORET system. However, due to weather conditions which restricted helicopter use and mechanical problems with the helicopters, sampling was severely restricted, as shown in Table II-1.

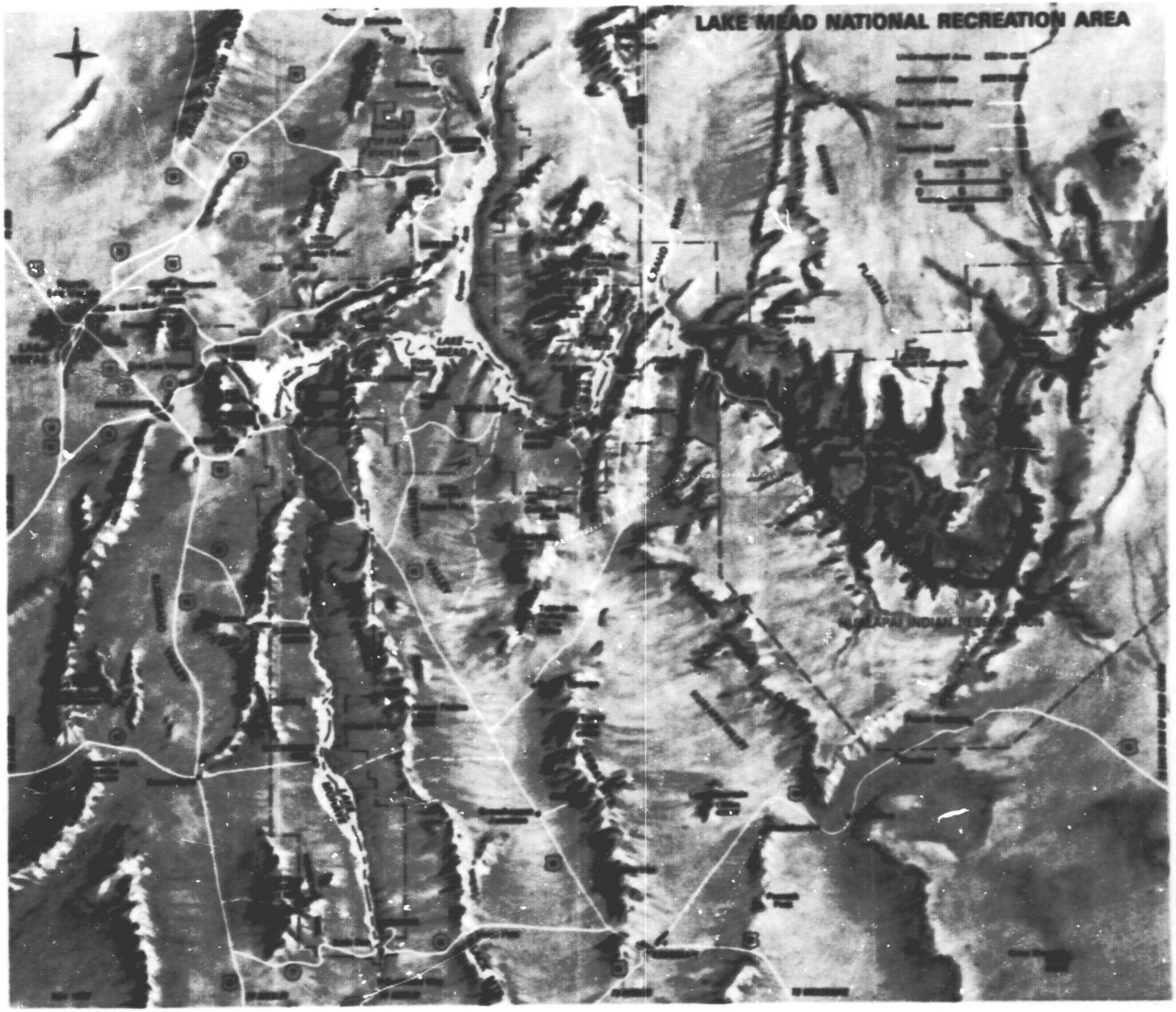

Figure II-1. Map of Lake Mead and surrounding area

Figure II-2. An EPA Bell-Huey helicopter prepares to land on a lake to obtain water samples

Table II-1. Landsat acquisition dates and corresponding ground truth sampling

Date, 1976	Frame No.	Sampled by EPA
June 11	5419-16584	
June 12	5420-17042	
June 20	2515-17254	
June 21	2516-17312	
June 29	5437-16572	
June 30	5438-17030	
July 8	2533-17251	
July 9	2534-17305	X
July 27	2552-17302	
August 4	5473-16542	X
August 5	5474-17000	X
August 15	2596-17241	
August 22	5491-16525	
August 23	5492-16583	X
August 31	2587-17234	
September 1	2588-17292	
September 18	2605-17231	
September 19	2606-17285	
September 27	5527-16495	X
September 28	5528-16553	X

The lack of ground truth data for all but 6 flyover dates limited accurate correlation of classification results vs trophic indicators to only 6 Landsat scenes. Of the six scenes, four were cloud-free over the lake surface area while two were obscured by cloud cover, as indicated in Table II-2.

Table II-2. Landsat frame number, date, and approximate cloud cover for sample acquisition periods

Date, 1976	Frame No.	Cloud Cover
July 9	2534-17305	Clear
August 4	5473-16542	Clear
August 5	5474-17000	Clear
August 23	5492-16583	Obscured
September 27	5527-16495	Clear
September 28	5528-16553	Obscured

4. Image Processing Methodology

The IPL has established a set of standard procedures to be applied to Landsat imagery before any follow-on processing is attempted. Figure II-3 depicts the processing sequence used on Lake Mead images. The preprocessing procedures consist mainly of cosmetic and geometric corrections. In terms of cosmetic processing, Landsat images are corrected for MSS line dropouts, such as slipped or missing lines and other obvious visual defects in the MSS imagery. In addition, Landsat-1 MSS images have been plagued by a striping problem which is the consequence of an imbalance in the sensor detectors. If this phenomenon occurs in a selected image, the frame is preprocessed to reduce the effect of the striping, also known as sixth-line banding.

Geometric corrections are achieved through the routine application of the VICAR program VERTSLOG. Data contained in the Landsat CCT's, as supplied by the EROS Data Center, are in a four spectral band interleaved format which is not compatible with the processing approaches established at the IPL. The IPL software program VERTSLOG unravels these interleaved data, producing a separate image for each spectral band. A history label is also produced which supplies the user with pertinent information about the image, such as latitude, longitude, scene identification and any processing history. The program also applies geometric corrections to the data, including corrections for mirror scan velocity, panorama correction and resampling of the data to create an approximately 80-meter instantaneous field of view (IFOV).

5. Registration of Extracted Lake Mead Subscenes

Due to the large number of Landsat images of Lake Mead which were to be used in the study, the resources of the IPL's library of over 300 documented image processing programs were used to aid in the development of a more easily wielded data base. The decision was made to extract subscenes of the Lake Mead image area from each frame and to register these subscenes to one reference image so that all pixel coordinates would be identical from scene to scene.

Frame 5473-16542 (August 4), depicted in Fig. II-4, was chosen to serve as the reference image to which all other Landsat frames would be registered. The frame appeared to be free from cloud cover and the entire lake surface was contained within the scene. The program PICREG was used to collect tiepoints which correlated the images to be registered. Figure II-5 schematically represents the operation of this program. PICREG is an interactive VICAR applications program designed to aid in the registration of two images. PICREG will display two pictures on a split-screen video display, with each picture window 512 lines by 320 samples. The user positions trackball-controlled cursors over identical features in each image and enters these tiepoints by keying in a predefined character. Several modes are available to improve the input tiepoint position, and the user has the capability of recursoring any bad points or of restarting at any position.

The correlation mode in PICREG was used, which attempts optimization of a tiepoint position using two areas of linear dimension 2^N. The default value N = 5 was used; thus the size of the correlation region was 32 pixels. PICREG requires that the user must choose at least five tiepoints. For the purposes of this

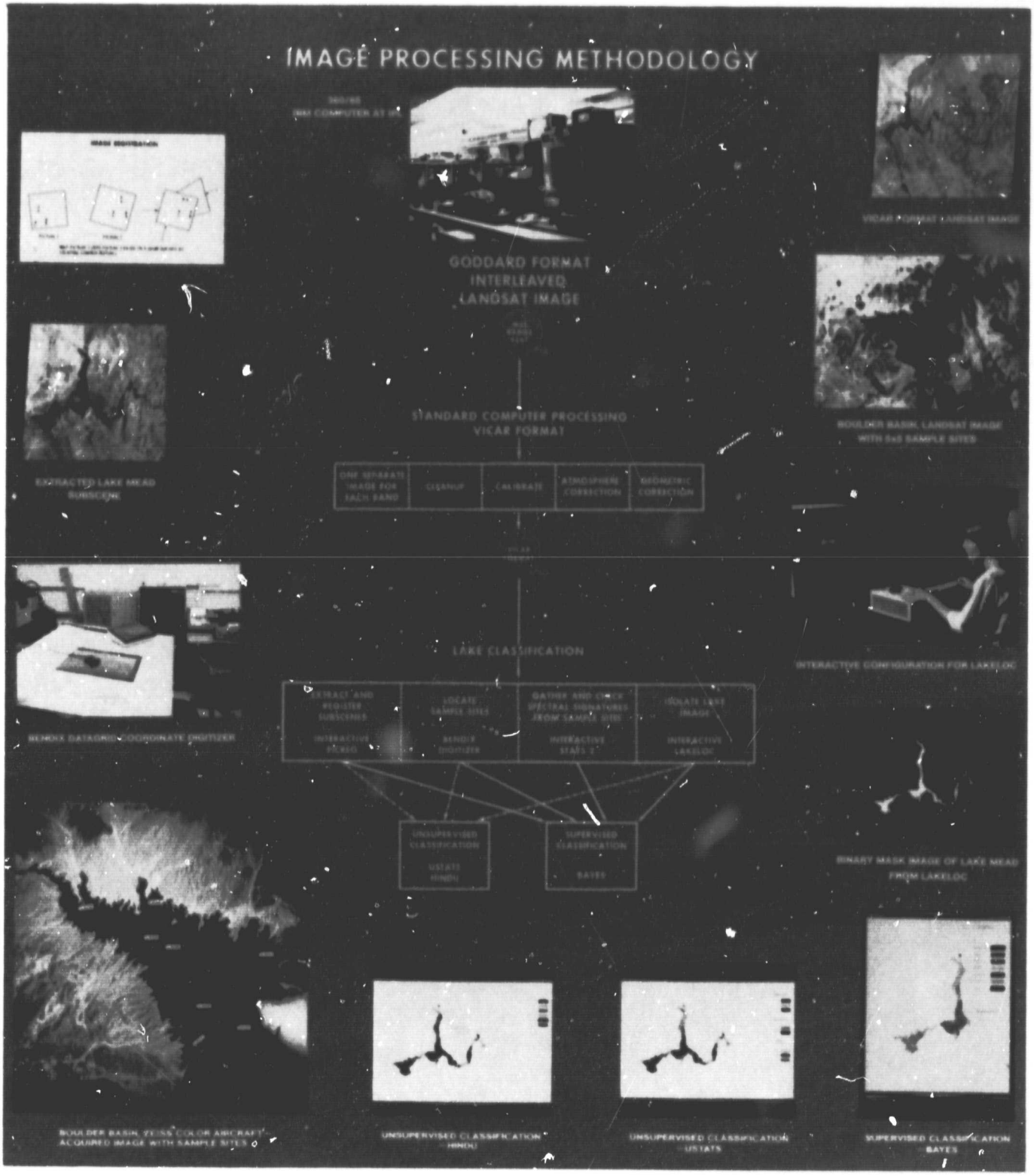

Figure II-3. A reproduction of a poster board used to describe the image processing methodology used at JPL for lake water quality monitoring

Figure II-4. Landsat frame 5473-16542, August 4, 1976, of the Lake Mead Study area. This image has been contrast-enhanced prior to false color production

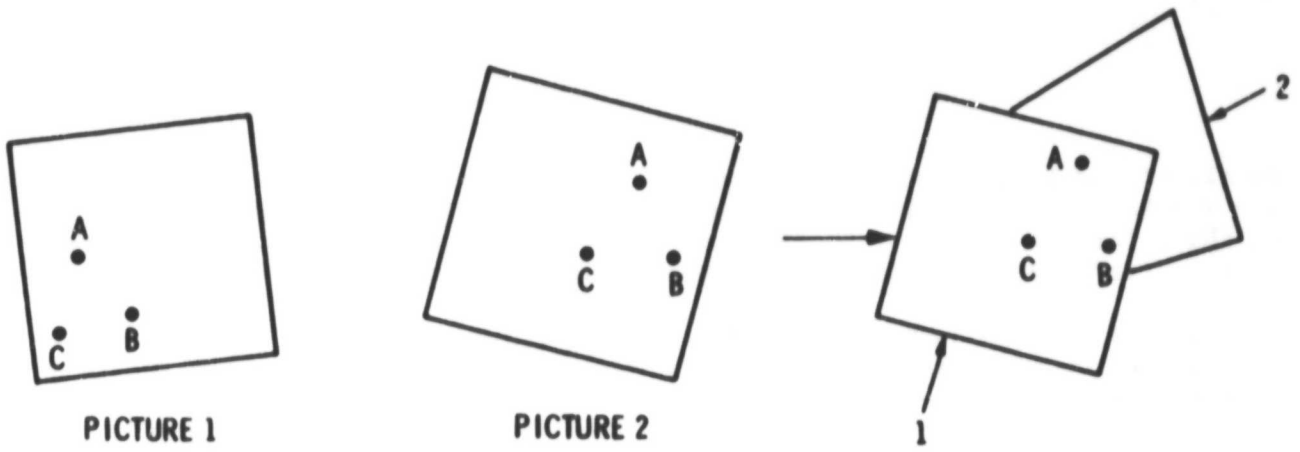

Figure II-5. A graphic interpretation of how the various Lake Mead images were registered to each other

project, governed by the size of the subscenes, at least 24 tiepoints were used, depending upon the amount of lake surface area which was contained in a particular subscene.

Tiepoints were located on geologic formations for the most part, as these formations were highly visible and relatively unchanging. The tiepoints were located within close proximity to the lake shoreline, but not on the actual shoreline. This procedure alleviated confusion associated with fluctuating water levels and shadows at the shoreline. The attempt was made to use the same geologic formations from frame to frame, if the formations appeared in the scene and were not obscured by cloud cover or shadow.

After the tiepoints were located and before the geometric transformation was performed, PICREG was allowed to reorganize the tiepoints through the "fit" option. The fit parameter allows PICREG to fit a surface, where (x,y) are the reference picture and (x',y') are the input picture,

$$y' = ax + by + cxy + d$$

$$x' = ex + fy + gxy + th$$

to all of the tiepoints and reinterpolate new tiepoints from this expression into a 5 x 5 matrix within the area encompassed by the originally specified tiepoints on the reference picture. At this point, the user could also check the accuracy of the "fit" and either try a new approach or accept the fit as computed.

If the fit was deemed acceptable, within 1 pixel accuracy, the geometric transformation was then performed. Following the transformation, the accuracy of the registration was checked visually by creating a difference picture using the reference image and the transformed image. All 20 Landsat scenes were registered to the control scene using PICREG. This created a data base in which the coordinates for each sample site would be identical for each Landsat image.

6. Sample Site Location

The NES handled the problem of accurately locating sample sites on the lake's surface by sighting the location of the helicopter as samples were gathered. The position was located by sighting on prominent land or cultural features with the helicopter compass and then calculating its distance from the shore and/or features. Locational data were then recorded in the field notes and later entered into the STORET system along with the trophic indicator data.

Color aerial transparencies of Lake Mead, taken at altitudes ranging from 23,200 feet to 25,900 feet using a Zeiss camera equipped with a 152.4 mm lens, were supplied by JPL to EPA/LV. EPA/LV, using the locational data from STORET, marked the location of each water sample station with a pinpoint and the corresponding STORET identification number and returned the photographs to JPL. Using these as a guide, the sample sites were then located on the reference image, frame 5473-16542 (August 4), with the aid of a Bendix Datagrid Digitizer. Figure II-6 shows the reference frame with sample sites outlined.

The matrix size of pixels to be used for the spectral analysis was 5 x 5 for all but four of the 49 total sample sites. Sites VR04, LV08, LV09 and IC01 were located in areas of Lake Mead in which the shoreline configuration necessitated that the site boundaries be adjusted to fit within the lake's confines. Table II-3 gives the dimensions of the pixel matrices for these four sites.

The Landsat imagery used by IPL is, in most cases, resampled to produce pixels measuring approximately 80 meters by 80 meters. More specifically, the resampling has produced square Landsat pixels which represent an earth surface distance of 79.98 meters per pixel edge. This resampled pixel represents an area of 6,396.8 meters^2 (68,854.6 feet^2) or 0.6396 hectares (1.58 acres). Therefore, the selected 5 x 5 pixel matrix represents 15.99 hectares or 39.54 acres on the lake's surface. Figure II-7 illustrates both one pixel and the 5 x 5 pixel matrix in relation to a standard U.S. one-mile-square section.

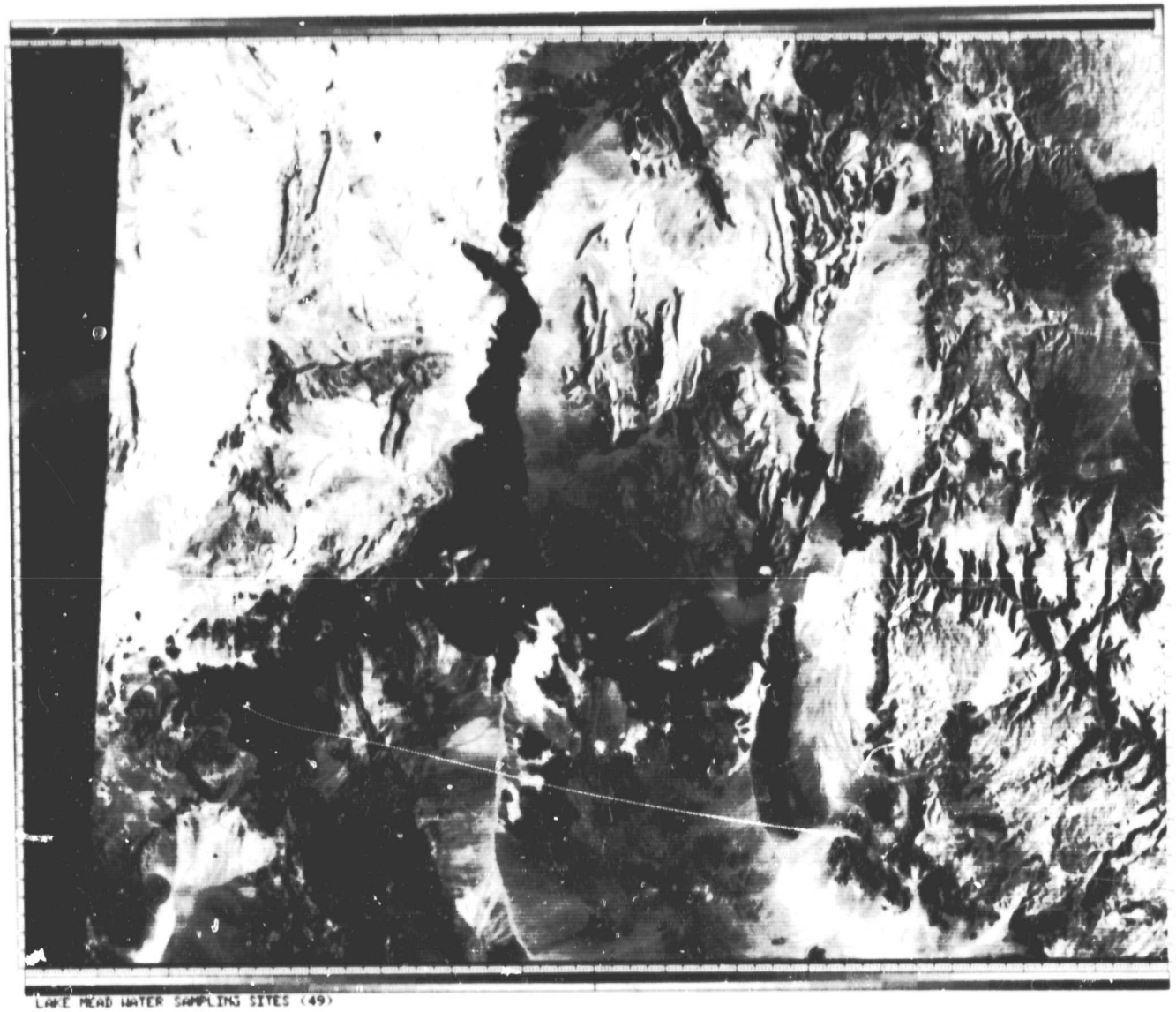

Figure II-6. Lake Mead with the water sample site locations marked by small white squares

Table II-3. Sample sites not conforming to the standard 5 x 5 matrix size

Sample site	Matrix size
VR04	5 x 2
LV08	2 x 2
LV09	1 x 5
IC01	5 x 2

The original intent of the project was to classify Lake Mead solely using supervised classification routines which were developed for the previous lake tasks (Refs. II-1,3,4). One limitation of this approach is that these programs depend heavily upon statistics gathered from sample sites. Many times sampling data for a lake are either out of date, simply too expensive to gather or not available. This necessitated the development of unsupervised routines which did

