NASA
National Aeronautics and
Space Administration

AUG 21 1980

JSC-13821
Revision A

Lyndon B. Johnson Space Center
Houston, Texas 77058

81-10010
CR-163562

EARTH OBSERVATIONS DIVISION

SPACE AND LIFE SCIENCES DIRECTORATE

EARTH OBSERVATIONS DIVISION VERSION OF THE LABORATORY FOR APPLICATIONS OF REMOTE SENSING SYSTEM (EOD-LARSYS) USER GUIDE FOR THE IBM 307/148

VOLUME II — USER REFERENCE MANUAL

Job Order 76-662

Prepared By

Lockheed Engineering and Management Services Company, Inc.
Houston, Texas

Contract NAS 9-15800

(E81-10010) EARTH OBSERVATIONS DIVISION VERSION OF THE LABORATORY FOR APPLICATIONS OF REMOTE SENSING SYSTEM (EOD-LARSYS) USER GUIDE FOR THE IBM 370/148. VOLUME 2: USER REFERENCE MANUAL (Lockheed Engineering and G3/43

N81-12487

Unclas
00010

June 1980

LEMSCO-12564
Revision A

JSC-13821
Revision A

EARTH OBSERVATIONS DIVISION VERSION OF THE LABORATORY FOR APPLICATIONS OF REMOTE SENSING SYSTEM (EOD-LARSYS) USER GUIDE FOR THE IBM 370/148

VOLUME II — USER REFERENCE MANUAL

Job Order 76-662

PREPARED BY

P. J. Aucoin and J. Stewart
Earth Observations Data Products Department

and

M. F. McKay
Technical Publications Department

APPROVED BY

NASA

J. M. Sulester, Technical Monitor, Systems and Facilities Branch

LOCKHEED

T. C. Mackin, Supervisor
Exploratory Investigations Section

Prepared By

Lockheed Engineering and Management Services Company, Inc.

For

Earth Observations Division

Space and Life Sciences Directorate

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

June 1980

LEMSCO-12564
Revision A

PREFACE

The system which is the subject of this documentation was designed originally for execution on the Univac 1108/1110 computer at the Laboratory for Applications of Remote Sensing, Purdue University. This volume II documented the conversion in 1978 of the EOD-LARSYS software for execution on the IBM 370/148. In 1979, the IBM 370/148 was replaced by the IBM 3031 computer, which is thoroughly compatible with the software as altered for execution on the IBM 370/148. Thus, no conversion of software is required for this system to be operable on the IBM 3031 computer. This revision primarily documents enhancements and additions to the system.

CONTENTS

Section		Page

Section | Page

Section | Page

Section | Page

Section | Page

Appendix		Page

TABLES

FIGURES

Figure		Page

1. INTRODUCTION

For several years, the Earth Observations Division (EOD) of the National Aeronautics and Space Administration (NASA) Lyndon B. Johnson Space Center (JSC) has supported research for the development of techniques to be used in processing remotely sensed imagery data obtained from the multispectral scanner (MSS) placed aboard various aircraft and satellites. One of the earliest operational computer systems to use pattern recognition techniques in the analysis of these data was developed at Purdue University's Laboratory for Applications of Remote Sensing (LARS). The earliest version of the LARS system (LARSYS) was converted in 1970 to a batch program for execution on the Univac 1108 EXEC 2 system at NASA/JSC.

The computer system described in this document originated from this early version of LARSYS. However, since 1970, personnel of the Earth Observations Division, Lockheed Electronics Company, Inc., now Lockheed Engineering and Management Services Company, Inc. (LOCKHEED), and other EOD support contractors have made many modifications and improvements to the Univac 1108 version of the LARSYS; thus, new techniques have been developed and programmed to perform additional functions in the interpretation of the data.

Although the basic structure of the system remains the same, a large portion of it has been reprogrammed. Modifications to existing techniques and the addition of new techniques have expanded the capabilities of the system. The current version is called EOD version of the LARSYS (EOD-LARSYS). In 1977, the software was converted from Univac 1108 Fortran to IBM Fortran IV for execution on the IBM 370/148 computer at the LARS. The EOD-LARSYS is now operational on the IBM 3031 computer located at

LARS, Purdue University, West Lafayette, Indiana. It may be accessed both on and off the JSC site by remote terminal.

The purposes of this document are to define the capabilities and limitations of the system and to provide the user with the information needed to execute the program and obtain the output desired. It is assumed throughout the document that the user is familiar with the terminology and the pattern recognition techniques involved. No attempt is made to assist the user in the analysis of data output by the system.