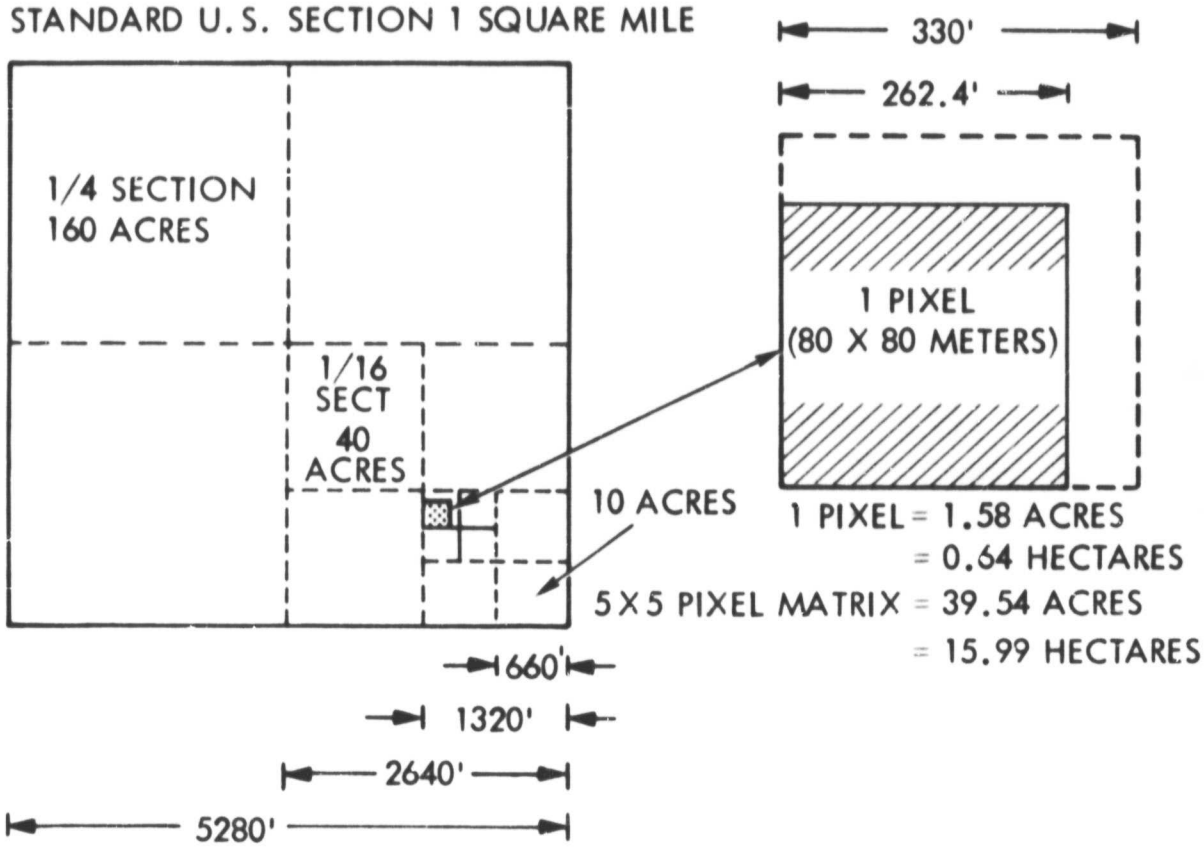

Figure II-7. A graphic representation showing the area of Landsat pixel(s) in relation to a standard U.S. section

not rely on training sites. During the course of the project, the IPL had become increasingly involved in the development of unsupervised clustering routines. An unsupervised program, USTATS, was applied to the August 4 scene 5473-16542. The results from the unsupervised routine were encouraging for this particular image, in terms of a comparison with the sample site measurements provided by EPA/LV. The unsupervised program had attractive benefits which enhanced its continued use by the project. For example, the routine did not require the somewhat timely gathering of MSS statistics and ranking of classes associated with the supervised routines used previously. A test of the supervised classification routine yielded classification maps which indicated confusion in class discrimination as indicated in Fig. II-8. It was felt that if the unsupervised classifier could prove to be as accurate as or more accurate than the supervised techniques, more emphasis would be placed upon its development and application to the Lake Mead study.

7. Statistical Gathering from Sample Sites

Although interest in unsupervised classification increased, it was decided to continue with procedures developed in earlier projects until the unsupervised technique could be refined and further investigated. To this end, sample sites which had been located and identified on the reference image were then located on all of the registered images through the application of the interactive program STATS2.

Follow-on supervised multispectral classification programs developed and used previously at IPL required the creation of MSS statistical data sets. The data sets contained pixel values for each "training" (or sample) site to be used in an MSS classification routine. The introduction of an interactive image processing system at IPL enabled programmers associated with this project to develop a fast routine for accurately locating training sites, computing the associated statistics and storing the information in a usable form. The program, STATS2, detailed in Section I-C, enabled the user to display a subject image of variable size on a video screen. The image was input in MSS format (interleaved) and contained up to four spectral channels. The image could be displayed in any one of the input channels and was viewed through a movable "window" of 512 lines by 640 elements. The program also offered a magnification option which enabled the user to accurately locate a training area and vary its size to remain within the boundaries of the subject area, such as a lake shoreline. The user could select training sites by either keying in predetermined coordinates or by locating a site visually and defining it through manipulation of a trackball controlled cursor.

Coordinates for this project were keyed in for the preregistered Landsat images and assigned a 5 x 5 matrix pixel density. The matrix size was changed from 5 x 5 only if the sample site overlapped the lake boundaries, as in the case of sites located in a few narrow areas. After a site was located and drawn, a command would be given to the program to display the statistical information for the site. Statistics were returned for all input channels in the form of $\bar{x}$ and sigma and the size of the pixel matrix was also displayed. A unique name was assigned by the user to the site at the time it was saved, which served as identification in future reference. The NES STORET number was used to identify Lake Mead water sample sites. Once a site was saved, the user returned to the processor to continue with the selection of the next site.

Figure II-8. A supervised classification image of Lake Mead using a principal component transformation of 10 trophic indicators as a basis to effect classification

Although eventually the sample site statistics were not heavily relied upon for classification, the statistics did provide a most important factor in the discovery and resolution of data discrepancies which plagued the project.

8. Data Discrepancies

As the MSS statistical data sets were being created by STATS2, comparisons were made with day-to-day spectral signatures to check the reliability of the MSS data. When spectral signatures from sample sites in the June 11 frame 5419-16584 were compared with signatures from the following day's coverage (5420-17042), the presence of serious inconsistencies became obvious, as illustrated in Table 3. As the operation of STATS2 continued on subsequent images, further comparisons were made and the problem was found to be occurring to some degree in all consecutive day coverage. In the case of a comparison between signatures on August 22 frame 5491-16525 and August 23 frame 5492-16583, the discrepancy ranged up to 10 DN, as seen in Table II-4.

Several possible causes for the inconsistencies were considered and rejected. Examination of the EPA water quality measurements, for conflicting sample sites on consecutive days, did not support a theory that the water quality had changed enough to seriously affect the MSS spectral response. Atmospheric changes were also considered. The effects of clouds and atmospheric haze on the spectral response of Landsat sensors has been documented (Refs. II-7,8). However, a comparison of images which contained cloud cover with cloudless scenes did not entirely support the atmospheric disturbance theory. Further, as suggested by Everett, Leonhart and Lepley in their investigation (Ref. II-1), the remoteness of the Lake Mead area serves to minimize the probability of atmospheric disturbances related at least to human activities. This was precisely a governing reason for selecting the area, both in their study and in this project. Had the data been acquired by both Landsat-1 and Landsat-2, a theory regarding the possible difference between the spectral response of the two spacecraft sensors would have been investigated. However, all consecutive day coverage was provided by the 14% overlap in scan paths from Landsat-2.

Having virtually ruled out all suggested considerations, it was concluded that 1) the discrepancies could have been caused by variations in the sensor response which were not detected and therefore not corrected in the ground calibration procedure or that 2) atmospheric changes had occurred which were unknown and undetectable.

9. Resolution of Data Discrepancies

Regardless of the cause, the discrepancies had to be corrected so that the project could continue. In deriving a solution to the problem, two considerations were of primary concern: 1) the proposed correction should not degrade the data quality; 2) due to time considerations and cost constraints, the correction had to be efficient and simple enough not to require large amounts of computer processing time.

Table II-4. Examples of discrepancies

Date, 1976	Site	Band channel	DN
June 11	0A02	4	45.1
June 12	0A02	4	40.1
June 29	BB05	4	50.2
June 30	BB05	4	44.9
June 29	0A02	4	57.8
June 30	0A02	4	53.7
August 22	BB05	4	46.7
August 23	BB05	4	36.6

Three ideas were proposed and judged on the basis of these two concerns. The first suggestion, to attempt to recalibrate the data using the calibration wedge which is provided with the Landsat imagery, did not satisfy criterion #1. Previous attempts to use the calibration wedge at IPL for similar purposes had proven ineffective. The calibration wedge provided with the Landsat CCT's contains only six sample points, of which two are typically saturated. The remaining four points do not provide enough information to properly recalibrate the data. It was felt that further data quality degradation was highly probable with this method.

The generation of an atmospheric model, which could account for some of the error, was also considered. Work has been performed in this area by other institutions with a degree of success (Ref. II-9). However, it was felt that this approach did not satisfy the time constraints of criterion #2. The procedure would have to be applied to 20 separate Landsat scenes. The computer processing time would have been costly in both hours and money.

The third alternative was to devise a normalization procedure which could operate on the MSS (interleaved) data in its present form, correcting all four spectral bands in a single operation. This hopefully would satisfy the time constraints. It was thought that by selecting a "control area" from one Landsat scene which was of relatively unchanging spectral response, it could be used to essentially normalize or conform the identical area in the remaining Landsat images. Similarly, all pixels comprising each Landsat frame would then be normalized, based upon the differences between the standard frame's control area and the same area in the subsequent frames. By constraining the spectral signatures to be constant in the control area, it was believed that a correction would be made for both the sensing system and any atmospheric effects simultaneously. Initial tests indicated that this procedure should be selected due to the ease of implementation and the results it could produce.

The control area, as seen in Fig. II-9, was selected in the Virgin Basin area of Lake Mead. A comparison of several test sites ranging throughout the lake indicated the Virgin Basin's spectral response to be the most stable in all the Landsat scenes. The spectral signature for the Virgin Basin Control area was calculated

Figure II-9. A false color image of Lake Mead showing the approximate location in the Virgin Basin where MSS data for 13 days were compared

for each of the Landsat scenes. The August 4 image 5473-16542 was selected as the standard control area to which all other scenes would be normalized. The August 4 scene was cloud-free and did not exhibit any haze. The DN differences between the standard control area and a given frame were then added to or subtracted from each picture element in the given frame. This procedure proved to be effective, inexpensive and easily implemented into the project plan.

The effect of the normalization procedure performed on the Lake Mead scenes was analyzed in two ways: 1) the statistical effect on the data was analyzed, and 2) the effect on clarsification performance was examined.

In the first case, the standard deviation of selected sample sites from all the scenes was computed for each spectral channel using STATS2. The standard deviations dropped significantly after the normalization was performed. As an example, the standard deviation for the Overton Arm sample site (OA02) dropped from 9.8 to 1.7 in band 4. Line plots (Fig. II-10) of the sample site DN's vs. scene (date) illustrate the effect the normalization had on stabilizing DN variation from scene to scene. Figure II-10 depicts line plots comparing Landsat data before and after normalization was applied.

To test the effect on classification performance, an unsupervised classification was performed on both June 11 (5419-16584) and June 12 (5420-17042) scenes. Using an unsupervised routine, clustering of the water signatures was performed on the June 12 frame before normalization was applied. The statistics for these clusters were applied to both the June 11 and the June 12 data. The same procedure was then followed for the normalized output of the same frames. The results indicated that the normalization had caused the June 11 scene to classify more closely to the June 12 scene, as shown in Table II-5.

10. Final Data Preparation for Classification

The data discrepancies having been resolved, the project moved into the final phase of data preparation prior to classification. At this point, each image subscene contained spectral information for the water surface area of Lake Mead as well as the surrounding land features. Since only the water data was of interest, the land surface area information was "removed" from the image before classification. This allowed the clustering routine to operate most efficiently, by avoiding confusion and unnecessary computer time in water/land discrimination. The resources of the LAKELOC program were used to isolate Lake Mead from the surrounding land mass. Figure II-11 depicts a Lake Mead subscene in the final binary mask form as output from LAKELOC.

11. Lake Mead Classification Summary

The thematic maps illustrated in this section were all derived from the application of USTATS. USTATS is an unsupervised routine which accumulates statistical information about a scene using local clustering. Results from unsupervised classifications using USTATS were more satisfactory than those obtained from supervised techniques used in the past. The algorithm implemented in USTATS was originally developed at Pennsylvania State University under the name CLUS, and has undergone extensive modification to run in the VICAR system. After USTATS has

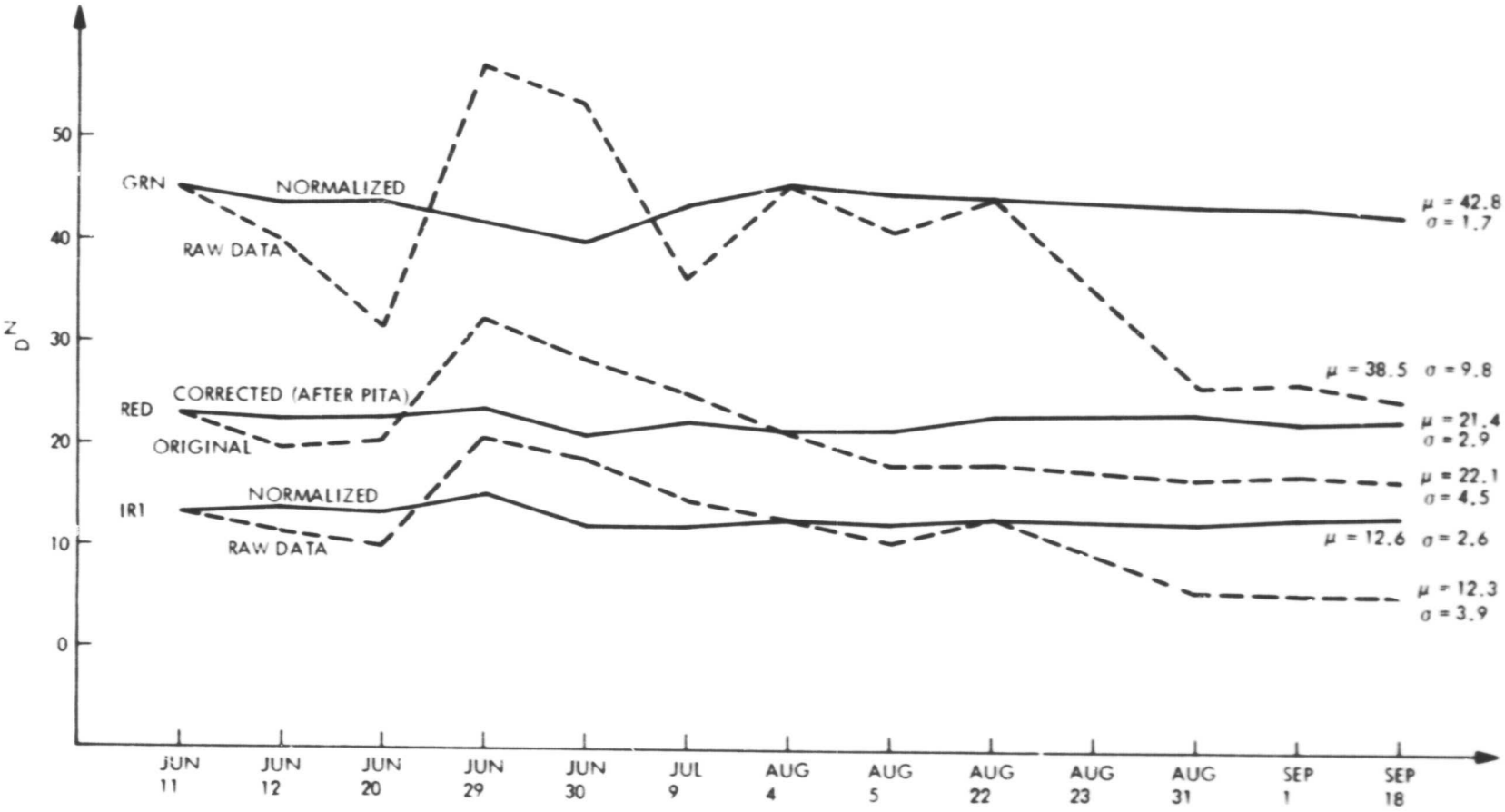

Figure II-10. A plot showing raw and normalized MSS means and standard deviations for the green, red and IR1 channels. Data shown are from the center of the Virgin Basin for the dates indicated

Table II-5. Effects of normalization on classification performance

Class No.	Percentage distribution by class		
	June 12	June 11 (Original)	June 11 (Normalized)
1	65.2	16.2	55.0
2	16.7	37.3	16.2
3	1.5	14.0	5.1
4	7.5	4.8*	0.6*
5	6.9	17.6	10.6
6	1.1	3.3	2.6
7	1.1	0.5	3.6

*Classification affected by clouds.

completed the cluster process, it passes statistics to the VICAR program BAYES, a Bayesian maximum likelihood classifier. Bayes performs the actual classification using a non-Euclidean distance function to decide which cluster signature is most similar to that of a given pixel. Figures II-12, 13, 14, and 15 are thematic photomaps of Landsat scenes for August 4, August 5, July 9 and September 27 respectively. These four scenes are the only ones which have supporting ground truth measurements which can be used to compare classification accuracy.

Visual examination of these maps indicated they provide limnologists with an excellent synoptic view of Lake Mead. In addition, the areal extent of river plumes, as seen in the Overton Arm and Pearce Ferry areas, was successfully mapped. Spectral classes were correlated most effectively with Secchi depth measurements, one percent light level and nephelometric turbidity. These indicators were expected to correlate well due to their direct correspondence with visual phenomena produced by turbidity and thus affect the spectral response. Chlorophyll content was also fairly successfully correlated with spectral classes. Complete analysis of the classification results is presented within the following section.

Two deficiencies evidenced by previous lake classification projects, conducted by the IPL, were encountered again by this project. Visual correlations between spectral classes with total phosphorous and total organic nitrogen measurements were difficult to construct. The sixth-line banding problem which has plagued Landsat data, and the data users, continued to degrade the resulting classification products, despite attempts to minimize its effect.

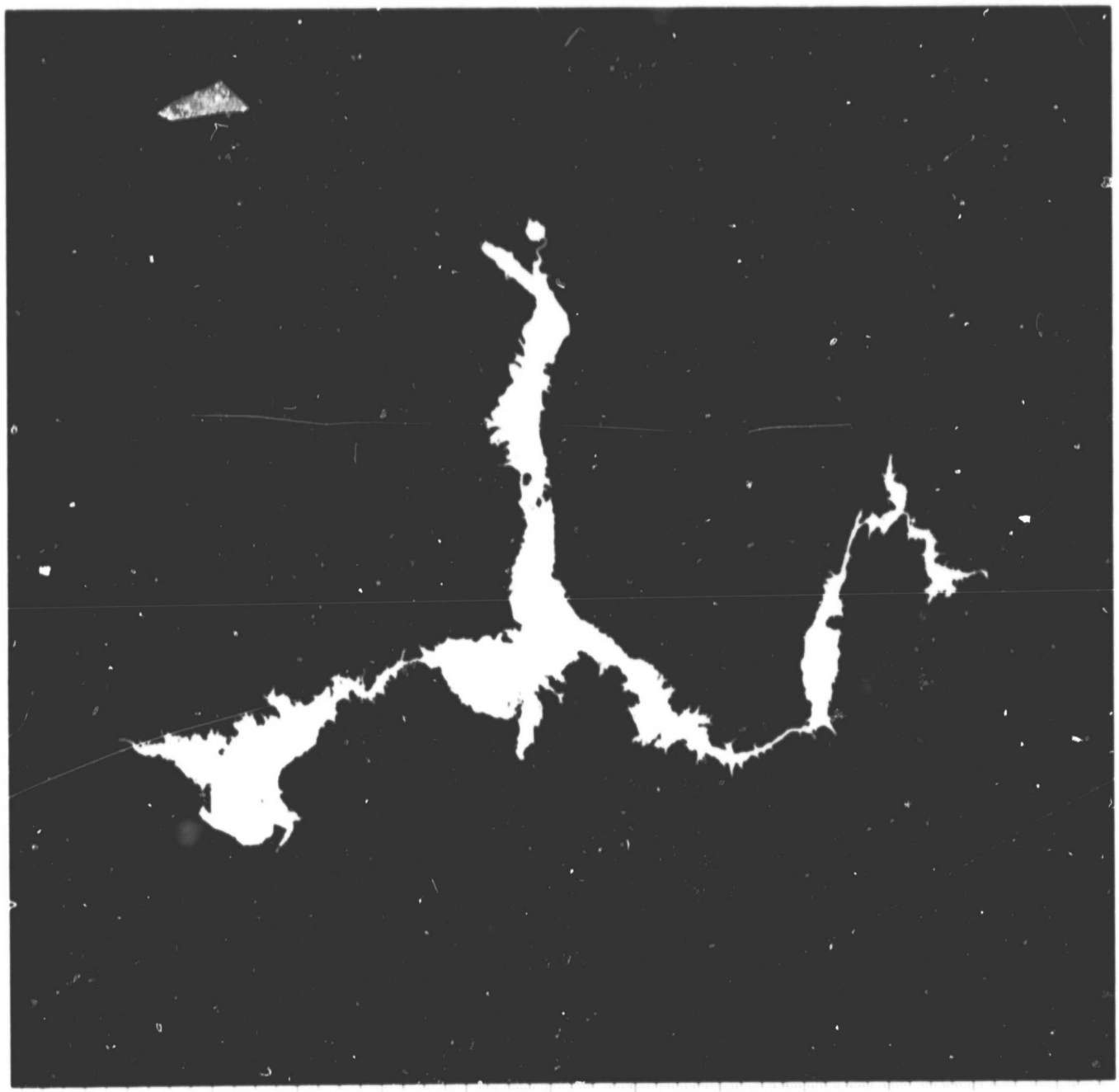

Figure II-11. A binary mask of Lake Mead

Despite these deficiencies, Landsat demonstrated great potential for use in the monitoring of changes in Lake Mead. This is evidenced by an examination of Table II-6, which correlated the results from the unsupervised classification of four Landsat scenes with corresponding water quality measurements made from surface samples on the same Landsat flyover dates.

Figure II-12. Unsupervised classification of Lake Mead for August 4, 1976

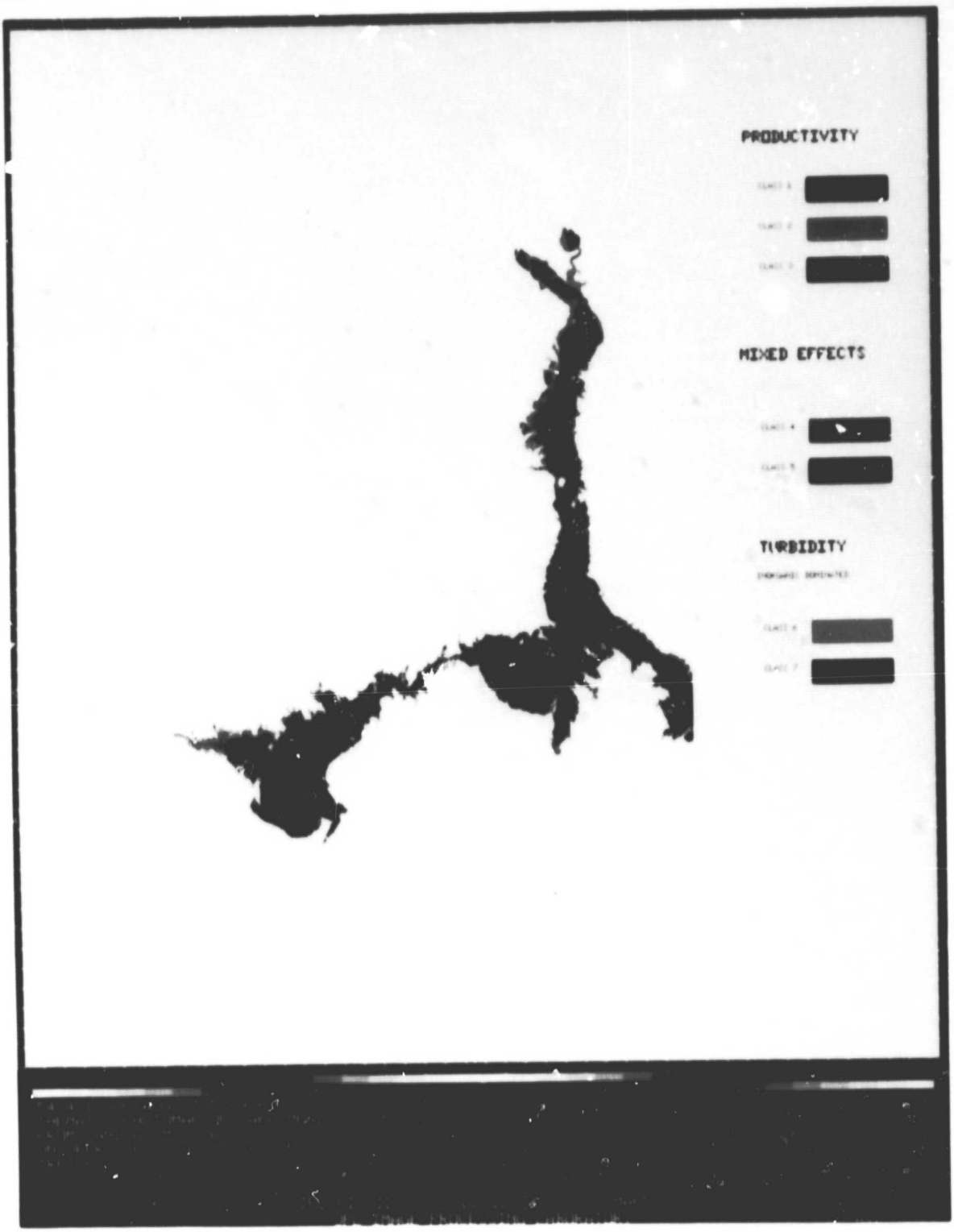

Figure II-13. Unsupervised classification of Lake Mead for August 5, 1976

PRECEDING PAGE BLANK NOT FILMED

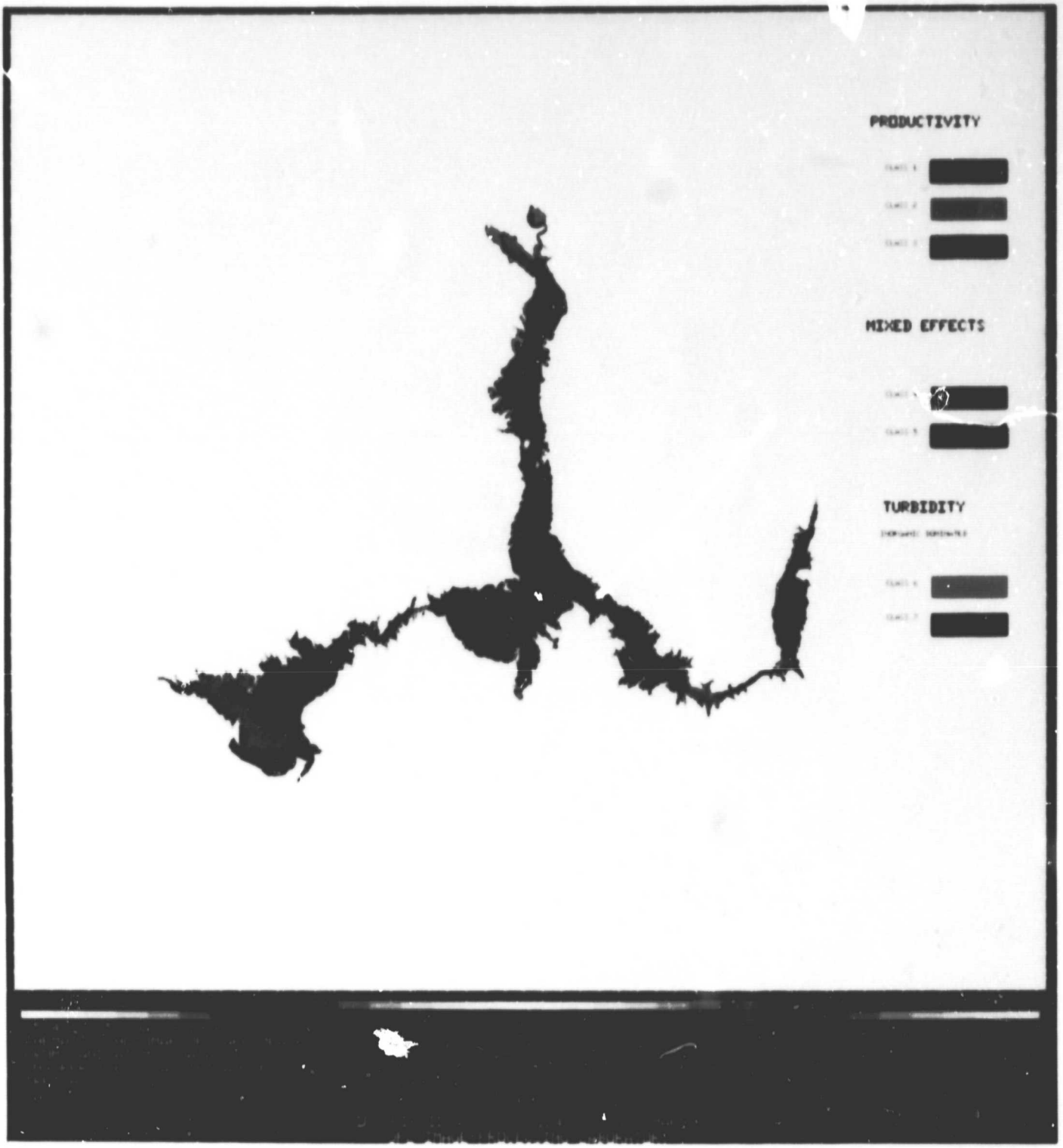

Figure II-14. Unsupervised classification of Lake Mead for July 9, 1976

ORIGINAL PAGE
COLOR PHOTOGRAPH

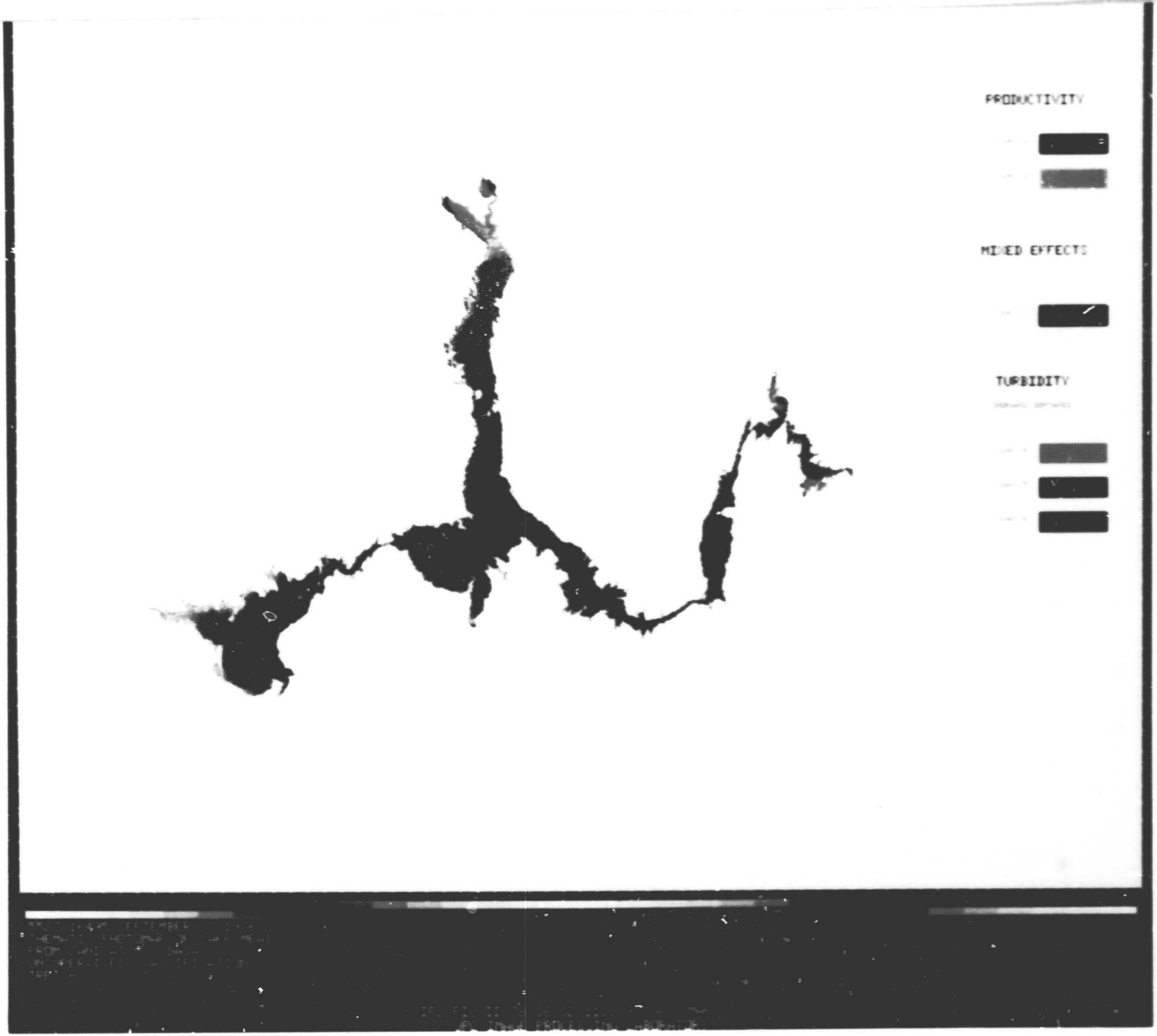

Figure II-15. Unsupervised classification of Lake M[illegible] for September 27, 1976

Table II-6. Correlation of four Landsat scenes with corresponding water quality measurements

		SECCHI DISC	1% LIGHT LEVEL	NEPHELOMETRIC TURBIDITY	CHLOROPHYL A	TOTAL PHOSPHORUS	TOTAL ORGANIC NITROGEN	PIXELS	HECTARES	% OF TOTAL AREA
PRODUCTIVITY	AUG 4	4 62 [illegible] – 6 10 N = 17	9 67 7 30 – 15 00 N = 22	89 30 – 1 70 N = 21	2 2 9 – 3 6 N = 9	022 010 – 097 N = 21	27 04 – 63 N = 17	5.465	32.298	63 4
	AUG 5	4 74 3 06 – 6 10 N = 13	10 02 7 80 – 15 00 N = 18	1 06 58 – 1 70 N = 18	2 2 8 – 3 6 N = 8	024 010 – 097 N = 17	33 14 – 63 N = 13	48.095	30.780	64 6
	JULY 9	5 49 4 88 – 7 01 N = 5	13 67 13 11 – 14 63 N = 2	32 28 – 37 N = 6	8 4 – 1 1 N = 5	012 009 – 015 N = 6	NO DATA	49.981	31.988	66 7
	SEPT 27	4 78 2 44 – 9 14 N = 16	9 14 6 50 – 13 00 N = 19	95 32 – 1 80 N = 19	1 2 1 – 2 60 N = 18	017 003 – 083 N = 19	61 38 – 1 03 N = 19	58.421	37.789	73 4
	AUG 4	2 65 2 13 – 3 68 N = 3	5 13 3 70 – 7 50 N = 3	1 39 68 – 2 00 N = 3	6 7 2 1 – 11 5 N = 9	027 016 – 048 N = 3	21 04 – 44 N = 3	10.732	6.889	13 5
	AUG 5	2 44 2 13 – 3 35 N = 4	5 33 3 70 – 9 00 N = 4	1 40 78 – 2 00 N = 4	2 2 8 – 3 6 N = 8	032 013 – 048 N = 4	23 12 – 44 N = 3	3.485	2.230	5 4
	JULY 9	3 96 3 96 – 3 96 N = 3	13 11 N = 1	30 27 – 32 N = 2	1 4 6 – 2 3 N = 3	014 011 – 015 N = 3	NO DATA	13.168	8.428	17 6
	SEPT 27	1 68 1 52 – 2 01 N = 4	4 75 3 50 – 5 60 N = 4	93 46 – 1 2 N = 4	2 1 9 – 3 4 N = 4	034 014 – 052 N = 4	70 57 – 86 N = 4	7.980	5.107	10 0
MIXED EFFECTS	AUG 4	1 83 N = 1	7 00 N = 1	62 N = 1	N = 0	016 N = 1	NO DATA	1.768	1.132	2 2
	AUG 5	1 42 91 – 1 83 N = 3	4 77 2 80 – 7 00 N = 4	1 11 62 – 1 80 N = 3	2 1 N = 1	023 016 – 030 N = 3	42 43 – 51 N = 2	1.301	833	2 0
	JULY 9	3 66 2 13 – 5 18 N = 2	8 26 N = 1	63 34 – 91 N = 2	5 55 6 – 10 5 N = 2	021 011 – 031 N = 2	NO DATA	2.158	1.381	2 9
	SEPT 27	NO CLASS GREEN FOR SEPTEMBER 27								
	AUG 4	CLASS GRAY ASSOCIATED WITH SIXTH LINE BANDING						3.845	2.461	4 8
	AUG 5	NO CONTACT SENSED DATA						3.947	2.526	6 1
	JULY 9	NO CLASS GRAY FOR SEPTEMBER 27						859	550	1 0
	AUG 4									
	AUG 4	CLASS BLACK ASSOCIATED WITH BOTTOM EFFECTS SHADOW AND SUBMERGED PLANT COVER NO CONTACT SENSED DATA						8.460	5.414	10 6
	AUG 5							3.513	2.248	5 5
	JULY 9							3.202	2.248	4 3
	SEPT 27							4.147	2.654	2 3
TURBIDITY	AUG 4	1 02 61 – 1 83 N = 6	3 18 2 30 – 4 50 N = 5	2 03 92 – 5 20 N = 6	2 8 1 0 – 5 6 N = 4	027 021 – 033 N = 6	24 0[illegible] – 51 N = 6	2.836	1.815	3 6
	AUG 5	1 46 1 02 – 1 83 N = 3	3 70 3 00 – 4 40 N = 2	2 23 1 40 – 3 60 N = 3	3 7 1 5 – 5 1 N = 3	023 020 – 028 N = 3	22 12 – 32 N = 2	3.614	2.313	5 6
	JULY 9	1 22 91 – 1 52 N = 2	NO DATA	32 30 – 34 N = 2	9 2 7 9 – 10 5 N = 2	061 049 – 073 N = 2	NO DATA	4.823	3.088	6 4
	SEPT 27	1 31 56 – 2 01 N = 12	3 89 2 50 – 5 50 N = 13	1 49 46 – 2 80 N = 14	2 3 1 – 11 1 N = 13	031 006 – 014 N = 14	71 41 – 1 17 N = 14	6.735	4.310	6
	AUG 4	53 31 – 86 N = 3	1 80 80 – 2 30 N = 3	73 1 80 – 18 00 N = 3	6 3 5 8 – 6 3 N = 2	024 016 – 030 N = 3	17 07 – 3 6 N = 3	1.268	812	1 6
	AUG 5	55 31 – 86 N = 4	1 93 80 – 2 30 N = 4	6 78 1 80 – 18 00 N = 4	6 3 5 8 – 6 8 N = 2	023 0 16 – 030 N = 4	36 N = 1	524	335	8
	JULY 9	46 21 – 89 N = 4	NO DATA	5 41 1 37 – 11 00 N = 4	5 85 4 7 – 7 4 N = 4	020 012 – 027 N = 4	NO DATA	794	508	1 1
	SEPT 27	61 61 – 61 N = 2	NO DATA	2 30 1 5 – 3 6 N = 3	8 1 – 1 2 N = 3	033 017 – 051 N = 3	91 48 – 1 16 N = 3	1.809	1.158	2 3
	AUG 4	10 N = 1	40	175 N = 1	9 N = 1	179 N = 1	1 06 N = 1	226	145	3
	AUG 5	NO CLASS ORANGE FOR AUGUST 5								
	JULY 9	NO CLASS ORANGE FOR JULY 9								
	SEPT	18 N = 1	N = 1	58 1	5 N = 1	155 N = 1	80 N = 1	458	293	6

12. Analysis of Classification Results

Reference is made in the following text to limnological measurements and conditions which may be unfamiliar. Briefly, the productivity of an ecosystem, such as a lake, connotes a quality whereby living substance is manufactured through interactions of community and environment (Ref. II-6). The productivity or fertility of a lake depends on nutrients received from regional drainage, on the geological age, and on the depth of the water body (Ref. II-10). Turbidity is the degree of opaqueness produced by suspended particulate matter (Ref. II-6). Secchi Disc transparency (SDT), one percent light level (1% level), nephelometric turbidity (NT) and chlorophyll-a (CHLA) are measurements which are indicators of productivity and turbidity.

The statistics, illustrated in Table II-6, were derived from the correlation of the thematic maps with contact-sensed data in the following manner. An image polygon output from STATS2, containing the 5 x 5 pixel sample site matrices, was overlayed on each thematic map. Each sample site matrix was then examined for its homogeneity as evidenced by the color or colors represented within it. Sites located within highly heterogeneous matrices were dropped from the analysis. Although it is recognized that a pixel measures 80 x 80 meters as processed by the IPL and that a sampling site is readily contained within a pixel, a degree of uncertainty exists as to which pixel within the matrix is the one containing the sampling point. Thus it was decided to drop all highly heterogeneous sites. All remaining sites which were identified as falling within a particular color-coded spectral class were then pooled and descriptive statistics generated for each ground truth parameter illustrated. The mean measurement is listed first, followed by the range of ground truth measurements for that class. The last statistic, N, equals the number of sampling sites used to calculate the mean. The statistics generated were then used to characterize the spectral class in terms of water quality.

Spectral classes from the thematic maps can be pooled into three general categories. The first category, PRODUCTIVITY, is represented in the thematic maps by dark blue, light blue and green. Areas color-coded dark blue can be described as oligotrophic, as evidenced by the SDT, low CHLA and deep 1% Lev. measurements. The light-blue-coded class can be considered to be a borderline between oligotrophic and mesotrophic, based upon supportive contact-sensed measurements. The third productivity-related class does not appear in all thematic maps. It is a relatively small class with little supportive ground truth. However, based upon SDT and 1% Lev. measurements and the geographic locations of the class, it appears to be more productive than either the dark blue or light blue class.

The second category has been designated MIXED EFFECTS and is represented by a gray class occurring in some thematic maps and a black class appearing in all the images. It is suspected that the gray class does not relate to water quality but is evidence of residual sixth-line banding which was not successfully removed. The black coded class has no corresponding water quality measurements. However, an examination of the lake in terms of where the black pixels occur, coupled with a knowledge of the lake, strongly suggests that the class consists of the intermingled effects of nearshore areas. These areas typically exhibit effects caused by visible lake bottom, submerged plant cover and shadows.