This EOD-LARSYS User Guide is being issued in four volumes:

Volume I — System Overview (JSC-13821; LEC-12563, Revision A), May 1980

Volume II — User Reference Manual (JSC-13821; LEC-12564, Revision A), June 1980 (this document)

Volume III — As-Built Documentation (JSC-13821; LEC-12565, Revision A), April 1980

Volume IV — Program Listings (JSC-13821; LEC-12566, Revision A), November 1979

2. GENERAL DESCRIPTION AND OPERATION

2.1 DESCRIPTION

The EOD-LARSYS is a processing program operational on the IBM 3031 system at Purdue University. The system is composed of a system monitor and a set of processors, each of which performs a specific function in the analysis of MSS imagery data. Linkage between processors is accomplished by the use of files in the computer; by files on disk or tape; or, less commonly, by card decks. The execution of a particular batch job may begin or end with any processor, provided the appropriate files are furnished.

Two pattern recognition classification schemes are provided by the system. One, the supervised classification algorithm known as the maximum likelihood classifier, is embodied in the CLASSIFY processor (ref. 1).* The other, an unsupervised classification or clustering algorithm, is embodied in the Iterative Self-Organizing Clustering System (ISOCLS) processor (ref. 2). ISOCLS, along with other processors, may be used to "train" the maximum likelihood classifier or to display the results of classification. TESTSP is provided as a more efficient version of ISOCLS (in terms of disk space required).

Having obtained an MSS image data tape (DATAPE) in one of the allowable formats (see section 3.1), the data analyst must train the classifier. The maximum likelihood classification algorithm is based on the assumption that the samples within a given class are distributed according to a multivariate normal probability density function. Such a distribution is specified completely in terms of a mean vector and a covariance matrix, which must be computed from known samples of the class being represented. This implies that the data analyst must have some prior knowledge

*References are given in appendix A.

(i.e., ground-verified information) of specific areas within the MSS image. Using this ground truth, the analyst must identify training samples for computation of statistics. The histogram (HIST) and gray map (GRAYMAP) processors may be used to aid the analyst.

The HIST processor provides a histogram of data values from the MSS image for use by the GRAYMAP processor. HIST may also be used independently to provide the analyst with information on the distribution of data values within specific user-defined blocks (or fields) of the image. The mean, standard deviation, and range of data within each user-defined field are standard outputs from the HIST processor for each requested channel. Histogram plots may be obtained optionally. With the histogram information, a file (HISFIL) is written automatically for the GRAYMAP processor.

The GRAYMAP processor provides the analyst with a pictorial gray-scale map of any channel of the MSS image for use in obtaining training field coordinates. The map is labeled by sample and scan-line numbers. From this map, the analyst may locate the fields within the image for which he or she has ground truth. Having identified the fields, the analyst must note the coordinates (sample and scan-line numbers at each vertex) to define the fields for the statistics (STAT) or ISOCLS processor.

Alternatively, the analyst may proceed using Procedure 1 methodology. Using an ERIPS* tape as input to the DOTDATA processor, the analyst writes a dot data file. A dot data file contains both type 1 and type 2 dots. Type 1 dots are used both as starting vectors for the clustering processor (ISOCLS) and as labeling

---

*Earth Resources Interactive Processing System.

4

vectors for the labeling processor (LABEL). Type 2 dots are used as a bias correction factor in computing the classification results output by the DISPLAY processor.

The coordinates for training fields may be input to either the STAT or the ISOCLS processor for computation of statistics for the classifier. The dot data file is input to ISOCLS to compute statistics. Both processors save the statistics and training field or dot information on a file (SAVTAP, section 4.1) for use in other processors.

In using the STAT processor, the user must group training fields into statistically similar subclasses. Subclasses may be grouped further into classes. For example, three statistically similar subclasses of spring wheat, winter wheat, and harvested wheat may be grouped into one wheat class. Statistics for each subclass are maintained on the SAVTAP file, along with the class grouping. Class groupings are maintained simply for convenience in defining categories in the CLASSIFY processor and for performance reports by the DISPLAY processor. The analyst may obtain the following output for each training field and/or subclass.

- Mean vector
- Covariance matrix
- Correlation matrix (among channels)
- Histogram plots
- Spectral plots

In using the ISOCLS processor with training classes, the user must group the training fields into classes. The clustering process breaks the class data into statistically similar subclasses (clusters). Subclasses are given names by taking the first two characters of the class name and two digits indicating the number of the subclass within the class. Again, the statistics for each subclass are saved on the SAVTAP file for use in other processors. ISOCLS is an iterative self-organizing clustering procedure which uses the measure of absolute (L_1) distance from a picture element (pixel) to the cluster center to determine the similarity of pixels. At each iteration the user may obtain a cluster summary and map. Optionally, a cluster image data file (MAPUNT) may be output in either LARSYS II/III or Universal format (appendixes B and C, respectively).

Under Procedure 1, designated other/designated unidentifiable (DO/DU) fields are delineated by card input. ISOCLS clusters the LACIE* segment using the starting vectors from the dot data file to begin the clustering process. Sun angle correction is provided. An unconditional cluster map and a set of cluster statistics are output.