The color coded areas of yellow, brown and orange represent the third category, TURBIDITY. The yellow coded areas are characterized by a lowering SDT and increased NT. The brown coded areas can be readily identified as river water which has entered the lake and is still in the process of dropping its suspended sediments. The most turbid class of water found in Lake Mead is represented in the orange coded areas. These areas consist of the heavily sediment laden plume of the Colorado River where it enters Pierce Basin and a small portion of the Muddy River branch of the Overton Arm. These areas signal the influx of a substantial quantity of inorganic materials into the lake which effectively reduce the zone in which photosynthesis occurs, thereby limiting productivity.

B. THE LAKE TAHOE STUDY

In order to investigate the possibilities of a comprehensive data base system which took into account the contributing factors as well as the effects of lake eutrophication, the IPL chose to construct an IBIS data base for the Lake Tahoe Basin. Due to increasing environmental pressures and rampant land development, this region has been the subject of intensive study by the U.S. Department of Agriculture and the Tahoe Regional Planning Agency (TRPA). In 1977 TRPA made available a volume entitled, Water Quality Problems and Management Program, which detailed present knowledge of the Lake Tahoe Basin Watershed. Using data compiled in this volume as a source, analysts at IPL proceeded to integrate these data with Landsat and digital t rain data produced by the Defense Mapping Agency (DMA). The IBIS data base method allowed cross correlation of Landsat imagery and topographic data with USDA environmental data relating to such parameters as drainage basin acreage, surface runoff and terrain configuration. Data base construction led to the production of integrated images created by "overlaying" conventional maps and digital maps on Landsat. Preliminary work was also performed which attempted to depict the terrain configuration for each drainage basin based on DMA topographic data.

1. The Lake Tahoe Basin

The Tahoe Basin occupies over 500 square miles (1295 km^2) situated in a graben straddling the boundary between California and Nevada. Figure II-16 reproduces a map of the Lake Tahoe Basin and its environs. Lake Tahoe contains 126 million acre feet (155.4 km^3) of water in a 190-square-mile (492.1 km^2) surface area receiving inflow from 63 tributaries with only one outlet at the Truckee River. The Tahoe Basin has traditionally attracted recreational activity due to its clear deep water and pine-forested shorelands coupled with its proximity to major metropolitan areas in northern California. Since the 1950's the basin has experienced escalating demands for land development at the expense of the natural watershed. Discharge of sediments to the lake has greatly increased due to accelerated human interference, and alterations to the natural drainage patterns are evident in some areas. According to the Water Quality Plan published by the Tahoe Regional Planning Agency, although most development has occurred in areas of mixed-pine and pine-fir vegetation types, the greatest impact of urbanization upon vegetation and hence the watershed environment has occurred on several marshes and meadows within the Basin.

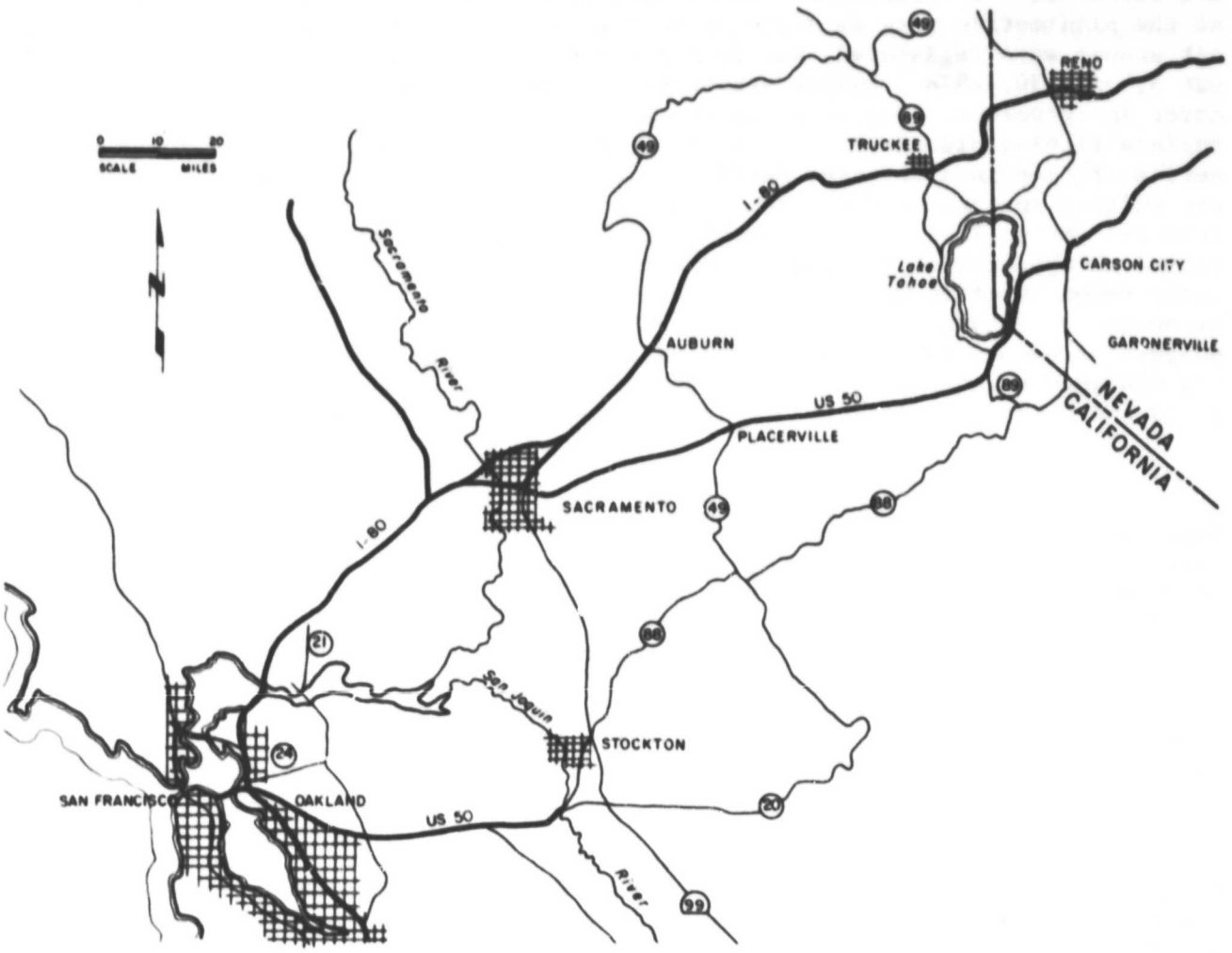

Figure II-16. Map showing the geographic location of Lake Tahoe

The problems which the Lake Tahoe basin is presently confronting are certainly not unique to this area. The consequences of man's alteration of the natural environment are symptomatic of the pressures which can be brought to bear by the activities of an increasingly mobile populace seeking new recreation and living space.

2. Construction of the Lake Tahoe Basin Data Base

a. Landsat Data. IPL chose to construct the Lake Tahoe Basin data base from data sources which were readily available and familiar to its analysts. Landsat imagery, which has formed the basis for the majority of IPL's water quality assessment efforts, was chosen as the planimetric base to which all other data would be registered. A subsection of a Landsat 2 scene from August 27, 1976,

was extracted and transformed into a Lambert conformal conic projection to serve as the planimetric base as depicted in Fig. II-17. To this scene two other Landsat scenes were registered, one from Landsat 2, July 21, 1978, and one from Landsat 3, July 30, 1978. Summer scenes were chosen to reduce the chance of snow cover interference. Registration was achieved through the aid of a piecewise surface fitting algorithm. The surface transformation was defined through a series of tiepoints selected during the use of an interactive spatial pattern recognition routine (PICREG) in which the analyst selected common geographical features (Refs. II-11, 12). (Details of the PICREG program can be found in the technical application section on Lake Mead). After all images were registered, color reconstruction was also performed on each scene using a color enhancement technique which approximates a gaussian distribution using the principal components of each Landsat image separately (Ref. II-13). This type of enhancement was designed to attempt even distribution for all regions within an image to avoid the saturating effects of conventional linear enhancements.

b. DMA Digital Topographic Data. Terrain data produced by the Defense Mapping Agency (DMA) was next integrated into the data base. Digital terrain tapes were acquired for the Lake Tahoe region from the National Cartographic Information Center. The tapes were prepared from U.S. Geological Survey 1:250,000 scale topographic quadrangle map series. Map contour lines falling within a 1-degree block of latitude and longitude were digitized and a matrix of elevations generated with one elevation for every 0.01 inch on the map (200 feet/60.96 meters on the ground). The terrain data were reformatted for use in IPL's VICAR operating system. In reformatting, halfword integer elevation values were converted to single byte integers and scaled to the terrain variation within the area (Ref. II-14).

After reformatting, the terrain data sets were rotated 90 degrees counterclockwise to orient the north-to-south values vertically in the sample direction, placing north at the top of the images. This compensated for data format as produced by the National Cartographic Institute. For the Lake Tahoe Basin, four separate terrain quadrangles were required to construct a complete image comprising the Tahoe Basin area. This necessitated the mosaicking of the four quadrangles before final registration to the data base. Figure II-18 reproduces the four digital terrain quadrangles required to completely encompass the Lake Tahoe Basin and the final mosaic image. The relief-like effects portrayed in the images were produced by digitally shading between contour intervals.

Registration of the terrain mosaic with the Landsat planimetric base was achieved through the application of a resampling algorithm which applied a two-dimensional correction grid derived from selected control points (Refs. II-11, 12). Tiepoint selection was achieved interactively (PICREG) relating surface features through comparison of the Landsat planimetric base image and a relief-like version of the terrain image.

c. Graphical Data. Graphical data for the Lake Tahoe Basin data base were acquired from conventional maps produced for the Tahoe Regional Planning Agency by the U.S. Department of Agriculture Forest Service. These maps presented hydrologic and climatologic data pertaining to the Lake Tahoe Basin. Before integration into the IBIS data base, graphical data must be transformed into image format.

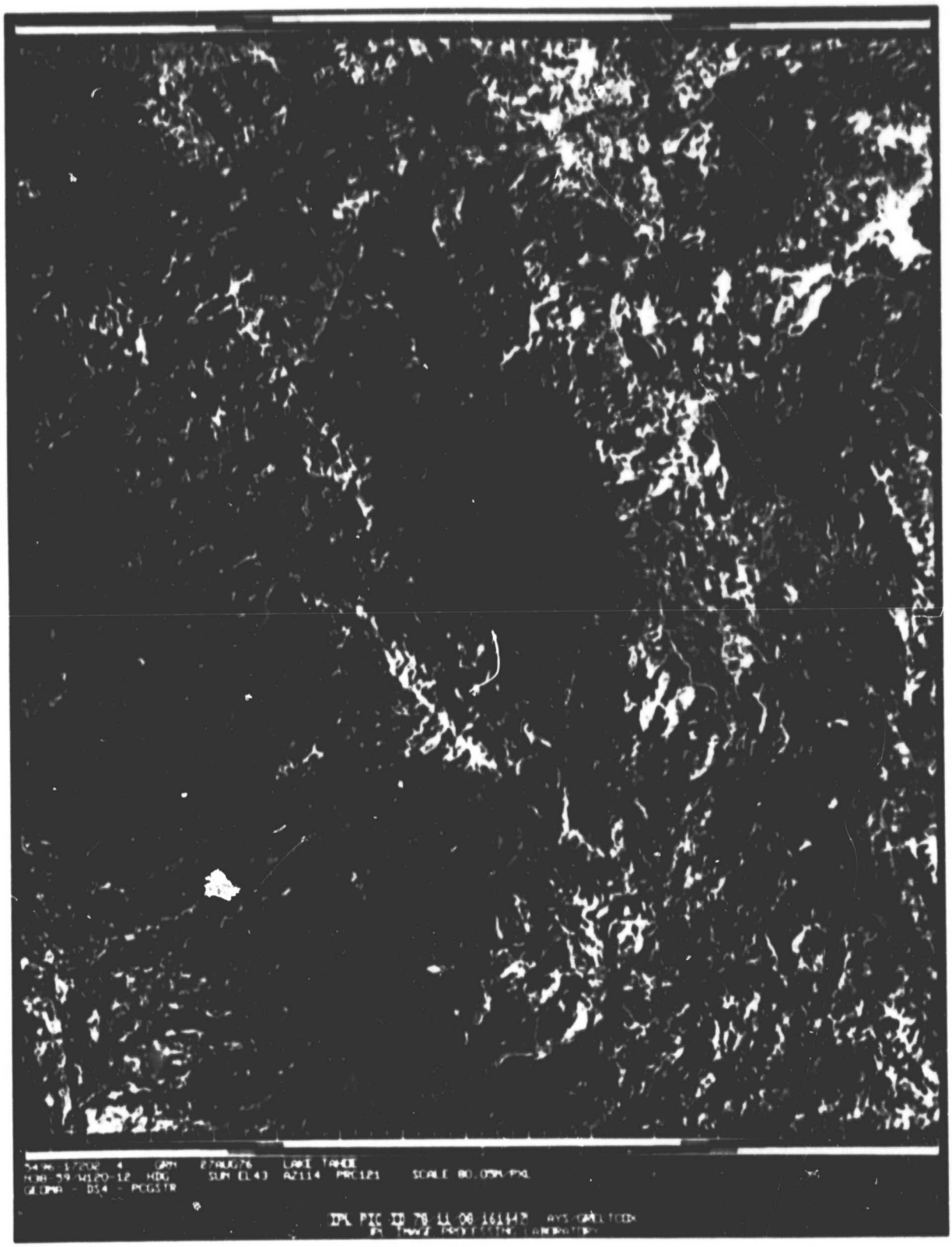

Figure II-17. False color composite image of Lake Tahoe from Landsat frame 5496-17202, August 27, 1976.

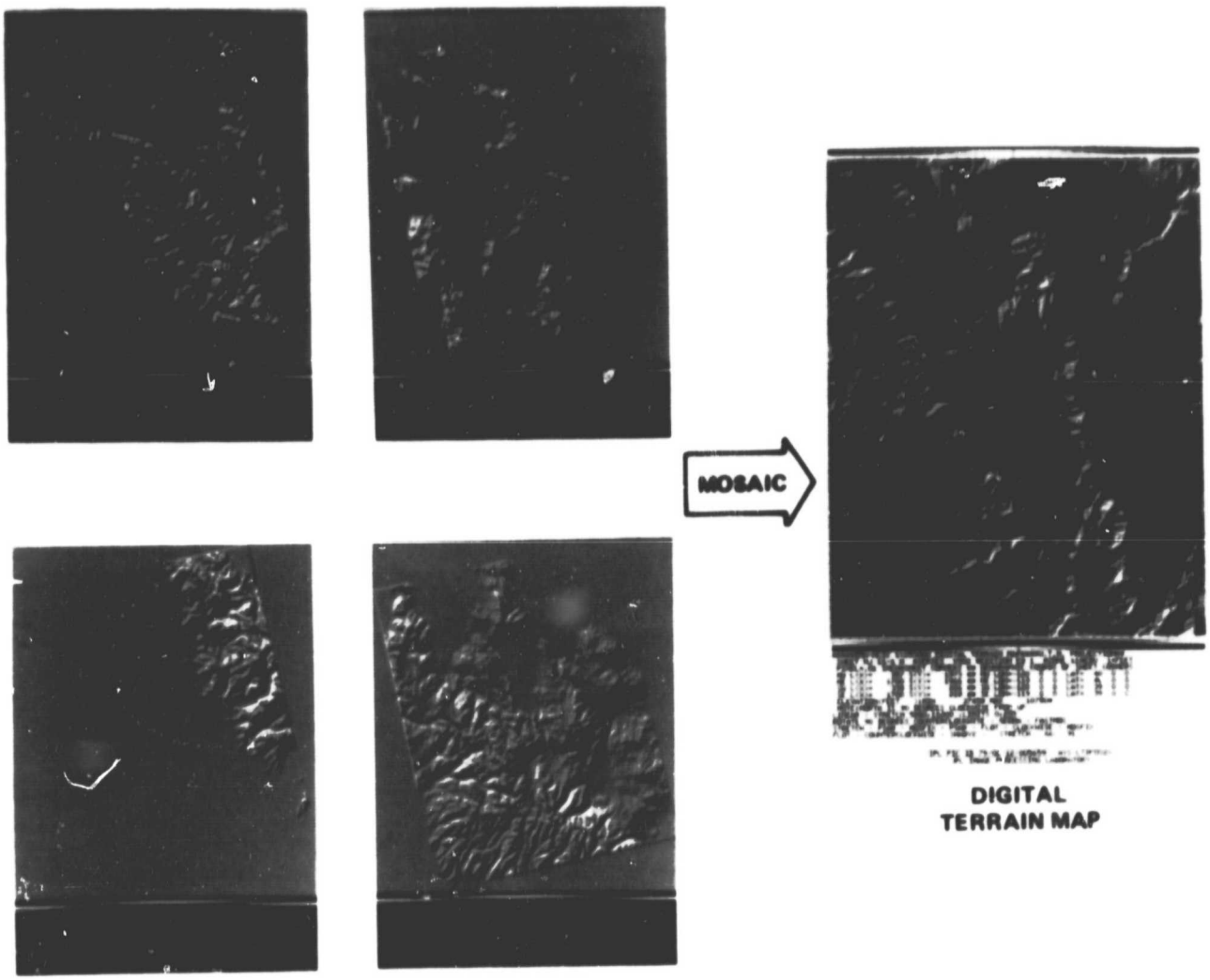

Figure II-18. Four digital terrain quadrangles were mosaicked to produce a terrain map which contained the Lake Tahoe Basin

This was achieved by first digitizing the graphical data onto magnetic tape. Next, a least squares affine transformation was applied which created a general surface fit. Corresponding line and sample coordinates, in terms of the Landsat planimetric base, were calculated from latitude and longitude coordinates taken from the graphical map data for geographical features common to both data sets. Registration was once again achieved through the interactive (PICREG) selection of control points which linked geographical features from the graphical data file and the planimetric data base. In Fig. II-19, the transformation of a conventional map to a graphical data plane is illustrated. Step A represents the digitizing procedure in which the drainage basin map is converted to graphical image format.

d. Georeference Plane. A georeference image plane which provided an interface between all dataplanes for the Lake Tahoe Basin data base was created from the drainage basin map provided by the Tahoe Regional Planning Agency. As seen in Fig. II-19, this map identified the location and outlined the boundaries for the 63 drainage basins identified as providing watershed inflow to Lake Tahoe. The map was digitized and subsequently registered to the planimetric base. Each separate drainage basin was then assigned a unique data number for all picture elements comprising that basin. Step B, in Fig. II-19, represents the process of encoding each drainage basin, a procedure known as "painting." The final product, the georeference plane, is reproduced at the right in Fig. II-19. All tabular data corresponding to the drainage basins were then entered into the data base by referencing the unique number assigned to each basin. The georeference plane is an integral element within the data base, for it is through this plane that all tabular and most image data are interfaced. Figure II-20 is a conceptualization of the georeference plane for the Tahoe region. The magnified portion of the figure indicates the coded picture elements which comprise each basin. In this figure a most basic application is illustrated, in which the picture elements for each basin are summed and transferred to an interface file in which the sums are stored according to basin code number. This interface file can be accessed at a later date to produce statistical output such as acreage estimates for each drainage basin.

Figure II-21 illustrates the Lake Tahoe Basin data base as completed, forming a series of registered data planes. Once all data planes are in registration and at least one georeference plane is registered, the data base can be used to develop additional products through manipulation and integration of the planes.

3. Application of the IBIS Concept to the Lake Tahoe Basin

a. Integrated Images. Manipulation and integration of data planes comprising the data base is achieved through the implementation of VICAR and IBIS programs developed at IPL. Standard products output from IBIS integration consist of overlay images. Figure II-22 reproduces a precipitation map produced by the Tahoe Regional Planning Agency. Using IBIS resources, this map was electronically digitized and registered to the planimetric base image. By adding the digitized map to the color version of the Landsat scene for August 27, 1976, an integrated image is created. As seen in Fig. II-23, this type of product can represent more clearly the relationship between mean precipitation and the Tahoe environment as depicted

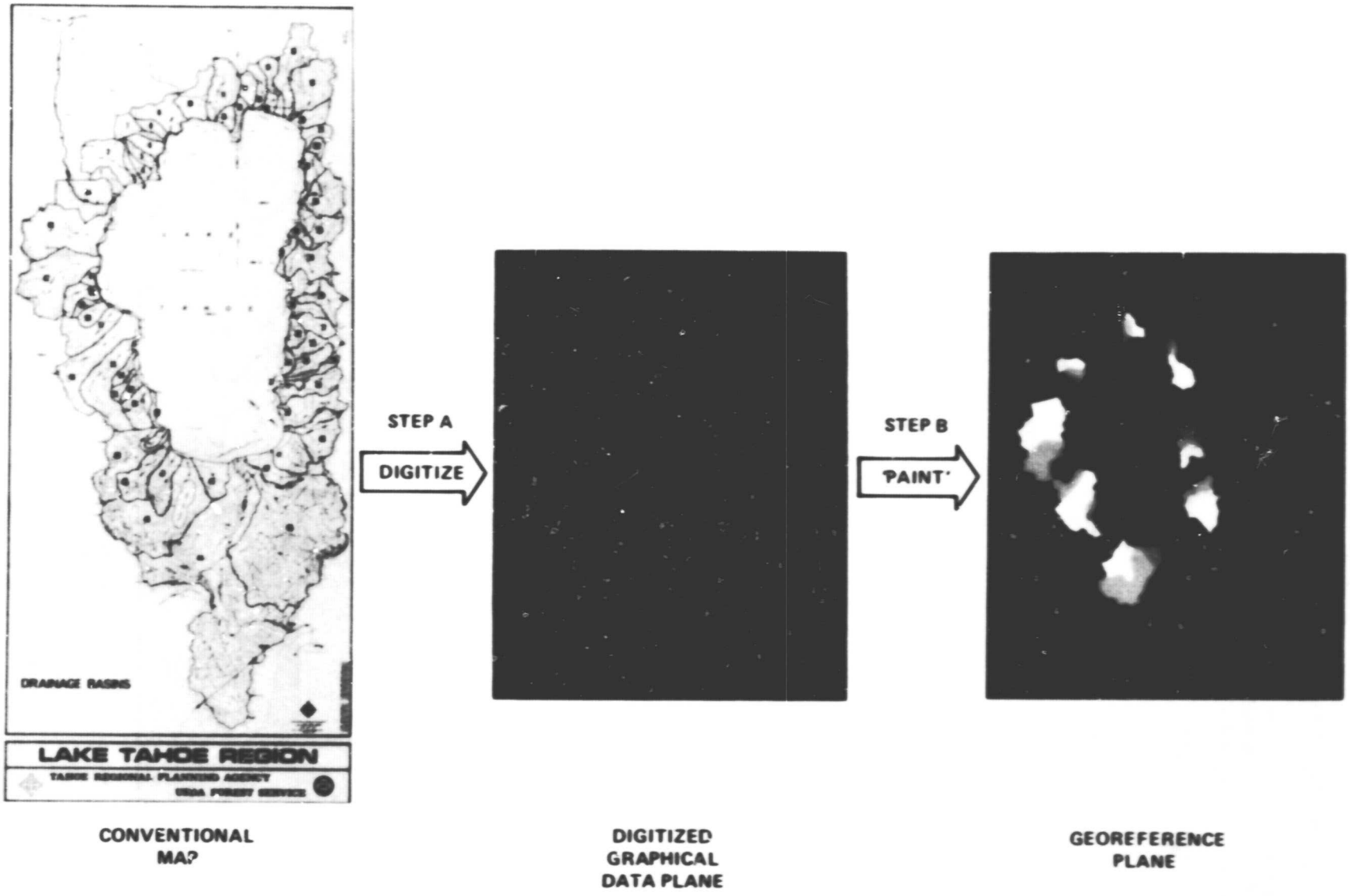

Figure II-19. Transformation of a conventional map to a graphical data plane and georeference plane

INTERFACE FILE

DRAINAGE BASIN NUMBER	NO. OF IMAGE ELEMENTS	IMAGE GRAY VALUE
1	1165	1
2	3027	2
3	403	3
4	1110	4
5	763	5
6	4639	6
7	5102	7
8	1708	8

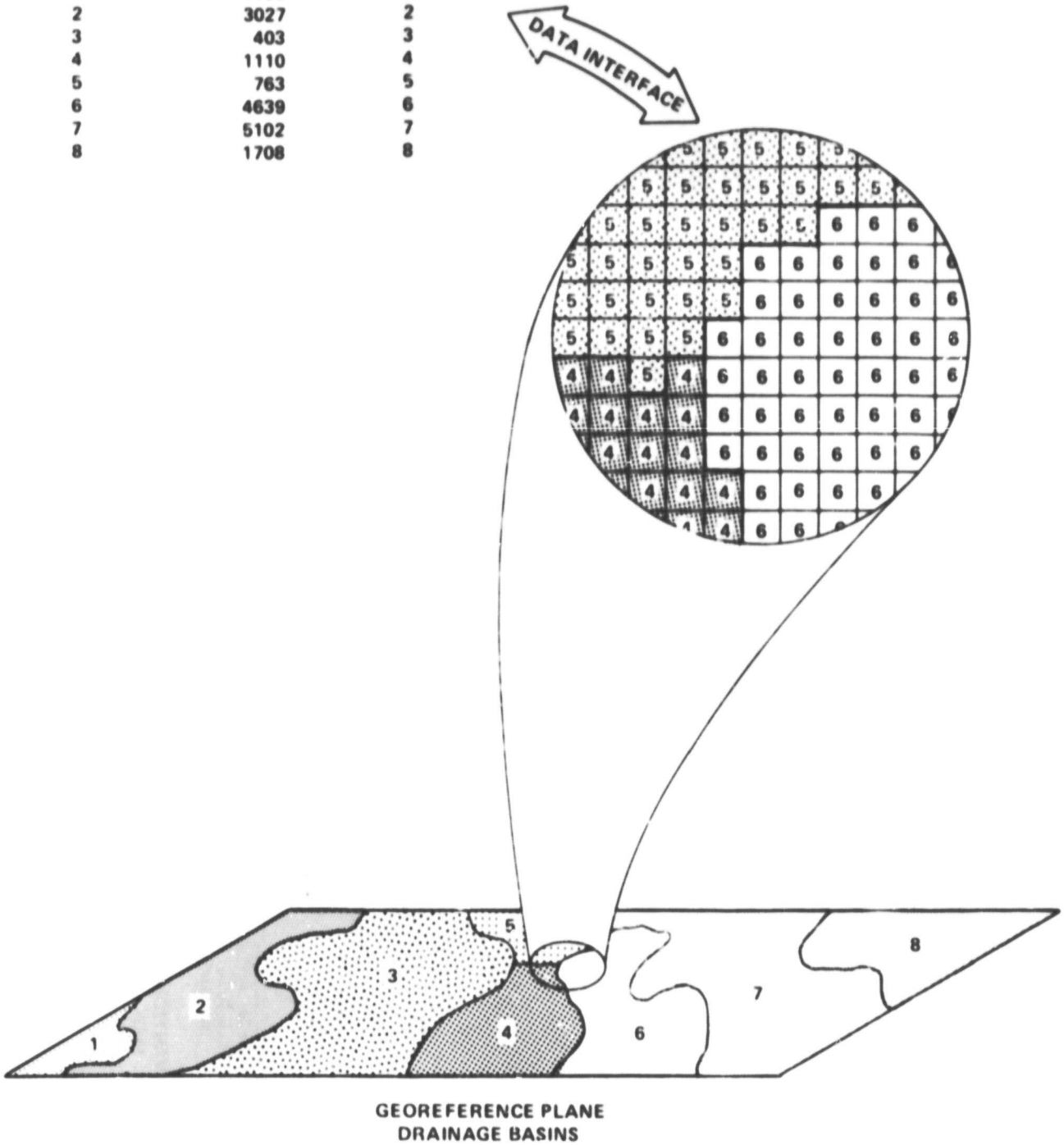

Figure II-20. Conceptual drawing of the georeference plane and tabular interface file

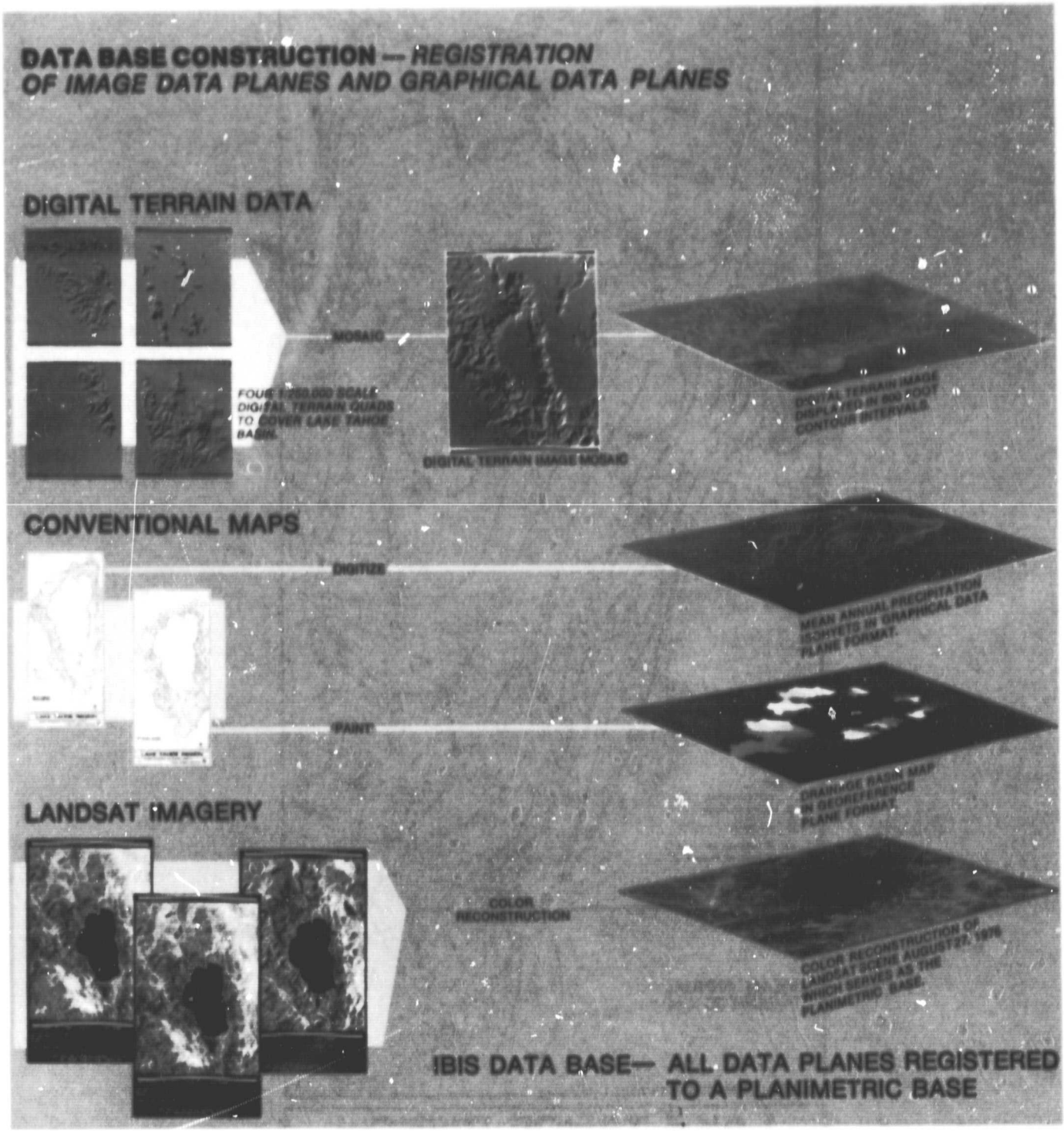

Figure II-21. A conceptual diagram illustrating the development of the Lake Tahoe data base from digital terrain data, conventional data and Landsat imagery

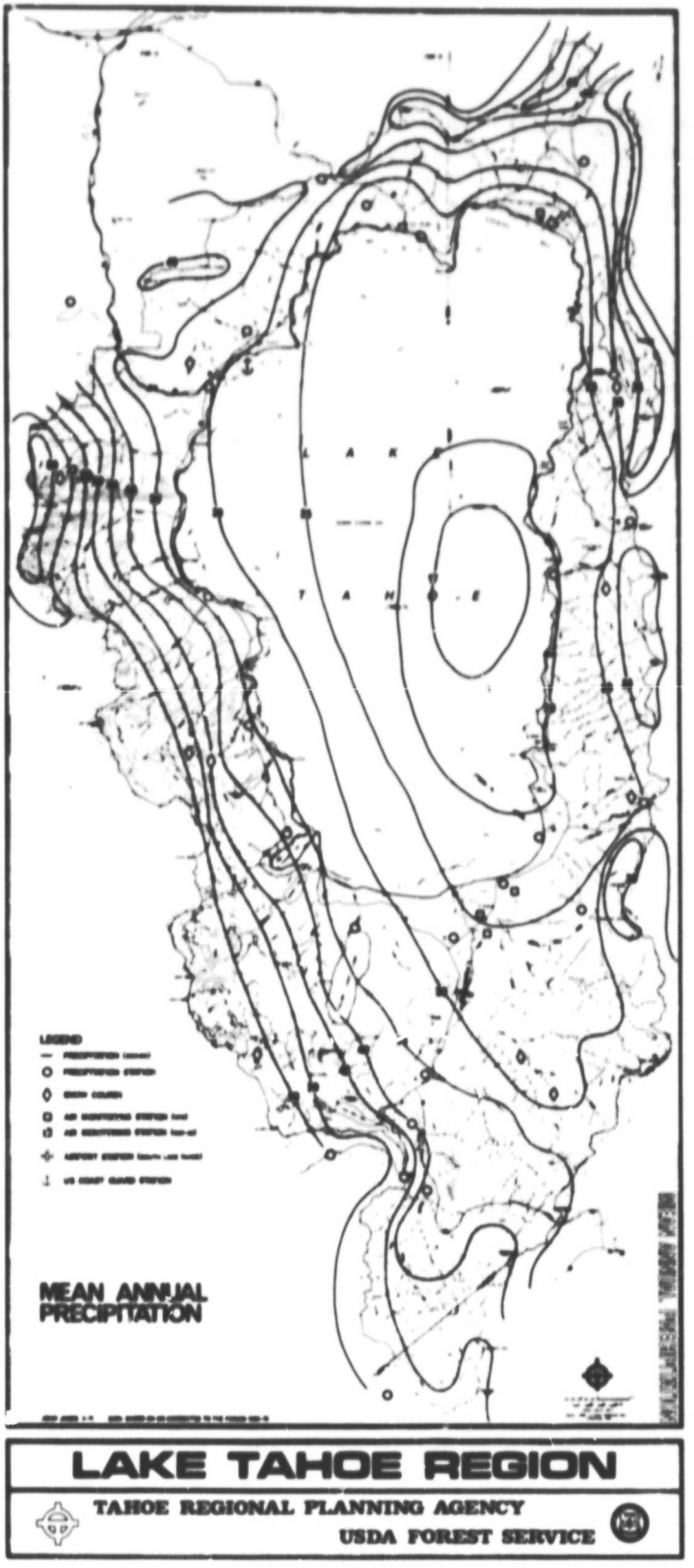

Figure II-22. Conventional map showing mean annual precipitation contours over the Lake Tahoe Basin

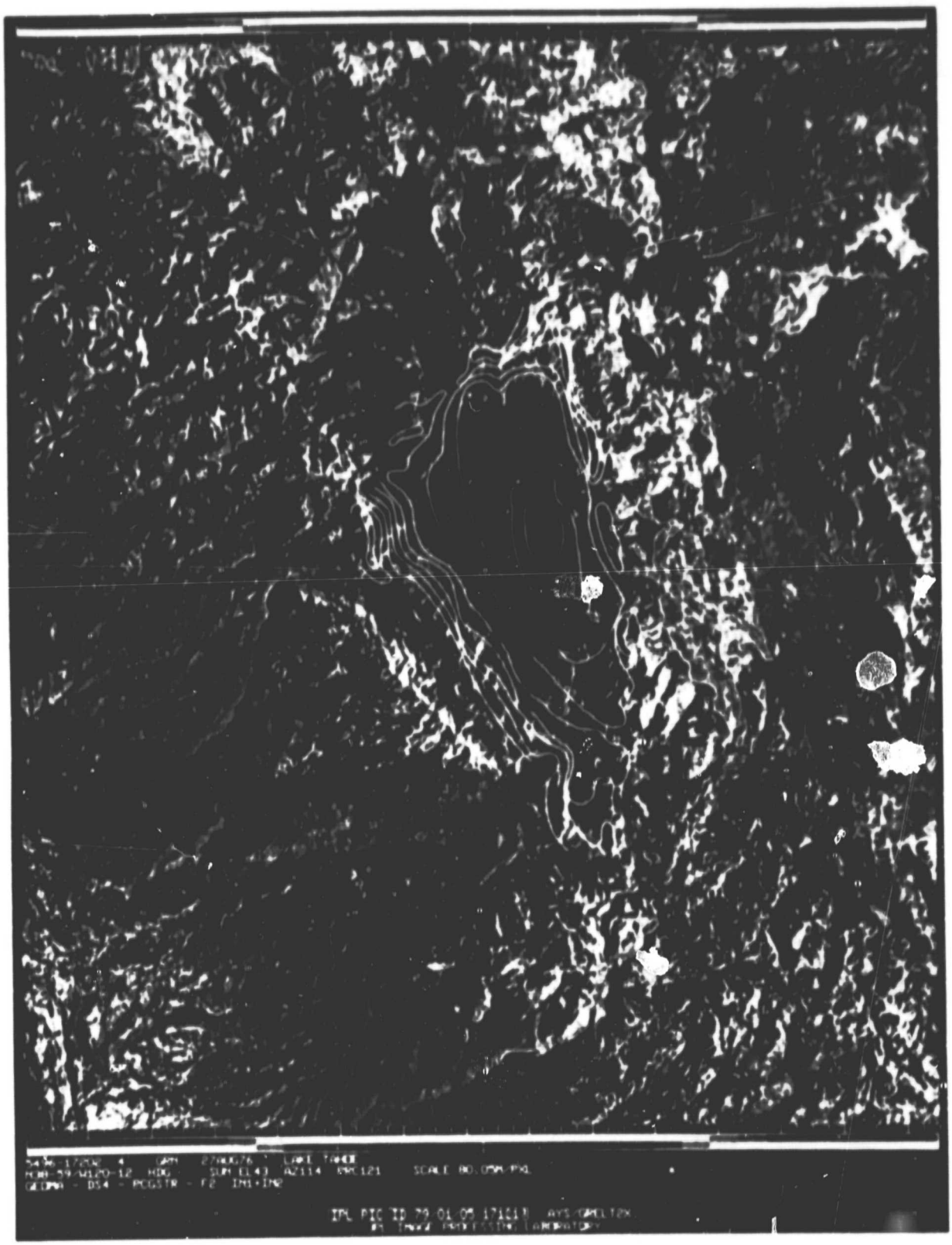

Figure II-23. Image showing the registration of precipitation data from a conventional map with Landsat imagery.

by Landsat. In future applications the mean annual precipitation isohyets could be converted to a continuous surface image. This process is similar to the conversion of USGS topographic maps to terrain relief images as depicted previously in Fig. II-19. The precipitation surface image could then be cross-tabulated with individual drainage basins, using the georeference plane, to produce estimates of runoff coefficients.

Similarily, another integrated image was produced by combining digital terrain data with Landsat imagery. The Defense Mapping Agency (DMA) mosaic was transformed to display 600-ft. contour intervals as illustrated in Fig. II-24. The digital terrain image was then integrated with the August 27, 1976, Landsat scene to produce the image shown in Fig. II-25. Integrated imagery of this type can be used to facilitate interpretation of terrain elevation in terms of a Landsat scene. DMA data were also used to calculate estimates of azimuth and slope for each drainage basin. Details of this procedure are described in the following section.

The georeference plane was integrated with each Landsat scene to depict the boundaries of each drainage basin mapped by the TRPA. Figure II-26 shows each of three Landsat scenes contained in the data base as it appears integrated with the drainage basin map. While imagery of this type is useful for cursory interpretation and display purposes, the ability to extract additional information from data plane manipulation is the key to the application of the IBIS concept.

b. Data Plane Manipulation. Manipulation of the georeference plane, in this case the drainage basin map, provides the opportunity for the analyst to more closely examine the unique characteristics of each drainage basin as revealed by the Landsat sensors. By referencing each uniquely coded drainage basin in the georeference plane separately, the analyst can extract and display any or all basins as individual images. This application is illustrated in Fig. II-27, in which the georeference plane was used to extract the Trout Creek Basin and the Upper Truckee Basin, which are displayed in color as separate images. If a multispectral classification is performed on the data, information can be similarly reproduced in a basin-by-basin format. Tabulations of land cover type can then be expressed for all basins collectively or each basin individually.

The georeference plane can also be interfaced with tabular files to produce a statistical output such as that reproduced in Table II-7. In this table, the area of each uniquely coded drainage basin comprising the georeference plane was determined by summing picture elements and converting the sum to total acres, producing a Landsat acreage estimate. These data were then cross-referenced to tabular files which contain acreage estimates produced by the TRPA. All estimates were then ranked and output in the form of a computer listing which provided a comparison of Landsat vs conventionally acquired acreage estimates.

The inclusion of DMA digital terrain data provided an opportunity to attempt integration of an important element of ancillary, ground-based information with remote sensing data. The DMA developed these data by interpolating existing USGS 1:250,000 scale topographic maps to produce ultrafine mesh digitized latitude, longitude and elevation contour data. As described earlier, four DMA quadrangles were digitally mosaicked to form a unified, continuous surface elevation image.