The cluster statistics, cluster map, and dot data file are input to the LABEL processor. Using one of two procedures, k-nearest-neighbor or all-of-a-kind, the cluster statistics are labeled. A conditional or mixed cluster map may be output to be displayed later on the Passive Microwave Imaging System/Data Analysis Station (PMIS/DAS) or the Interactive Multispectral Image Analysis System, Model 100 (Image 100).

By the use of the transform statistics (TRSTAT) processor, the statistics on the SAVTAP file output by the processor STAT or ISOCLS may be transformed according to

$$\left.\begin{aligned} \mu' &= A\mu + b \\ K' &= AKA^T \end{aligned}\right\} \qquad (2\text{-}1)$$

*Large Area Crop Inventory Experiment, for which Procedure 1 was developed.

where

μ' = transformed mean vector

A = a matrix

μ = mean vector from SAVTAP file

b = a vector

K' = transformed covariance matrix

K = covariance matrix from SAVTAP file

A^T = transpose of matrix A

The transformed statistics are output as a new file on SAVTAP.

Before proceeding to classification of the MSS image, it may be desirable to reduce the dimensionality of the data vectors by selecting a smaller set of channels or a linear combination of the channels which maximizes some class separability measure. The SELECT processor (ref. 3) provides this capability. In order to compute the value of the separability measures, the statistics calculated using the STAT or ISOCLS processor must be made available to the SELECT processor by card deck, tape, or disk file, usually the last. The SELECT processor allows the analyst to work with subsets of the statistics on the file, if he or she desires. Subsets of the statistics are indicated by the CHANNELS and SUBCLASSES control cards, which are defined further in table 10-1.

In addition, the statistics for two or more subclasses may be grouped together and considered as one subclass. Grouping the statistics for two or more subclasses is equivalent to going back through the STAT processor and combining all training fields for those subclasses being grouped into one subclass. The grouping option is exercised via the GROUP control card defined in table 10-1. The subsets and groupings of the statistics provided to the SELECT processor for computation are used only in SELECT and are not passed on to other processors.

The SELECT processor also allows the analyst to evaluate a given set of channels using one of three different separability measures or to select the best set of channels (k) out of the total channels (n) on the basis of one of these separability measures. The three separability measures provided are

a. Weighted average interclass divergence

b. Weighted average transformed divergence

c. Weighted average Bhattacharyya distance

To select the best set of k out of n channels, the analyst may use either the without replacement procedure or the exhaustive search procedure. A third procedure will find k linear combinations of n measurements which extremize a given separability measure. This procedure, known as the Davidon-Fletcher-Powell procedure, outputs the linear combinations in matrix form. All the procedures and equations for separability measures referred to above are discussed in detail in reference 4.

After the SELECT processor has determined the subset or the linear combination of channels which maximizes subclass separability, the supervised classification of the imagery data is performed by the CLASSIFY processor. The options available in SELECT for grouping and selecting subsets of the SAVTAP file are also available in CLASSIFY. However, once the statistics for classification have been specified, the classes and subclasses are renumbered and referred to in the DISPLAY processor by the new numbers.

The CLASSIFY processor allows the user to group classes previously defined by the STAT or ISOCLS processor into categories for the sum-of-densities classification. When a category is defined (by class names), all subclasses in each class are assigned to the category. The density function for category r, $P_r(X)$, is the sum of densities for all subclasses in the category; that is,

$$P_r(X) = \sum_{i=1}^{M_r} P_{ri}(X) \qquad (2\text{-}2)$$

where

P_{ri} = the probability density function for subclass i multiplied by the a priori value

M_r = the number of subclasses in category r

(Note: More detailed equations are given in section 11.)

Pixel X is assigned to category r if $P_r(X) > P_s(X)$ for all categories s, $s \neq r$. Pixel X is further assigned to subclass i if (1) i belongs to category r and (2) $P_{ri}(X) > P_{rm}(X)$ for all subclasses m, $m \neq i$, and m belongs to category r.

Obviously a one-to-one correspondence between categories and subclasses reduces the above equation to

$$P_r(X) = P_{ri}(X) \qquad (2\text{-}3)$$

When this is the case, the amount of computation required for classification can be greatly reduced by the use of thresholds. CLASSIFY then has two procedures for classification which use this computational reduction to advantage. The sum-of-densities rule is used only when categories are defined by the user. Otherwise, the classification-by-thresholding procedure detailed in reference 4 is used.

The CLASSIFY processor writes a file (MAPTAP, section 4.4 and appendix D) containing the subclass number and confidence level for each pixel classified; the training fields and statistics for the classes and subclasses actually used in classification; and the correspondence between categories, classes, and subclasses.