Figure II-24. Digital elevation contours produced by processing digital terrain data

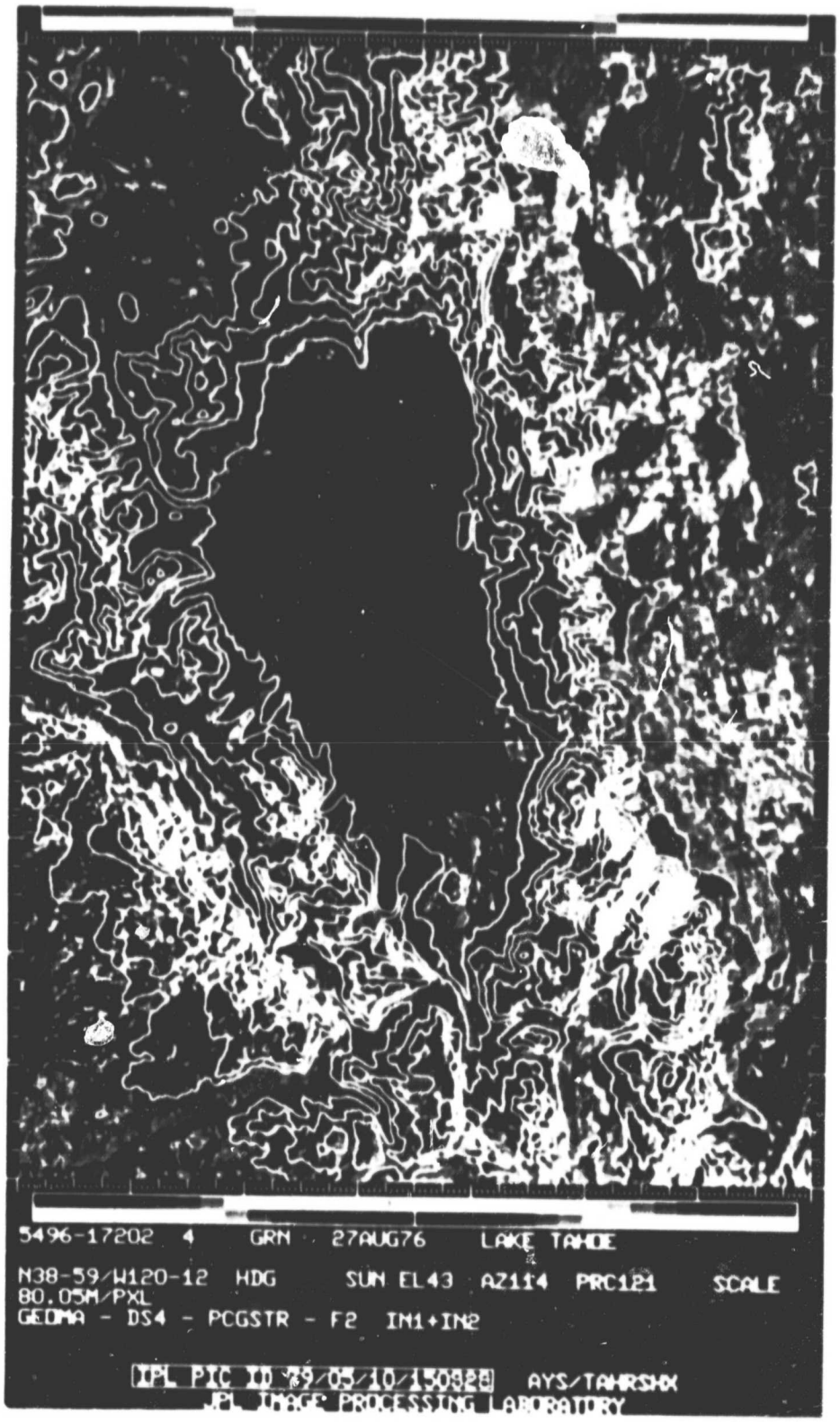

Figure II-25. Image showing the registration of elevation contours with Landsat imagery.

Figure II-26. Integration of data planes produces overlay images creating new visual tools which combine data from various sources and for different dates.

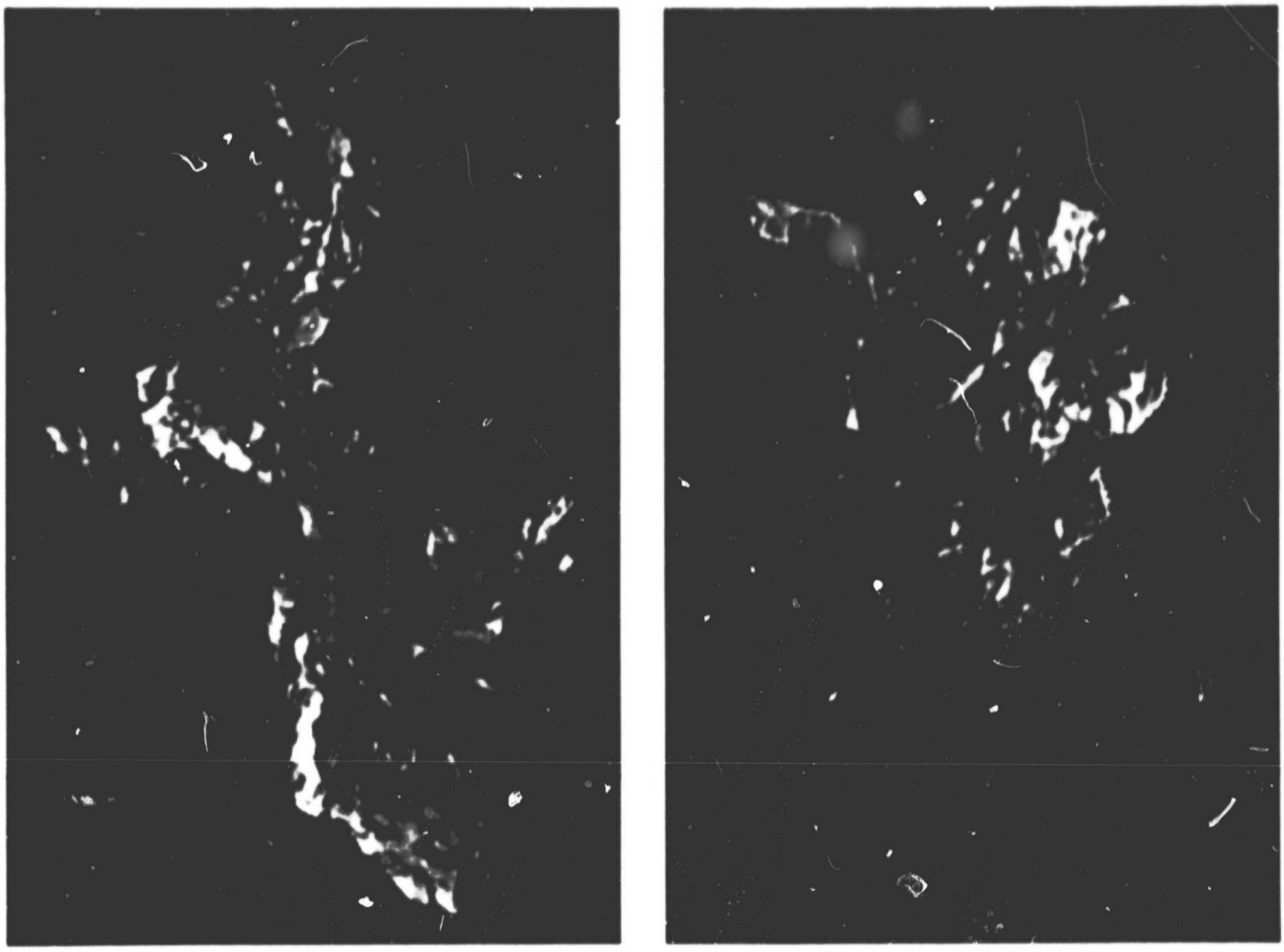

Figure II-27. Interfacing the georeference plane with Landsat imagery allows access to each basin separately and the generation of tabular reports:
(a) = Trout Creek Basin
(b) = Upper Truckee Basin

This image was then geometrically registered with the Landsat imagery to permit future cross-tabulations of elevation information with other information data planes within the data base. In an effort to quantify and process the digital terrain data, the elevation image was processed with VICAR software to produce a component representing slope magnitude. To develop this component it was necessary to compute the vector crossproduct between the horizontal (east-west) image elements and the vertical (north-south) picture elements. This vector product then provided an estimate of the slope between adjacent east-west and north-south elevation image elements. The outputs of this process were then coded to reflect slopes between 0° (no slope) and 90° (vertical slope). The angles in degrees were further coded for image output by rescaling 0° to be equal to 0 digital number (DN) or black and 90° to be equal to 255 DN or white. Figure II-28 illustrates an application of these concepts. First, the slope magnitude was cross-tabulated with

Table II-7. Lake Tahoe drainage basins acreage tabulation

BASIN NO.	BASIN NAME	TRPA ACREAGE ESTIMATE 1971	TRPA ACREAGE ESTIMATE 1977	AVERAGE ACREAGE ESTIMATE	LANDSAT ACREAGE
37	SOUTH ZEPHYR CREEK	320	407	364	223
5C	ALTISS STATE PARK	467	486	504	458
21	EAGLE ROCK	500	459	515	302
34	SKYLAND	480	481	481	508
19	BURNT CEDAR CREEK	530	521	531	535
52	UNNAMED	870	501	434	460
46	[illegible]WOOD CREEK	500	519	568	551
4	LAKE FOREST CREEK	403	642	561	508
28	BLISS CREEK	560	618	604	608
3	[illegible]ASTON CREEK	660	619	614	616
33	PARADISE FLAT	623	637	704	609
35	LONELY GULCH CREEK	790	616	703	612
27	DEADMAN POINT	670	954	815	654
12	KINGS BEACH	870	791	815	752
54	SIERRA CREEK	700	771	786	758
8	CARNELIAN BAY CREEK	550	706	838	789
13	EAST STATELINE POINT	640	659	838	829
36	ZEPHYR CREEK	870	695	785	863
1	TAHOE STATE PARK	800	946	901	854
21	TUNNEL CREEK	1010	1040	1027	961
30	NORTH LOGAN HOUSE CREEK	1100	1053	1077	1013
32	CAVE ROCK	1040	1018	1029	1080
58	QUAIL LAKE CREEK	1100	1035	1061	1088
14	FIRST CREEK	900	1037	984	966
3	[illegible]LAR CREEK	1100	1028	1063	1100
15	SECOND CREEK	1120	1135	1108	1168
9	CEDAR FLAT	1200	1164	1182	1178
23	SAND HARBOR	1250	1184	1252	1206
50	MILL CREEK	1200	1307	1254	1249
31	LOGAN HOUSE CREEK	1450	1542	1496	1349
7	WOOD CREEK	1400	1467	1418	1862
7	WATSON	1540	1547	1546	1642
33	LINCOLN CREEK	1610	1692	1605	1684
60	MADDEN CREEK	1530	1523	1526	1588
38	NORTH ZEPHYR CREEK	1570	1693	1601	1461
42	BIJOU CREEK	1670	1737	1671	1723
41	RUBICON CREEK	2010	2128	2084	2123
45	CAMP RICHARDSON	2140	2058	2094	2207
38	MCFAUL CREEK	2150	2022	2036	2364
25	SECRET HARBOR CREEK	2110	2500	2048	2091
6	CARNELIAN CANYON	2400	2580	2783	2064
47	TALLAC CREEK	2830	2781	2805	2651
10	TAHOE VISTA	2670	2624	2647	2771
44	BIJOU PARK	2680	2602	2671	2859
48	CASCADE CREEK	3100	3075	3088	[illegible]
24	MARLETTE CREEK	3100	3179	3140	3069
20	SLAUGHTER HOUSE	3180	4032	3906	3119
57	MCKINNEY CREEK	3260	3235	3248	3267
39	BURKE CREEK	3380	3317	3304	3308
29	GLENBROOK CREEK	3200	3309	3395	3347
11	GRIFF CREEK	3400	3373	3282	3376
3	BURTON CREEK	3190	3438	3564	3590
40	EDGEWOOD CREEK	3700	3476	3637	3683
18	THIRD CREEK	3780	4065	3872	4101
18	INCLINE CREEK	3720	4083	4177	4172
69	EAGLE CREEK	4250	5406	5073	4189
55	MEEKS	5640	5247	5585	5226
50	GENERAL CREEK	5930	5525	5658	5422
62	BLACKWOOD CREEK	5790	7338	9464	7240
63	WARD CREEK	7600	8035	8138	7958
60	TAYLOR CREEK	8240	11730	12003	11318
43	TROUT CREEK	15400	15549	15475	16728
44	UPPER TRUCKEE RIVER	36330	35475	35903	36045
		200980	197973	199507	200095

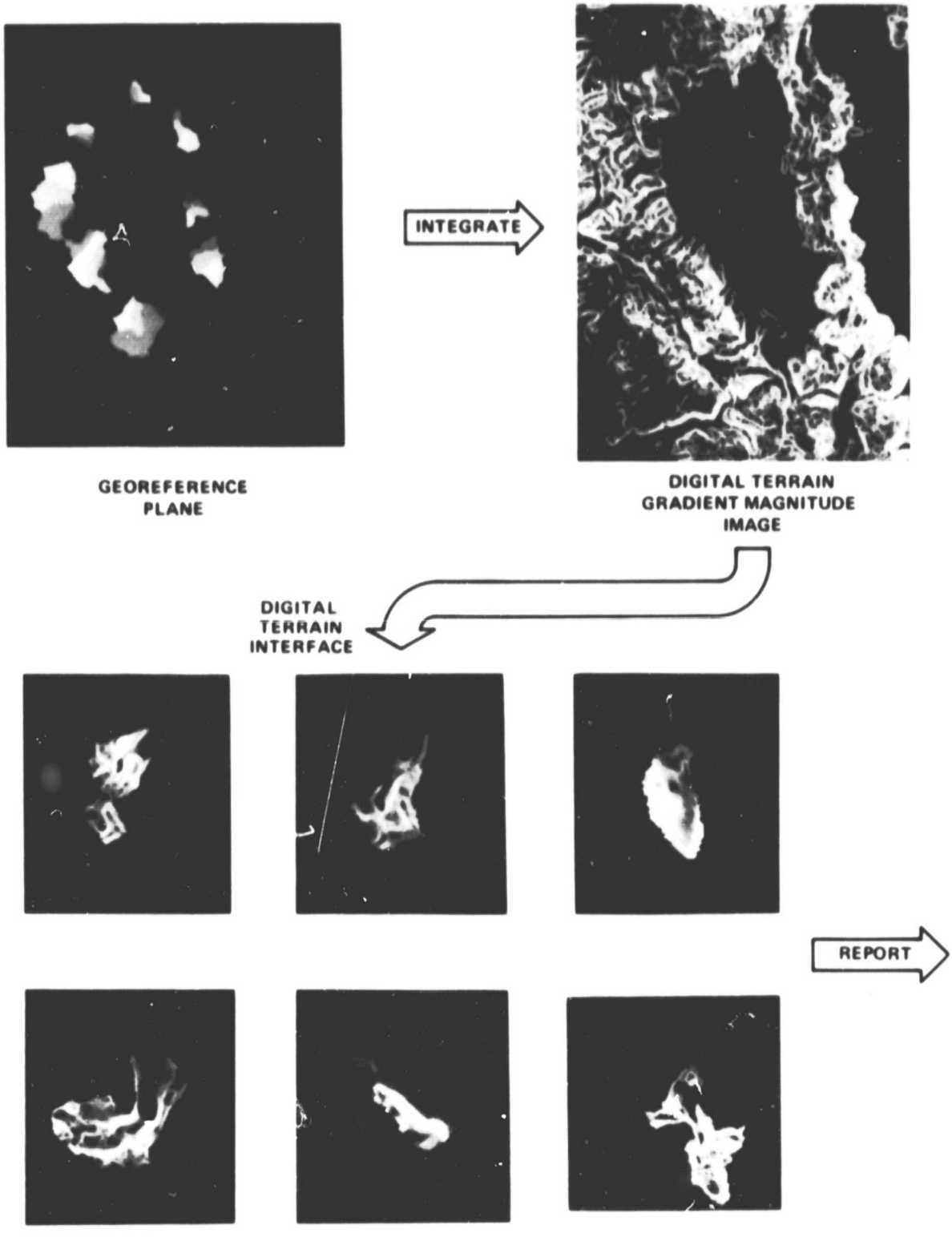

Figure II-28. Using the drainage basin map as the georeference image, the slope and gradient magnitude for each drainage basin may be displayed and tabular information produced

the drainage basin georeference plane. Each slope magnitude image element associated with the individual drainage basins was then extracted. Further cross-tabulation permitted the computation of a mean slope estimate by averaging the slope magnitude image elements for each drainage basin.

Table II-8 reproduces an IBIS table generated from the digital terrain imagery which lists azimuth and slope for each drainage basin. Eventually, it is hoped to develop a model for drainage basin terrain based on digital imagery. At present, however, difficulties have been encountered with discrepancies between individual DMA map quadrangles, especially along map edges, which have precluded the development of accurate models. Other types of digital terrain imagery are being investigated with the hope of integrating more reliable data into the Lake Tahoe Basin data base.

Table II-8. Lake Tahoe drainage basins azimuth (aspect) and slope magnitude statistical summary report.

				MODAL ASPECT				MODAL SLOPE		
BASIN NUMBER	DRAINAGE BASIN NAME	AREA (ACRES)	AVERAGE ASPECT	AZIMUTH (DEG.)	ACREAGE	PCT. TOTAL	AVERAGE SLOPE	SLOPE (DEG.)	ACREAGE	PCT. TOTAL
1	TAHOE STATE PARK	952	26.8	90.4	185.3	19.47	5.5	4.9	158.7	16.66
2	BURTON CREEK	3390	33.9	90.4	357.9	10.56	9.1	4.9	310.3	9.15
3	BARTON CREEK	616	38.5	180.7	196.3	31.88	6.6	3.2	145.7	23.65
4	LAKE FOREST CREEK	592	34.0	135.5	229.6	38.77	9.1	3.2	194.8	32.89
5	DOLLAR CREEK	1100	27.7	90.4	166.3	15.11	7.3	4.9	120.3	10.94
6	CEDAR FLATS	1178	24.5	90.4	422.8	35.89	7.7	6.7	202.7	17.20
7	WATSON	1542	28.6	90.4	186.8	12.11	9.1	10.6	107.7	6.98
8	CARNELIAN BAY CREEK	789	34.1	90.4	161.5	20.48	7.7	6.7	109.3	13.86
9	CARNELIAN CANYON	2644	30.3	135.5	278.7	10.54	8.4	3.2	216.9	8.20
10	TAHOE VISTA	2655	36.5	180.7	446.5	16.82	8.3	2.1	277.1	10.44
11	GRIFF CREEK	3376	47.0	180.7	650.8	19.28	10.4	4.2	253.3	7.50
12	KINGS BEACH	725	55.8	225.9	185.3	25.55	9.0	4.9	110.8	15.28
13	EAST STATELINE POINT	825	33.7	90.4	80.8	9.79	15.2	10.6	61.8	7.49
14	FIRST CREEK	1094	35.3	135.5	79.2	7.24	16.0	18.0	52.3	4.78
15	SECOND CREEK	1116	40.4	180.7	155.2	13.90	14.2	16.9	58.6	5.25
16	BURNT CEDAR CREEK	535	41.6	180.7	156.8	24.29	6.5	6.7	87.1	16.27
17	WOOD CREEK	1384	38.2	180.7	142.5	10.30	13.7	4.9	114.0	8.24
18	THIRD CREEK	4101	38.6	180.7	471.9	11.51	11.5	4.9	259.7	6.33
19	INCLINE CREEK	4172	50.1	225.9	427.5	10.25	12.2	10.6	220.1	5.28
20	MILL CREEK	1344	55.3	271.1	123.5	9.19	16.6	3.2	60.5	4.55
21	TUNNEL CREEK	983	56.0	271.1	82.3	8.37	16.2	6.7	44.3	4.51
22	UNNAMED	540	67.4	271.1	76.0	14.08	15.5	14.8	36.4	6.74
23	SAND HARBOR	1336	60.4	271.1	109.3	8.18	20.7	6.7	69.7	5.21
24	MARLETTE CREEK	3069	48.7	0.0	494.0	16.10	11.4	0.0	438.6	14.29
25	SECRET HARBOR CREEK	2608	55.5	271.1	275.5	10.56	13.1	10.6	115.6	4.43
26	BLISS CREEK	608	69.4	316.2	109.3	17.97	9.9	7.8	45.9	7.55
27	DEADMAN POINT	654	56.5	0.0	63.3	9.69	13.2	16.9	55.4	8.47
28	SLAUGHTER HOUSE	3118	43.5	0.0	552.6	17.72	9.7	0.0	443.4	14.22
29	GLENBROOK CREEK	3347	55.1	0.0	296.1	8.85	11.5	4.9	166.3	4.97
30	NORTH LOGAN HOUSE CREEK	1013	61.4	0.0	80.8	7.97	11.9	2.1	63.3	6.25
31	LOGAN HOUSE CREEK	1349	60.9	271.1	171.0	12.68	11.1	10.6	90.3	6.69
32	CAVE ROCK	1029	62.8	271.1	137.8	13.38	15.7	16.9	47.5	4.62
33	LINCOLN CREEK	1568	54.1	0.0	148.8	9.49	9.2	7.8	109.3	6.97
34	SKYLAND	508	60.9	271.1	99.8	19.63	9.1	6.7	55.4	10.90
35	NORTH ZEPHYR CREEK	1599	56.4	271.1	232.8	14.55	9.8	6.7	107.7	6.73
36	ZEPHYR CREEK	846	52.6	271.1	110.8	13.11	9.7	2.1	63.3	7.49
37	SOUTH ZEPHYR CREEK	225	58.8	271.1	60.2	26.76	7.6	4.2	28.5	12.68
38	MCFAUL CREEK	2481	54.1	271.1	305.6	12.32	9.3	4.9	183.7	7.40
39	BURKE CREEK	3308	45.3	225.9	554.2	16.75	7.4	0.0	492.4	14.89
40	EDGEWOOD CREEK	3683	58.7	271.1	481.4	13.07	11.8	0.0	330.9	8.99
41	BIJOU PARK	2771	56.7	271.1	590.6	21.31	8.5	0.0	544.7	19.66
42	BIJOU CREEK	2044	33.8	0.0	1021.3	49.96	4.2	0.0	950.0	46.48
43	TROUT CREEK	26728	49.1	0.0	3904.7	14.61	13.0	0.0	2883.4	10.79
44	UPPER TRUCKEE RIVER	34845	32.7	0.0	8803.8	25.27	10.4	0.0	7114.3	20.42
45	CAMP RICHARDSON	2207	11.5	0.0	1135.3	51.43	4.2	0.0	606.4	27.47
46	TAYLOR CREEK	11318	32.4	0.0	2606.3	23.03	11.1	0.0	2174.0	19.21
47	TALLAC CREEK	2651	15.7	0.0	486.1	18.34	13.8	0.0	272.5	10.27
48	CASCADE CREEK	2885	35.7	0.0	343.6	11.91	17.1	0.0	240.7	8.34
49	EAGLE CREEK	5189	35.3	0.0	539.9	10.41	17.4	0.0	356.3	6.87
50	BLISS STATE PARK	458	32.3	0.0	76.0	16.61	9.0	4.9	42.8	9.34
51	RUBICON CREEK	2123	18.4	45.2	191.6	9.02	17.6	0.0	66.5	3.13
52	PARADISE FLAT	616	24.2	45.2	44.3	7.20	14.4	9.5	34.8	5.66
53	LONELY GULCH CREEK	632	29.2	0.0	52.3	8.27	13.6	16.9	41.2	6.52
54	SIERRA CREEK	739	40.6	0.0	77.6	10.49	14.3	0.0	30.1	4.07
55	MEEKS	5224	37.1	0.0	1007.0	19.28	12.1	0.0	755.5	14.37
56	GENERAL CREEK	5422	28.4	0.0	1786.1	32.94	7.4	0.0	1236.6	22.81
57	MCKINNEY CREEK	3240	23.1	0.0	853.5	26.34	8.1	0.0	449.7	13.88
58	QUAIL LAKE CREEK	1048	15.9	45.2	118.8	11.33	12.3	0.0	69.7	6.65
59	HOMEWOOD CREEK	551	27.9	90.4	52.3	9.48	12.7	8.8	31.7	5.75
60	MADDEN CREEK	1598	28.9	0.0	101.3	6.34	14.8	16.9	82.3	5.15
61	EAGLE ROCK	502	29.3	90.4	101.3	20.19	13.9	4.2	57.0	11.36
62	BLACKWOOD CREEK	7250	37.0	0.0	842.4	11.62	13.6	0.0	606.4	8.36
63	WARD CREEK	7958	30.1	0.0	858.2	10.78	11.2	2.1	574.8	7.22
64		4093	30.0	0.0	440.2	10.75	11.9	0.0	316.7	7.74
		200095								

A data base, no matter how easily constructed and manipulated, is only as reliable as the elements of which it is made up. Therefore, as new data are acquired it is essential that these be tested and implemented into the data base to insure its viability. Although existing digital terrain imagery has proved difficult to adapt, Landsat data continues to work well within the data base concept. However, in terms of water quality, the subject lake and watershed area must be of such dimensions as to accommodate the resolution of the Landsat sensors.