The DISPLAY processor accepts the file output by CLASSIFY and generates a line-printer map of the classified data, along with several performance tables. In the map, each subclass has a symbol associated with it. A threshold option is provided for the analyst to print no symbol (blank) for samples classified with a confidence level less than some specified threshold value.

Performance summaries are provided on subclass, class, and category levels for pixels within each classified field, training field, and test field which is input to the DISPLAY processor. The training or test field performance summaries may be obtained by fields and/or classes. The DISPLAY processor also provides optional output of a classification map (MAPUNT) in either Universal or LARSYS II/III format.

The data transformation (DATATR) processor allows the analyst to use the linear transformation matrix computed by SELECT to create a new image data file (TRFORM). Since the matrix is computed to extremize subclass separability, the k linear combinations out of n channels represented by the matrix produce better class contrast when the image is displayed on the PMIS/DAS. In addition, the best linear combination of the data can be used to enhance the image. The TRFORM file may be output in either Universal or LARSYS II/III format.

The NDHIST processor performs an n-dimensional histogram of areas on the MSS data file, for which the user wishes to create scatter plots. The fields may be histogrammed on a class, subclass, or per-field basis. A line-printer summary of the fields, the number of data vectors in each field, and the number of unique data vectors histogrammed is given.

Optionally, if a scatter plot of a classified or clustered area is requested, a classification or cluster image data file (MAPUNT) from the DISPLAY or ISOCLS processor must be input to NDHIST. If this option is exercised, the field or fields input to this processor and their order of input must be the same as those input to CLASSIFY or ISOCLS.

Information such as the field, the cluster or subclass number, the frequency of occurrence, and the color code for each histogrammed radiance vector is written on the n-dimensional histogram (NHSTUN) file.

- The SCTRPL processor reads the NHSTUN file, and a two-axis color-coded spectral plot (SCTRUN) is output in Universal format. The background for the plot may be black or white.

If more than two channels were histogrammed by the NDHIST processor, the data vector is reduced to two components by

$$y' = B\vec{x} + \vec{c} \qquad (2\text{-}4)$$

where

y' = transformed image

B = a matrix

$\vec{x}$ = a data vector

$\vec{c}$ = a bias vector

The location on the scatter plot for each vector in the NHSTUN file is determined by its radiance values (if only two channels were histogrammed) or by two linear combinations of radiance values (if more than two channels were histogrammed).

The color for a pixel is assigned by

- Original radiance values
- Mean vector of the cluster or subclass to which the pixel was assigned during clustering or classification
- Mean vector of the test or training field from which the pixel was extracced
- User-defined colors
- Color extraction from a different pass when using multi-temporal Landsat data

Optionally, for pixel color assignment, the SAVTAP file created by the STAT or ISOCLS processor or card images containing color codes may be input.

Optionally, a line-printer pixel frequency or log of pixel frequency (base 2) plot is given. The plot is printed with up to 16 different symbols.

There are support processors to assist the analyst. DAMRG performs channel or spatial merger of image data. GTTCN and GTDDM accomplish the labeling of pixels and dots (selected pixels) on the basis of ground truth files.

Figure 2-1 is a diagram showing the principal processing options and paths in EOD-LARSYS.

2.2 OPERATION

The processors described here can be executed in two modes, batch or interactive. Both can be run from the direct-line terminals of the computer facilities in building 17, or anywhere that telephone contact can be established with the computer at Purdue. At present, this includes two terminals in the Lockheed buildings in Nassau Bay.

12

For batch or interactive operations, decks of cards (or files of card images) need to be prepared as shown in this document. From a direct-line terminal, an analyst issues instructions to read decks of cards, using the card reader.

Use of the editor to prepare files of control card images from the terminal is rather more complex, but it is much to be preferred.

The sections that follow give the practical steps for using EOD-LARSYS as implemented at Purdue LARS. Sections 2.2.1 and 2.2.2 tell how to log on and off the system. Section 2.2.4 describes the use of the editor to create, modify, and erase card image files. Section 2.2.5 tells how to run a program. These processes are also traced in figure 2-2.

2.2.1 TO START: LOG ON TO THE SYSTEM

To operate in either batch or interactive mode, the user must activate a terminal and establish telephone contact between terminal and computer. For the two dial-up terminals at Lockheed, make certain the terminal and coupler are turned on. Then dial the number of the multiplexer in building 17. When a high-pitched computer noise is heard, insert the telephone in the coupler.

From this point on, instructions for the dial-up terminals are identical to instructions for the direct-line terminals.

To get the attention of the system, depress the N key, then the carriage return (CR). Appearance of the carat (>) signifies that you are in control program (CP) mode. Now log on by typing

LOGON XXXXXX YYY

where XXXXXX is a user identification number and YYY is the password. If this information is acceptable, the system will ask for your name. If not, you will need to make a correction or simply try again. The system will then type a log-on message (LOGMSG), giving certain information regarding system operation.

At this point, if you plan to run interactively, you should type

DEF STOR 1M

This will allocate 1 megabyte of virtual memory for the job you are about to run. Without this, the program will not be able to operate for lack of space in the computer. (For batch operations, storage has been previously defined.) Now type

IPL LARSYSP1 (CR)

The system will respond with

EOD LARSYS READY

signifying that you have successfully logged on to the system and are ready to begin. Control now lies with the IBM CMS 3031 and with the JSC disk attached. To start the prompter, type

EODLARSYS

This calls the EOD-LARSYS program. From here, follow instructions in section 2.2.4 or 2.2.5. The first part of figure 2-2 shows this entire sequence, together with all responses by the system.

2.2.2 TO STOP: LOG OFF THE SYSTEM

To stop using the system, the user must log off. To do this, depress the BREAK key, and then type

LOG (CR)

In batch processing, this is normally done after the terminal displays the READY message (if done earlier, the job may be lost). In interactive operations, it can be done at any time.

14

2.2.3 CARDS OR CARD IMAGES

EOD-LARSYS can be run either with physical cards, submitted to a card reader, or as files of card images. These cards and card images are the control cards described in this document. The physical cards must be used with the card reader adjacent to the direct-line terminals in building 17. Card images can be used from any location.

When physical cards are used, the foregoing instructions are sufficient. The cards are fed to the card reader when indicated by the program. Output must be directed to HOUSTON. When cards are used, operations are normally run as batch, usually overnight. Figure 2-2 gives a complete set of computer messages and responses for such a job.

If physical cards are not to be used, the user must know how to create, modify, and destroy card image files. A file is a collection of data stored on disk or tape. A file is identified by three attributes — file name, type, and mode. The filename is specified by up to eight characters. Files to be read by Fortran programs should have filetypes of the form

$$FT_n\ F_m$$

where n is the Fortran unit number (00-99) and m the relative file number (000-999). If there is only one file for a particular unit, the filetype is arbitrary; e.g., CC. Filemode, for a file on the user's permanent disk, is An, where n is defaulted to 1.

To use EOD-LARSYS in the interactive mode, the user should create a control card file on disk. This is done using the editor, as described in the next section.

2.2.4 USING THE EDITOR TO CREATE, MODIFY, AND ERASE FILES

After typing EODLARSYS, you may edit a file by responding EDIT to the next prompt. (You may also call EDIT directly from CMS 370.) EDIT here means that you can create a new file, or modify an existing one, or delete one. Several options are described below.

To create a new file, type I (CR), for input. At this point you can type out the card images, much as though you were preparing cards on a keypunch device. After furnishing the name of the file (such as BARLEY CC), you might type the following:

```
FORMAT      1                          (CR)
$ISOCLS     INPUT/UNIT=11,FILE=1       (CR)
DATA                                   (CR)
  :
$EXIT                                  (CR)
                                       (CR)
```

To stop furnishing text, type two carriage returns in a row.

To change an existing file, you go into the EDIT mode in the same way. You can print out the entire file on the console by typing

```
TOP
T*       (For TYPE *)
TOP
```

The last TOP returns you to the top of the file.

If you type a wrong character, you may erase it by typing @. For example, to finish typing CAROLYN after you started with CARIL, you will have the following line of text on your screen:

CARIL@@OLYN

Notice that the BACKSPACE key will not function properly for this use. To erase the whole line, type [(CR) at any time.

16

The simplest way to locate a line to be changed, or a place to insert more lines, is by typing

N (CR)

until that line appears. N stands for NEXT. At this point you might insert more lines by typing

I (CR)

for input and then the lines to be inserted. Or, you might change characters by typing

C /XXX/YY/

where XXX is the incorrect string of characters and YY is the string to replace it. You might also simply delete the line by typing

DEL

for deletion.

To erase a file, enter

ERASE (filename) (filetype) (filemode) (CR)

At the end of an editing session, you should type FILE to store the modified file. If you neglect to do this, your changes will be lost.

More details on editing can be found in the "IBM Virtual Machine Facility 1370: CMS Command and Macro Reference."

2.2.5 TO RUN A PROGRAM

Before running a program, you must define the complete set of EOD-LARSYS control cards. You can punch the cards physically, or (preferably) you can prepare a file of card images as just desc.ibed.

You may execute in two ways: interactive or batch. The interactive way allows you to sit at the terminal, run a program, and receive the results on the terminal (or, in building 17, on the line printer). In batch, you submit a job and receive the results later on the line printer.

To run a program, type the location of the cards or card images when prompted (READER or DISK). If on disk, the system will ask for the filename. Although the name will appear with "CC" after it, you will not need to type more than the name.

The system will then ask if you wish to run interactively or batch.

The system will ask where to furnish the printer output. You will reply HOUSTON to this question if results are to be written to the line printers in building 17, the normal case. In certain cases it may be reasonable to write to the terminal screen; write TERMINAL for this case.