Accurate correlation of changes in Lake Tahoe with various data elements contained within the data base will require study over several years. It is hoped that such a system over time can be used to monitor and evaluate causes for changes which affect the lacustrine environment. This will require the development of precipitation modelling, surface runoff models and classification of drainage basin cover types. These elements must in turn be integrated and evaluated for accuracy before the system can be considered usable. Such a system is feasible given the continued improvement of the remote sensing tools used to construct the data base and the data integrated into it.

C. STUDY OF SELECTED LAKES IN MONTANA, UTAH AND MICHIGAN

Four areas were chosen for study as a final demonstration of lake classification technology developed at JPL. In Montana, Flathead Lake, located in the northwest area of the state, and several lakes within the Kootenai National Forest nearby were selected. The Flaming Gorge Reservoir astride the northern border of Utah and southern Wyoming was also chosen. The Illinois State EPA also expressed a desire to be included in the final demonstration, agreeing to coordinate sampling efforts, on the same lakes studied previously (Ref. I-3), with Landsat flyover dates during the summer of 1978. The U.S. Forest Service, Ottawa National Forest, also expressed an interest in surveying several lakes within their area in the northern part of Michigan. Therefore, in April and May of 1978 the IPL attempted to coordinate ground sampling of selected lakes in these areas with the USEPA Western Energy Office in Denver, Colorado (Kootenai National Forest and Flathead Lake and Flaming Gorge Reservoir), Illinois State EPA in Springfield, Illinois (selected Illinois lakes) and the U.S. Forest Service, Ottawa National Forest, in Ironwood, Michigan (selected Michigan lakes). The intent was to once more test and refine the techniques for lake isolation and classification which had been under development since 1974.

1. Data Acquisition

One of the most significant problems associated with any study using Landsat data is with data acquisition. The problem is basically twofold, with each part independent of the other. Landsat scene availability is always dependent upon favorable weather conditions; cloud cover can virtually rule out the use of any number of Landsat scenes. The other problem, assuming weather conditions to be fair, is scene availability from the data source, such as Goddard Space Flight Center or the EROS Data Center. Although the Landsat satellites theoretically make coverage of a particular area available every 9 days, the combination of the preceding two factors can easily limit availability to zero scenes. In the case of the Illinois lakes selected, although the Illinois EPA was interested in coverage from June until September 1978, due to weather conditions, the periodic

shutdown of the Landsat sensors and presumably data loss during retrieval and recording, only two scenes were available over the selected lakes during this period. Of these scenes, neither was acceptable for processing as both had thin streams of small cumulus clouds running diagonally across the images rendering them essentially useless. Owing to these problems, cooperation with the Illinois State EPA was necessarily halted.

Scene availability for the other areas of interest was somewhat better. However, coordination of ground sampling teams on selected flyover dates was plagued by weather, finances and lack of personnel available to sample. Table II-9 presents a list of Landsat scenes which were acquired for the study areas and indicates the dates when ground truth sampling was performed on these lakes.

2. The Ottawa-Nicolet Forest Lakes

The U.S. Forest Service, Ottawa National Forest, had been monitoring the trophic status of lakes within the boundaries of their jurisdiction. Interest in a Landsat overview of their lakes was expressed in correspondence with USEPA personnel at the Las Vegas Monitoring Laboratory. Two scenes were acquired which

Table II-9. Landsat scenes acquired and sampling dates

Area	Scene I.D.	Date, 1978	% Cloud cover	Lake(s) obscured?	Date of ground truth
Kootenai National Forest and Flathead Lake (upper section)	30146-17515	7/29	10%	no	8/07, 8/08 8/25
Flathead Lake (lower section)	30146-17521	7/29	10%	no	8/07 8/08
Flathead Lake (lower section)	21293-17320	8/07	10%	no	8/07
Nicolet National Forest	21292-15432	8/06	10%	no	8/04
Ottawa National Forest	21293-15491	8/07	30%	yes	8/05
Flaming Gorge Reservoir	21288-17044	8/02	10%	no	8/02
Flaming Gorge Reservoir (upper section)	21287-16590	8/01	10%	Partially	8/02
Flaming Gorge Reservoir (lower section)	21287-16592	8/01	20%	Partially	8/02

covered the lakes of interest in the Ottawa and Nicolet National Forest areas. Landsat 2 scene 21293-15491 contained 30% cloud cover which obscured most of the lakes selected for study. No further processing was performed on this area.

Landsat 2 scene 21292-15432 contained no cloud cover over the area of interest. Processing was limited to color reconstruction of the Nicolet National Forest area. Figure II-29 depicts the Nicolet National Forest subsection from Landsat scene 21292-15432. Two binary masks were created using the program LAKELOC to separate all the water bodies in this subsection from the surrounding land mass. Figure II-30 and II-31 are mask images, in which the threshold for distinguishing water pixels was set at two different levels. As can be seen in these images, the subject lakes were of such small size that further processing using only Landsat data would be of little use. Six bench mark lakes which the Forest Service had been monitoring were processed through LAKELOC in preparation for classification and the results sent to USEPA in Las Vegas. It was determined that due to the relatively small size of these lakes, unsupervised classification would not yield much useful information. No further processing was performed in this area.

3. Flathead Lake and the Kootenai National Forest

USEPA personnel from the Western Energy office in Denver, Colorado, were planning to monitor Flathead Lake in northern Montana during the summer of 1978. In addition the U.S. Forest Service, Kootenai National Forest, expressed interest in a Landsat overview of lakes located within their jurisdiction. Accordingly efforts were made to coordinate sampling of these lakes with Landsat flyover dates. As can be seen in Table II-9, Flathead Lake was sampled on August 7 and 8. The closest Landsat scenes, covering the entire lake area, available for this time frame were acquired on July 29. As Flathead Lake fell between Landsat scanner positions it was necessary to acquire two consecutive scenes to reconstruct the entire lake surface area. A scene of the southern portion of Flathead Lake was also acquired which coincided with the August 7 sampling. However, the August 7 Landsat scene containing the northern portion of Flathead Lake was listed as unavailable at the Goddard Space Flight Center. It is unknown whether this was due to shutdown of the scanner or loss of data during processing by the center.

Sixty-three lakes were identified by the Forest Service as being of interest in the Kootenai National Forest region. Of these, nine were selected to serve as bench mark lakes for the study by agreement of EPALV and Forest Service personnel. The bench mark lakes were sampled by Forest Service personnel on August 25. Again the closest Landsat scene available covering these lakes was the scene acquired on July 29. All other scenes were listed as unavailable by the Goddard Space Flight Center.

a. Processing Kootenai National Forest Lakes. Figure II-32 shows a color reconstruction of the Landsat 3 scene 30146-17515 for July 29, 1978. Flathead Lake appears almost in its entirety in the lower-right-hand section of the image. The Kootenai National Forest area under study lies within the boundaries marked to the left of center within the image. Figure II-33 shows the same area enlarged. Using LAKELOC, analysts at JPL located and isolated the 63 lakes of interest. Figure II-34 is a binary mask of these lakes positioned exactly as they appear in the original Landsat scene. JPL supplied EPALV with this image and the statistical output from the program STATUS for interpretation and comparison with ground

Figure II-29. False color composite showing major portion of Nicolet National Forest. Frame 21292-15432, August 8, 1978.

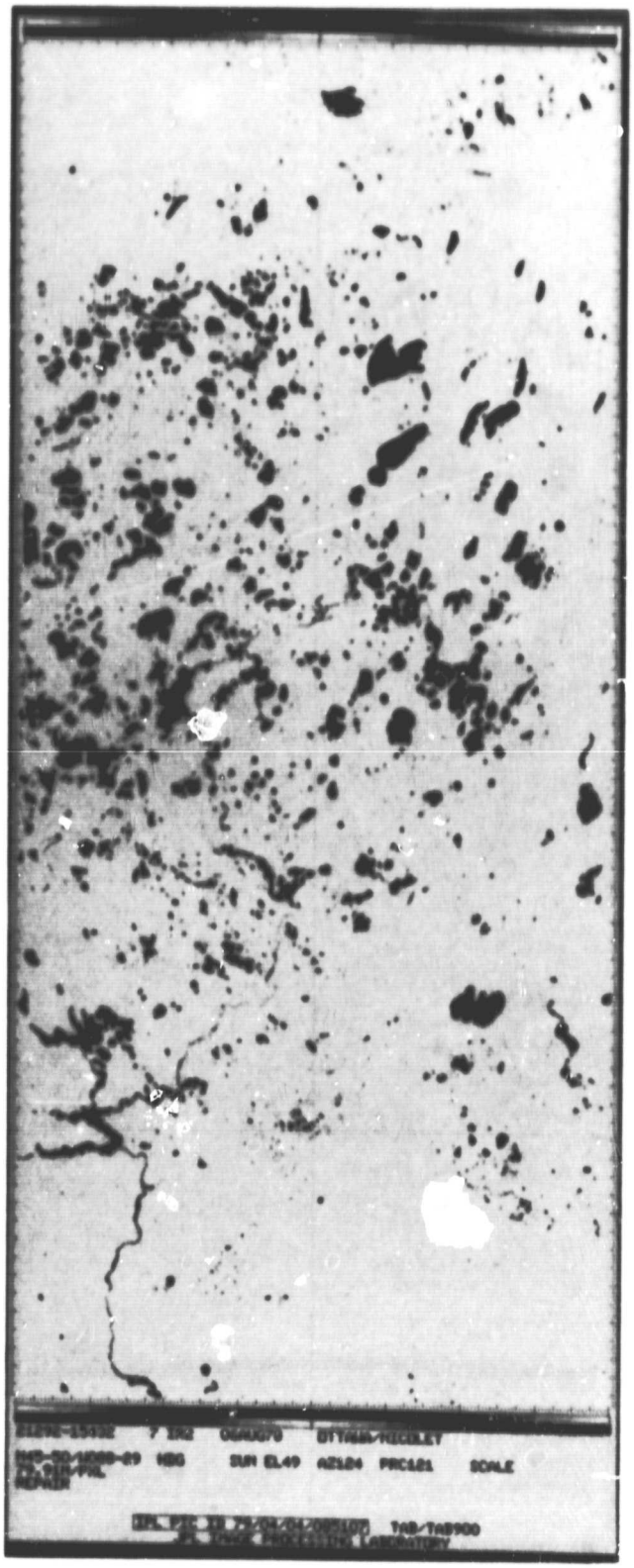

Figure II-30. Binary mask of Landsat NSS Channel 7. First threshold. Landsat frame 21292-15432

Figure II-31. Binary mask of Landsat MSS Channel 7. Second threshold. Landsat frame 21292-15432

Figure II-32. False color reconstruction of Landsat frame 30146-17515. Flathead Lake is seen in the lower right corner.

sampling measurements. Since ground measurements were acquired one month after the Landsat flyover, analysis and comparison of these data were limited. Again due to the small size of the study lakes, no classifications were performed.

b. The Flathead Lake Study Area. Flathead Lake is the largest natural body of fresh water to be found west of the Mississippi River. The lake is located in the northern portion of a large glacial valley near Glacier National Park in northwestern Montana. Flathead Lake is 56.4 by 25.8 kilometers in its extremes of length and width, and it has a maximum depth of 112 meters. (1) The lake shoreline is 185 kilometers, more than half of which is composed of rock and gravel which extends lakeward to a depth of approximately 30 meters. (2) Approximately 50% of the shoreline is composed of rocky cliffs. The steepness of much of the drainage area and its close proximity to the lake greatly limits the amount of drainage into the lake.

Figure II-33. A portion of Landsat frame 30146-17515 showing the Kootenai National Forest study area.

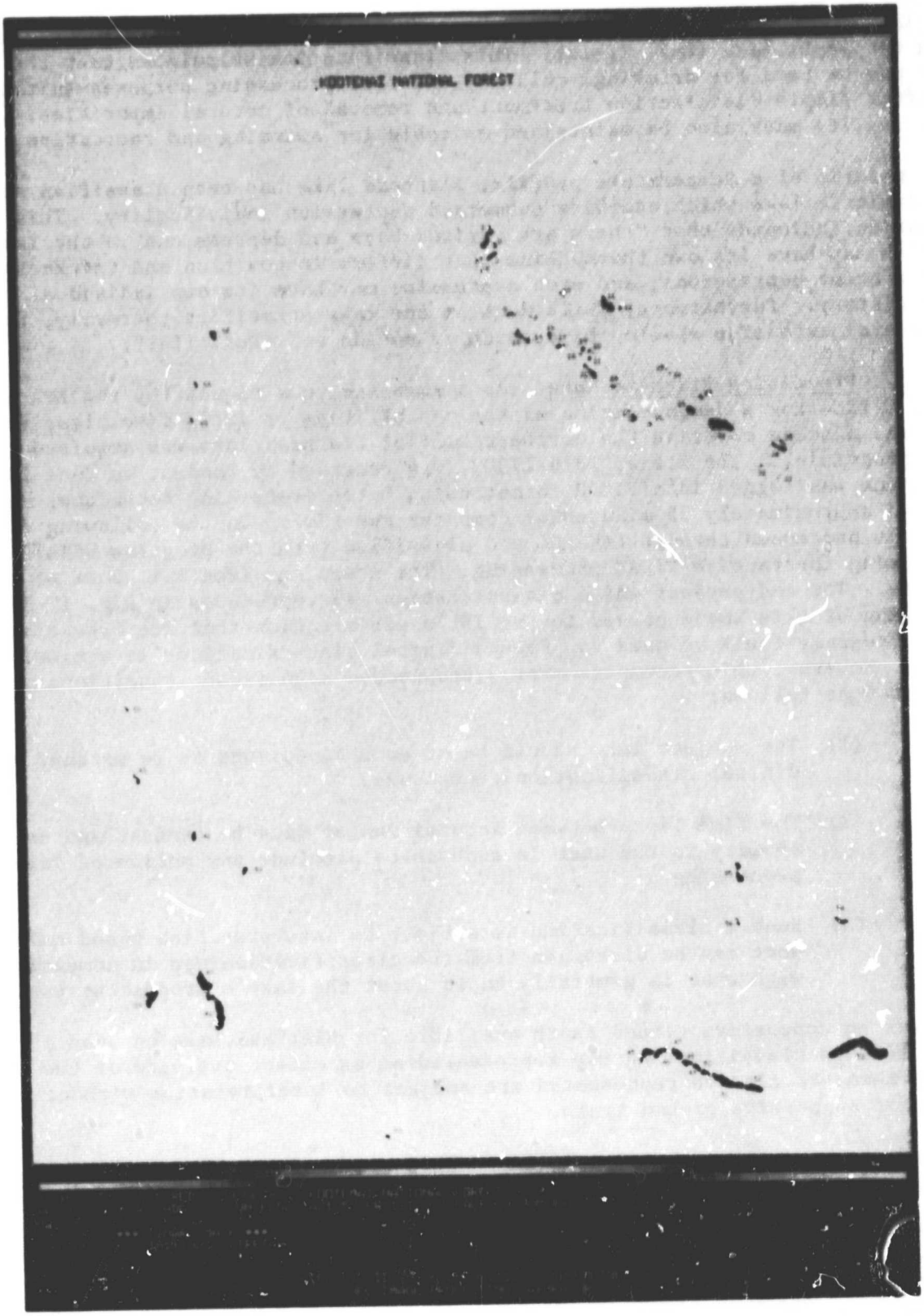

Figure II-34. A binary mask of the 63 lakes of interest in the Kootenai National Forest

The Montana Water Pollution Control Authority has classified Flathead lake as an A-open D1 lake (Ref. II-15). This classification stipulates that the lake's water may be used for drinking, culinary and food processing purposes suitable for use after simple disinfection treatment and removal of natural impurities. The water quality must also be maintained suitable for swimming and recreation.

In terms of a temperature profile, Flathead lake has been classified as a cold dimictic lake which exhibits submerged depression individuality. This classification indicates that "there are various bays and depressions in the lake each of which may have its own thermocline that differs in position and thickness in the different depressions, and each depression may have its own individual seasonal history. Furthermore, while much of the lake stratifies thermally, there are several extensive shallow bays which do not do so" (Ref. II-15).

c. Processing Flathead Lake. As a response to a request by the EPA Western Energy Office for a demonstration of the capabilities of IPL's lake classification programs, a scene covering the northern half of Flathead lake was acquired for trial processing. The scene, 2879-17303, was recorded by Landsat on June 19, 1977. This scene was logged into VICAR format using batch processing techniques which required approximately 25 minutes of computer run-time. On the following day the scene was processed through LAKELOC and classified with the programs USTATS and BAYES using interactive VICAR processing. The steps required 1.1 hours of computer run-time. The end product was a classification map reproduced in Fig. II-35. Processing of this scene proved to the IPL's satisfaction that the lake classification programs could be used to produce digital classifications in a timely manner. However, this accomplishment acknowledged significant conditions which are listed as follows:

(1) The subject lake should be of such dimensions as to warrant using digital classification techniques.

(2) The time lapse between acquisition of data by Landsat and availability to the user is such as to preclude any notion of "real-time" processing.

(3) Such a classification is subject to interpretation based only on what can be discerned from the classification map in comparison with what is generally known about the lake's trophic status.

There was no supportive ground truth available for Flathead Lake on June 19, 1977. Therefore this classification map represented an excellent overview of the lake, but the thematic classes represented are subject to interpretation without the benefit of supportive ground truth.

Efforts were subsequently made to include Flathead Lake as a study area in the final demonstration phase of the Lake Classification projects. EPA personnel at the Western Energy Office in Denver agreed to provide sample data coinciding with Landsat flyover dates. As indicated previously, Landsat data was limited to 1 scene recorded on August 7, for which ground truth directly corresponded. Figure II-36 shows a map of Flathead Lake with ground sampling sites as marked by EPA personnel. Two other scenes were also acquired which were recorded by Landsat 9 days prior to the ground sampling.