Type in the data tape number when requested. The tapes must be physically located in "Flexlab 2," the computer room at LARS.

The system will then ask whether to save various intermediate results on file or tape. Since the questions are asked even when the intermediate results will not be produced, the normal response is N (CR). Respond as appropriate.

2.2.6 NOTES ON PROCEDURES FOR MINISCAN

After you have entered the IPL LARSYSP1 command, you may enter EODSCAN instead of EODLARSYS. This command will bring in a module that can be used to scan the control card file for possible additional options. These new options are as follows:

1. Whenever the DATA card occurs in the control card image file, you may choose to use one of three new formats:

   a. If the data are known to the LARS data base by segment/acquisition, you may use one of the following formats:

   ```
   DATA      UNIT=VV, SEG=XXXX, ACQ=YYY
   DATA      INPUT/UNIT=VV, SEG=XXXX, ACQ=YYY
   ```

   SEG may be abbreviated S. ACQ may be abbreviated A. The three keys may be in any order.

   b. For any data set on tape (for example, a merged data set), you may use one of the following formats:

   ```
   DATA      UNIT=VV, FILE=XX, TAPE=YYY
   DATA      INPUT/UNIT=VV, FILE=XX, TAPE=YYY
   ```

   TAPE may be abbreviated T, and the three keys may be in any order.

   c. If the data set desired is on a disk file, you may use one of the following formats:

   ```
   DATA      UNIT=XX,NAME='FILENAME FILETYPE FILEMODE'
   DATA      INPUT/UNIT=XX,NAME='FILENAME FILETYPE FILEMODE'
   ```

   NAME may be abbreviated N. The two keys may be in either order. The single quotation marks are required around the name DATA.

   d. The currently existing options are still acceptable.

2. Whenever any card that uses one of the following formats

   ```
   DATA      UNIT=, FILE=
   DATA      INPUT/UNIT=, FILE=
   DATA      OUTPUT/UNIT=, FILE=
   ```

   is in the control card file, then you may use only format type 1b or 1c.

3. The READ control card image in processor GTTCN may be coded using format 1a, 1b, or 1c.

If you choose to use one of the new formats for your DATA card type, then when the prompting EXEC asks if the job requires an MSS tape, you should answer NO.

All other options currently available under EODLARSYS are still valid with EODSCAN.

2.2.7 TROUBLESHOOTING

It is not feasible to furnish responses to all problems that may occur. The following are the principal ones that have occurred in the past.

Sometimes, line noise may cause a small jumble of letters to be printed, and the system will return to CP status (shown by the carat). When this happens, type B (CR) to begin again.

If the diagnostic READ ERROR appears, merely touch the CR key and the system may recover.

If the system is not responding properly to typed commands, the telephone connection may have been broken. When this happens at remote terminals, the carrier light will not be lit. When this happens, start over by logging on. If you were editing, check to see the status of your files. If more than 15 minutes elapses, you may lose all edits.