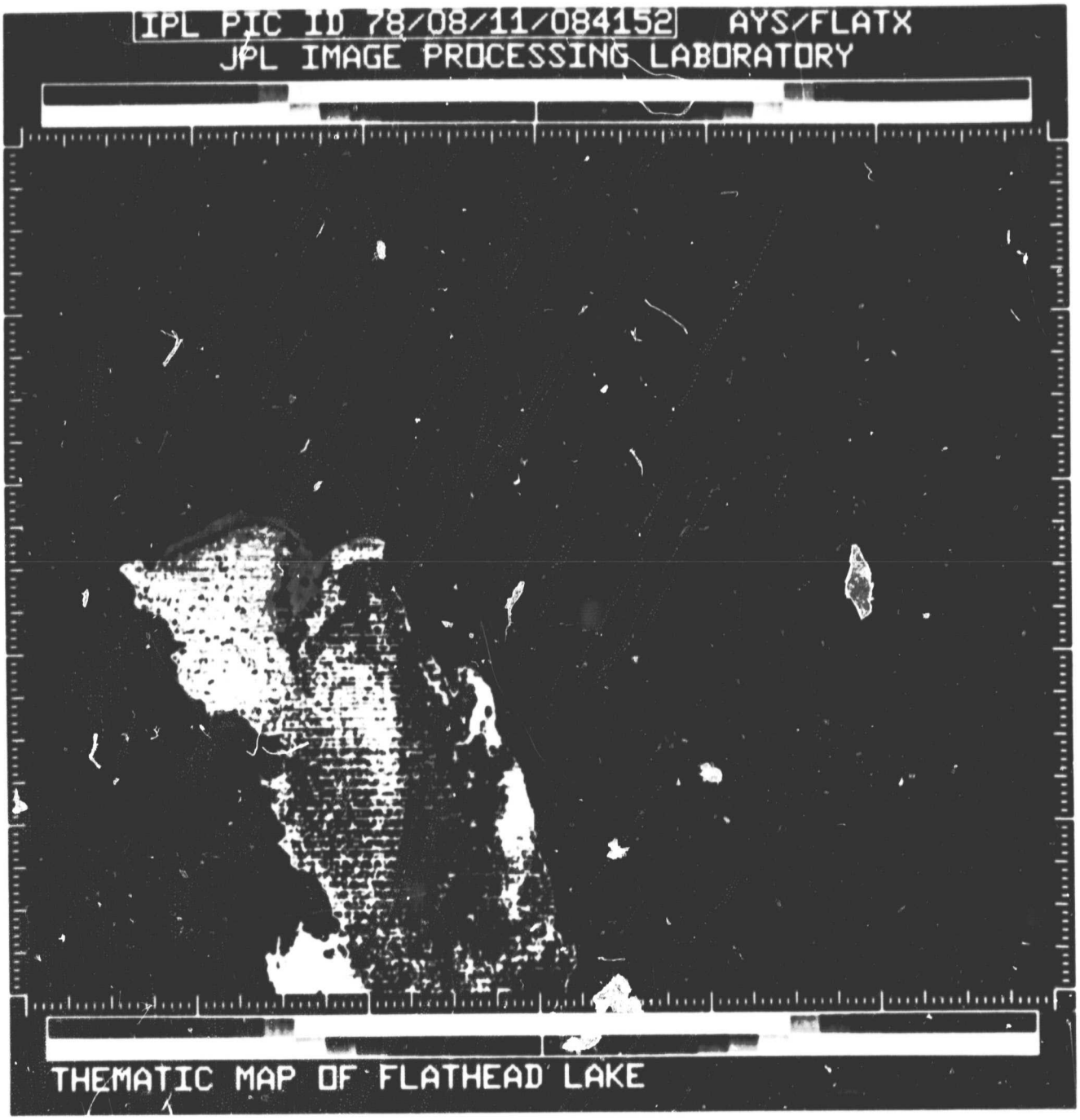

Figure II-35. An unsupervised classification map of Flathead Lake. Landsat frame 30146-17515.

ORIGINAL PAGE
COLOR PHOTOGRAPH

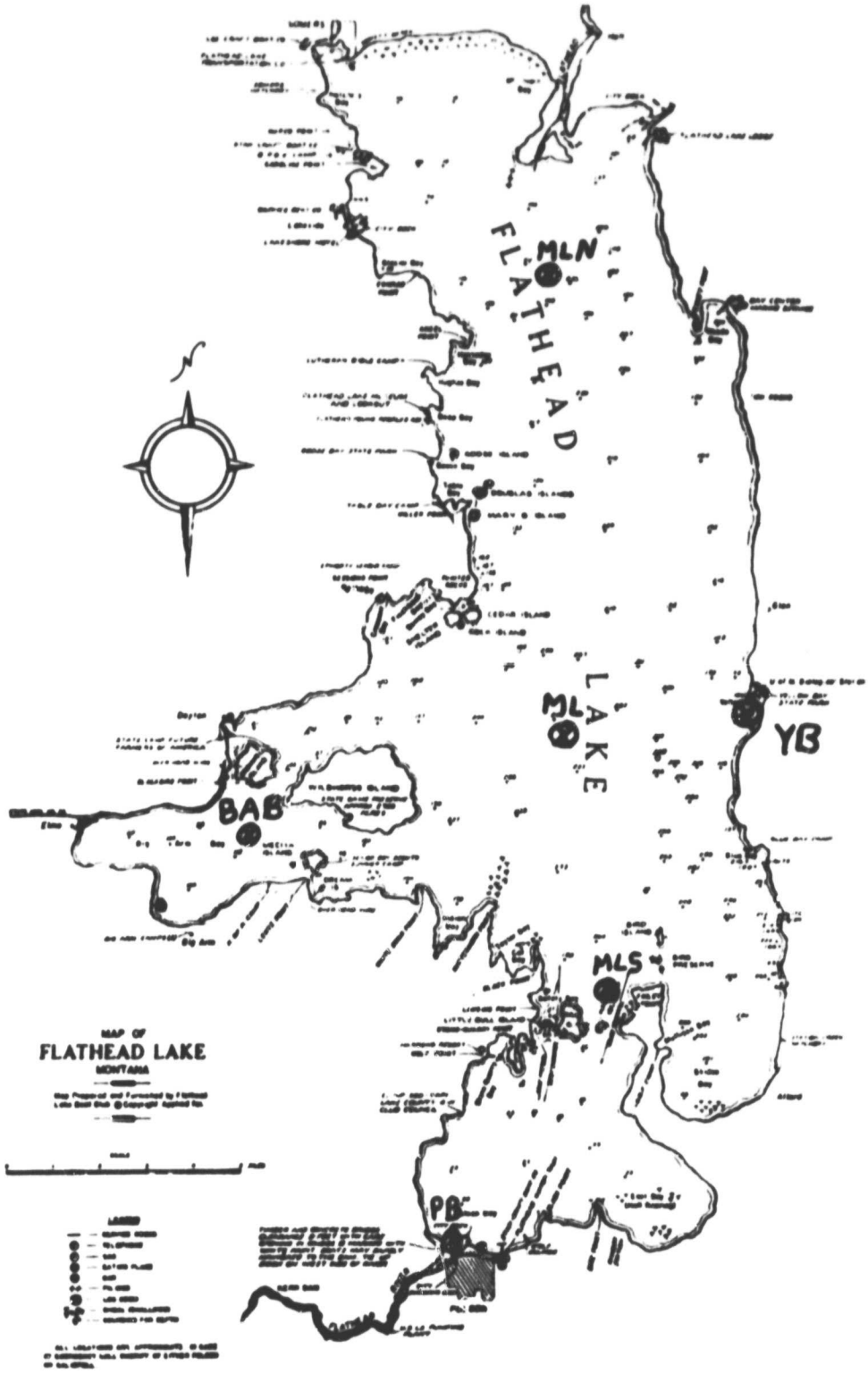

Figure II-36. A conventional map of Flathead Lake. Water truth sites are marked on this map

Since Flathead Lake lies within two consecutive Landsat scenes, it was necessary to digitally connect two scenes to create one image of the total lake surface area. Figure II-37 is a color composite image constructed from the two Landsat scenes of Flathead Lake acquired on July 29, 1978. The VICAR program STATS2 was used to gather multispectral statistics for each sampling site to be used in conjunction with analysis of classification statistics and ground measurements taken nine days prior to the Landsat flyover. Table II-10 lists the Landsat statistics from STATS2. These products, accompanied by classification statistics, were sent to EPALV for analysis. As of this writing these data are still under interpretation by EPA laboratory personnel.

Landsat scene 21293-17320 acquired on August 7, 1978, was also processed. As mentioned previously this scene contained only the lower half of the lake due to the Landsat scanner position. Since corresponding ground truth measurements were available for this scene, it was classified using the unsupervised routine in USTATS. Three spectral classes were readily identified using USTATS. These multispectral statistics for each class were sent to EPALV for analysis. These data remain under study by EPA personnel.

4. The Flaming Gorge Study Area

The Flaming Gorge Reservoir is a long body of water located in Sweetwater County, Wyoming, to the north and Doggett County, Utah, as its southern extreme. Flaming Gorge is approximately 146 kilometers long with a surface area of 137.67 km^2. The mean depth of the reservoir is 33.9 meters with a maximum depth of 153.0 meters (Ref. II-16).

EPA observation conducted at Flaming Gorge Reservoir indicate the following conclusions regarding its trophic condition.

> Flaming Gorge Reservoir ranked third in overall trophic quality among the 14 Wyoming lakes and reservoirs sampled in 1975 when compared using a combination of six water quality parameters.

5. Processing Flaming Gorge Reservoir

Coordination of ground truth sampling and Landsat flyover was most successful for the Flaming Gorge Reservoir. Located on the border of Utah and Wyoming the reservoir is a significant site to warrant the use of Landsat data. Three Landsat scenes were acquired covering Flaming Gorge on August 1 and 2, 1978. The scenes were acquired on consecutive days due to the 14% overlap in Landsat scanner paths making available a "second look" on August 2.

Two scenes were required to entirely cover the reservoir surface on August 1 due once again to the Landsat scanner position. Figure II-38 reproduces a composite image which was created from the two consecutive scenes acquired on August 1. Cloud cover obscured a portion of the reservoir surface. It was determined that the presence of clouds and cloud shadow would affect the classification results. However, the decision was made to attempt an unsupervised classification despite the interference. Figure II-39 presents a thematic map produced through USTATS and BAYES for the August 1 composite scene. As can be discerned, there was considerable confusion in class discrimination, which caused classes to

Figure II-37. A color composite image of Flathead Lake using two Landsat scenes (30146-17515 and 30146-17521).

Table II-10. Landsat MSS statistics from training sites on Flathead Lake mosaic (frames 30146-17515 and 30146-17521).

Training site	Matrix size	Mean DN				Sigma			
		Green	Red	IR1	IR2	Green	Red	IR1	IR2
MLN	5 x 5	29.3	13.3	5.2	3.4	2.1	1.8	1.1	2.0
ML	5 x 5	24.1	10.8	5.7	3.0	1.6	1.5	1.3	1.7
YB	5 x 5	23.8	12.1	6.8	5.8	1.4	2.3	0.9	3.1
BAB	5 x 5	25.8	14.4	7.5	6.1	1.6	1.5	1.3	2.0
MLS	5 x 5	23.0	10.3	5.0	2.3	1.3	1.0	1.8	1.6
PB	5 x 5	28.3	13.1	6.6	4.7	2.6	1.1	1.2	1.2

become fragmented or speckled. This was probably due to atmospheric interference from shadows cast by clouds. The effects are most pronounced in areas of cloud concentration.

More reliable results were obtained when scene 21288-17044, acquired on August 2, was processed. As shown in Fig. II-40, this scene was not affected by cloud cover. As a result the multispectral classification for this scene, reproduced in Fig. II-41, does not exhibit the same confusion in classes. Class color distribution is fairly uniform throughout the reservoir. Corresponding ground truth measurements were available for this scene from EPA's Western Energy Office. In addition, the locations for sample sites indicated by EPA personnel were identified. Multispectral statistics for each sample site were also acquired using STATS2. These statistics are reproduced in Table II-11. This information was sent to EPALV for comment and analysis. As of this writing it remains under EPA study.

The products presented in this section are examples of the results of digital image processing techniques which can be performed on water quality data supplied by Landsat. No attempt has been made to analyze the accuracy of the multispectral classifications without expert input from EPA laboratory personnel. Until EPA analysis is complete no significant conclusions can be drawn regarding the processing performed during this study.

Figure II-38. A color composite image of Flaming Gorge Reservoir developed by mosaicking two Landsat scenes together (21287-16590 and 21287-16592), August 1, 1978.

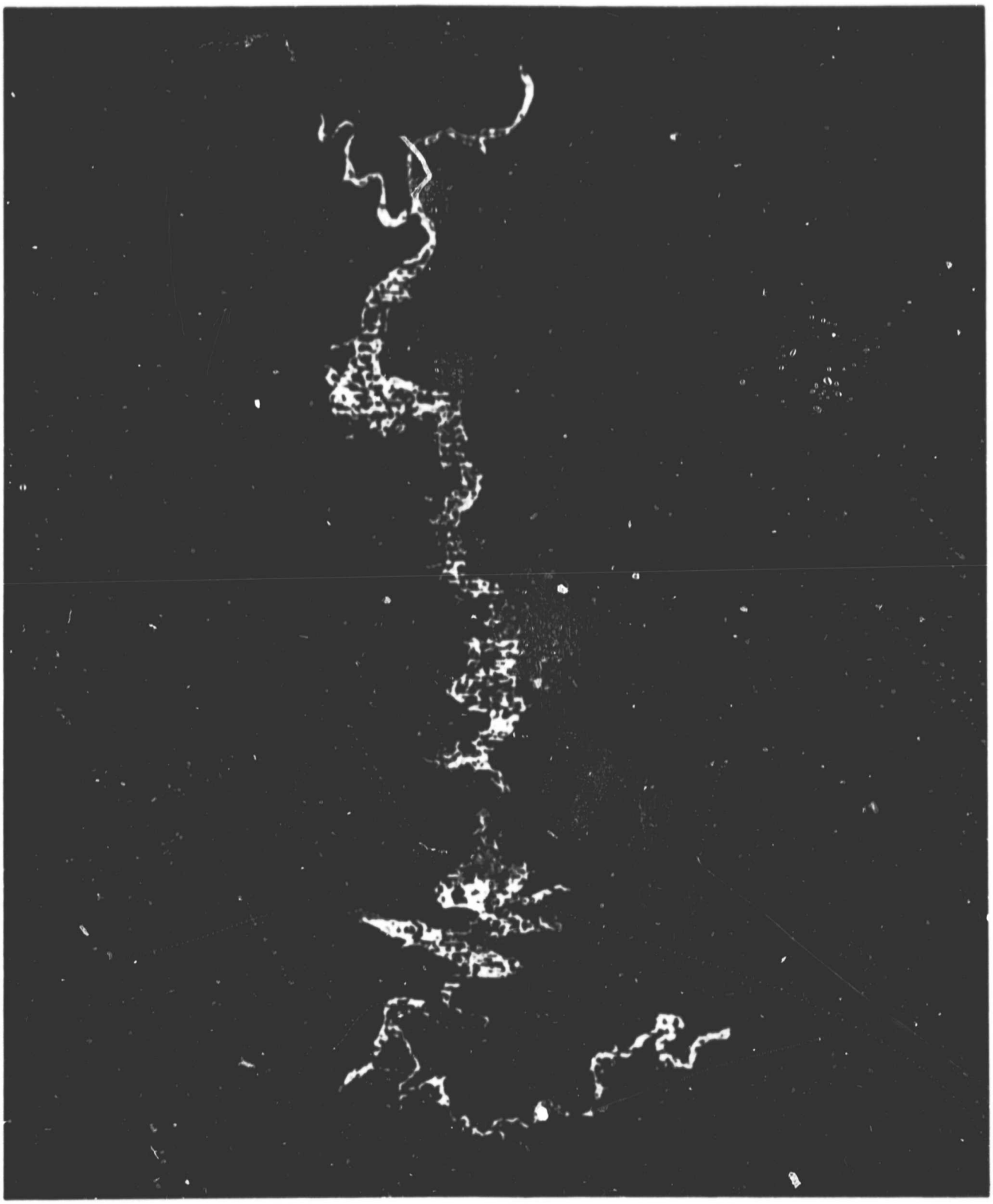

Figure II-39. An unsupervised classification map of Flaming Gorge Reservoir for August 1, 1978.

Figure II-40. A color composite image of Flaming Gorge Reservoir, Landsat scene 21288-17044, August 2, 1978.

Figure II-41. A supervised classification map of Flaming Gorge Reservoir for August 2, 1978.

Table II-11. Landsat MSS statistics from training sites on Flaming Gorge Reservoir (Frame 21288-17044)

Training site	Matrix size	Mean DN				Sigma			
		Green	Red	IR1	IR2	Green	Red	IR1	IR2
Site 1	2 x 2	39.2	37.2	13.5	3.0	.8	.8	.8	0
Site 2	4 x 4	31.0	25.1	10.1	3.0	1.0	1.2	1.8	0
Site 3	1 x 4	45.0	40.0	15.0	5.0	1.8	1.2	1.2	0
Site 4	4 x 4	25.3	20.5	8.4	2.0	1.1	1.1	1.6	0
Site 5	4 x 4	26.2	21.1	8.8	2.0	1.0	.8	1.8	0
Site 6	5 x 5	21.0	15.6	5.8	1.0	.9	1.2	1.2	0
Site 7	5 x 5	20.6	15.9	6.5	1.0	1.3	1.2	.8	0
Site 8	5 x 5	22.2	16.4	7.0	2.0	.6	.8	1.5	0
Site 9	5 x 5	19.8	13.3	4.3	1.0	.9	1.0	1.0	0
Site 10	5 x 5	22.3	15.3	5.8	2.0	.8	.8	1.2	0
Site 11	5 x 5	21.0	14.6	16.2	1.0	.9	.9	1.7	0
Site 12	5 x 5								
Site 13	5 x 5	21.7	14.4	6.9	1.7	1.4	3.3	5.7	4.8
Site 14	5 x 5	22.5	14.4	4.8	2.0	1.3	1.2	1.4	1.0
Site 15	5 x 5	21.9	14.3	7.1	2.1	1.2	2.6	5.7	6.1
Site [illegible]	5 x 5	21.7	14.5	5.3	1.0	.8	1.2	1.2	0

REFERENCES

I-1. Blackwell, R.J., and D.H.P. Boland, "The Landsat-1 Multispectral Scanner as a Tool in the Classification of Inland Lakes," 1st Symposium on the Practical Application of Earth Resources Data, NASA TM-58168, JSC 09930, pp. 419-442, June 1975.

I-2. Blackwell, R.J., and D.H.P. Boland, Trophic Classification of Selected Colorado Lakes, Publication 78-100, Jet Propulsion Laboratory, Pasadena, Calif., January 1979.

I-3. Boland, D.H.P., D.J. Schaeffer et al., Trophic Classification of Selected Illinois Water Bodies Using Landsat Data, EPA-600/3-79-123, Las Vegas, Nevada, December 1979.

I-4. Boland, D.H.P., Trophic Classification of Lakes Using Landsat-1 (ERTS-1) Multispectral Scanner Data, EPA-600, Covallis, Oregon, April 1976.

I-5. Smith, A.Y., and J.D. Addington, "An Interactive Lake Survey Program," S.P.I.E., Volume 119, 7 pp. 1977.

I-6. Horowitz, H.M., et al., "Estimating the Proportions of Objects Within A Single Resolution Element of A Multi-Spectral Scanner," 7th International Symposium on Remote Sensing, U. Michigan, Ann Arbor, Michigan, pp. 685-694, 1971.

I-7. Work, E.A., and D.S. Gilmer, "Utilization of Satellite Data for Inventorying Prairie Ponds and Lakes," Photogrammetric Engineering and Remote Sensing, pp. 685-694, 1976.

I-8. McCloy, K.R., "The Vector Classifier," 11th Symposium on Remote Sensing of Environment, U. Michigan, Ann Arbor, Michigan, pp. 52, 1977.

I-9. Bryant, N.A., and A.L. Zobrist, "An Image Based Information System: Architecture for Correlating Satellite and Topological Data Bases," 1st Users Conference on Computer Mapping, Harvard University, Cambridge, Mass., 1978.

II-1. Everett, L.G., L.S. Leonhart and L.K. Lepley, "An Evaluation of ERTS-1 Imagery in Reservoir Dynamics," 4th Annual Conference on Remote Sensing in Arid Lands, U. Arizona, Tucson, Arizona, p. 259, 1976.

II-2. Deacon, J.E., and W.L. Haskell, "Observations on the Ecology of the Freshwater Jelly Fish in Lake Mead, Nevada," The American Midland Naturalist, p. 155, 1966.

II-3. USEPA, "Report on Lake Mead, Clark County, Nevada; Mohave County, Arizona," EPA Region IX, Working Paper No. 808, Las Vegas, Nevada, p. 1, 1977.

II-4. Anderson, E.R., and Pritchard, Physical Limnology of Lake Mead, Navy Electronics Laboratory, Rept. 258, pp. 1-152, 1951.

II-5. La Rivers, I., Fishes and Fisheries of Nevada, Nevada Fish and Game Commission, Reno, Nevada, 1962.

C-2

REFERENCES (Contd)

II-6. Reid, G.K., Ecology of Inland Waters and Estuaries, Reinhold Publishing Corp., New York, p. 116, 1961.

II-7. Ahern, F.J., and D.G. Goodenough, "Landsat Atmospheric Corrections at CCRS," 4th Canadian Symposium on Remote Sensing, Quebec City, Canada, May 1977.

II-8. Fraser, R.S., O.P. Bahenhi and A.H. Al-Abbas, The Effect of the Atmosphere on the Classification of Satellite Observations to Identify Surface Features, 6th Symposium on Remote Sensing of the Environment, U. Michigan, Ann Arbor, Michigan, pp. 229-249.

II-9. Rochon, G., and F.J. Langham, "Sur la Transformation des Radiances en Reflectances pour L'etude de Qualite de L'eau," 4th Canadian Symposium on Remote Sensing, Quebec City, Canada, pp. 393-403, May 1977.

II-10. Odum, E.P., Fundamentals of Ecology, W.B. Saunders Company, Philadelphia, pp. 295-323, 1971.

II-11. Ruiz, R.M., et al., "IPL Processing of the Viking Orbiter Images of Mars," J. Geophys. Res., Vol. 82, No. 28, 1977.

II-12. Davis, J.B., and S.Z. Friedman, "Assessing Urbanized Area Expansion Through the Integration of Landsat and Conventional Data," Proceedings, 45th Meeting of the American Society of Photogrammetry, pp. 776-79[illegible], 1979.

II-13. Madura, D.P., and J.M. Soha, Color Enhancement of Landsat Agricultural Imagery, Publication 78-102, Jet Propulsion Laboratory, Pasadena, Calif., December 1978.

II-14. Strahler, A.H., T.L. Logan and C.E. Woodcock, "Forest Classification and Inventory System Using Landsat, Digital Terrain and Ground Sample Data," 13th International Symposium on Remote Sensing of the Environment, Ann Arbor, Michigan, April 1979.

II-15. Brink, C., Water Quality Criteria, Montana Water Pollution Control Council, Helena, Montana, 1967.

II-16. EPA Region VIII, Report on Flaming Gorge Reservoir Working Paper No. 885, EPA Region VIII, Denver, Colorado, December 1977.