30

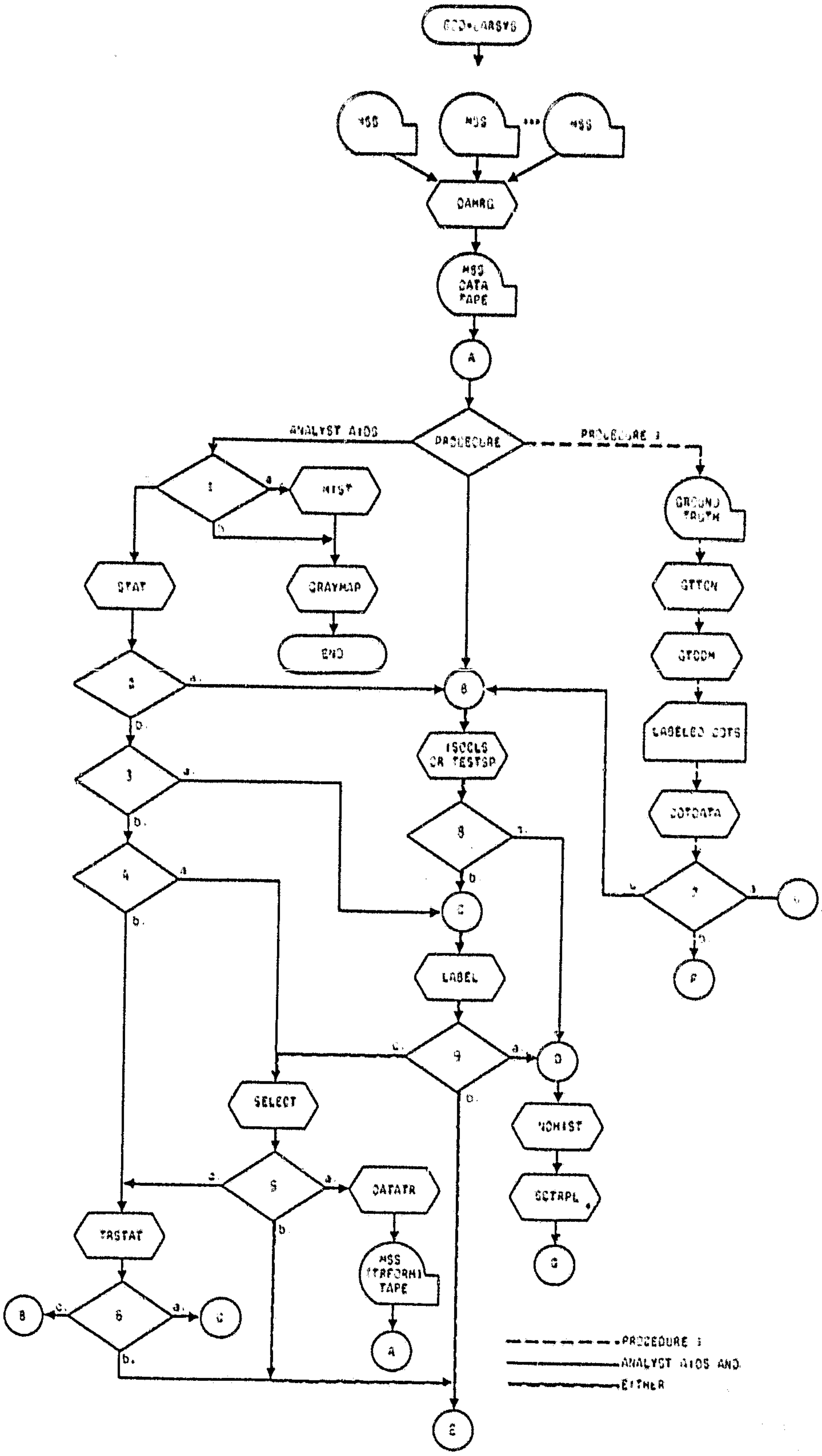
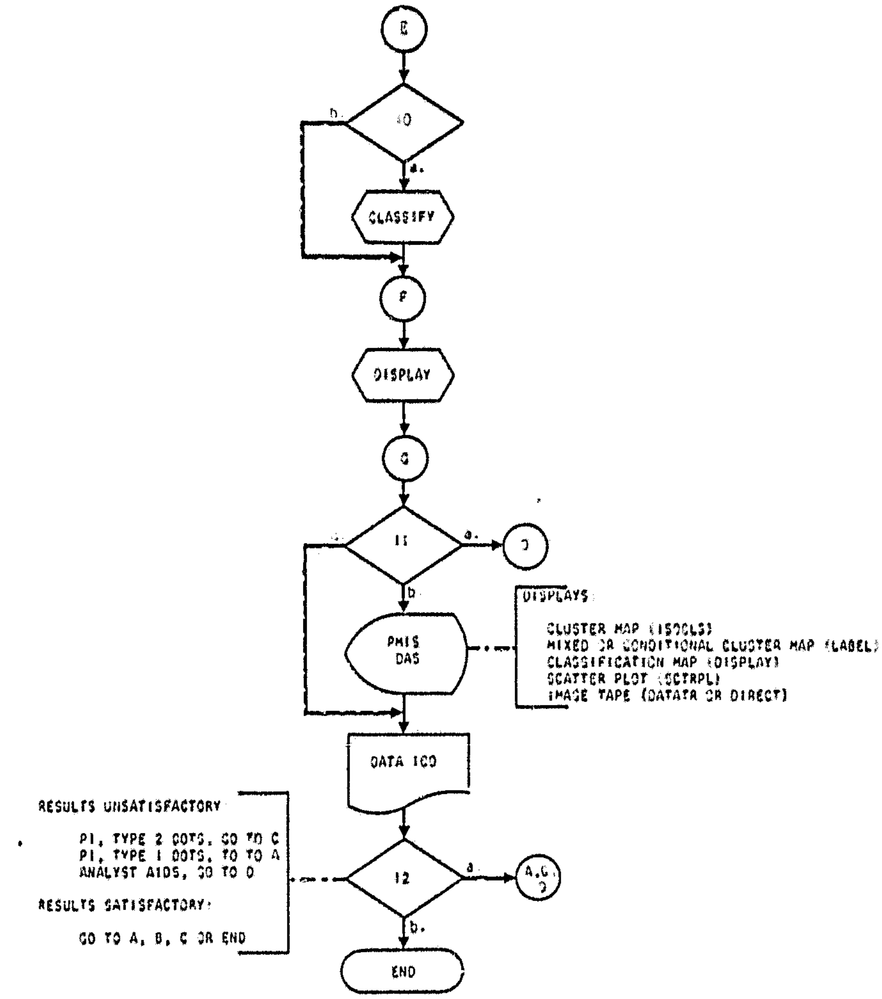


Figure 2-1.— Major processing paths in EOD-LARSYS.



Key to decision points 1 through 12

1 a. Compute histograms and print pictorial gray-scale map of data from any channel, using HIST and GRAYMAP.
  b. Print pictorial gray-scale map only, using GRAYMAP.
  c. Compute training field statistics and write SAVTAP file, using STAT.

2 a. Group training fields into classes, using ISOCLS.
  b. Omit clustering.

3 a. (Re)label training field statistics.
  b. Omit labeling.

4 a. Determine subset or linear combination of channels that maximizes subclass separability, using SELECT.
  b. Transform training field statistics, using TRSTAT.

5 a. Create new image data tape, applying linear transformation matrix computed by SELECT, using DATATR.
  b. Perform supervised classification of image using CLASSIFY.

6 a. Relabel previously labeled statistics file SAVTAP, using LABEL.
  b. Proceed to classification, using CLASSIFY.
  c. Use statistics file SAVTAP to provide starting cluster mean vectors for ISOCLS.

7 a. Relabel dots in dot data file, using LABEL.
  b. Display dots, using DISPLAY.
  c. Cluster image, using ISOCLS.

8 a. Compute n-dimensional histogram of selected data areas, using NDHIST, and scatter plot, using SCTRPL.
  b. Proceed to labeling, using LABEL.

9 a. Compute n-dimensional histogram of selected data areas, using NDHIST, and scatter plot, using SCTRPL.
  b. Proceed to classification, using CLASSIFY.
  c. Evaluate discriminatory capability of channels, using SELECT.

10 a. Proceed to classification, using CLASSIFY.
  b. Proceed to classification summary, using DISPLAY.

11 a. Compute n-dimensional histogram of selected data areas, using NDHIST, and scatter plot, using SCTRPL.
  b. Display image on display station and print

12 a. Results unsatisfactory (see annotation on flowchart).
  b. Results satisfactory (see annotation on flowchart).

Figure 2-1.– Concluded.

32

LOGON

```
>LOGON JSC700 SUE

ENTER NAME:  >GIDDINGS

LOGMSG - 08:03:47 EST SATURDAY 05/05/79
* NEXT SCHEDULED SHUTDOWN IS SATURDAY 5/5 AT 17:00
LOGON AT 10:29:41 EST SATURDAY 05/05/79

>IPL LARSYSP1

EOD LARSYS READY,
DEV 19C DOES NOT EXIST
Y (19C) R/O
R;  T=0.17/0.27 10:30:07

>EODLARSYS
```

Question 1. ARE THE EOD-LARSYS INPUT CARDS IN THE CARD READER, ON DISK, OR DO YOU WISH TO CREATE OR MODIFY THEM, OR DO YOU WISH TO GET A STANDARD SET FROM THE EOD-LARSYS SYSTEM DISK? (READER, DISK, EDIT, OR GET)

Possible responses

A. READER (Default)
Implies EOD-LARSYS input cards are in your card reader.

B. DISK <>FILENAME (Will ask for filename, if not given.)
WHAT IS THE FILENAME OF THE EOD-LARSYS INPUT CARDS?
FILENAME
Implies EOD-LARSYS input cards are in disk file "FILENAME CC."

C. EDIT <>FILENAME
WHAT IS THE FILENAME OF THE EOD-LARSYS INPUT CARDS THAT YOU WISH TO CREATE OR MODIFY?
FILENAME
EDIT
    EDIT commands
EDIT
FILE
Implies EOD-LARSYS input cards will be edited on disk, or will be created on the disk via edit.

D. GET <>FILENAME (Will ask for filename, if not given.)

Figure 2-2.- EOD-LARSYS queries and responses.

Possible errors

E00001 FILE 'FILENAME' CC A NOT FOUND

This error occurs when question 1 is answered by "DISK FILENAME" and the specified file does not exist. (If this error occurs, question 1 is reissued.)

Question 2. DO YOU WISH TO RUN INTERACTIVELY AT THE TERMINAL OR HAVE YOUR EOD-LARSYS JOB SENT TO A BATCH MACHINE? (INTER OR BATCH)

Possible responses

A. INTER (Default)
Implies you wish to run interactively.

B. BATCH
Implies EOD-LARSYS input cards will be sent to a batch machine for processing.

NOTE: If a batch job is to be run and the user specified "READER" for question 1, the following error may occur.

E00002 THERE WERE NO CARDS IN YOUR CARD READER.
READ THEM IN AND THEN TYPE 'READY'
(Default response here is READY.)

Question 3. AT WHICH SITE DO YOU WISH TO RECEIVE THE PRINTER OUTPUT (AND OPTIONALLY PRINT STATUS)?

Possible responses

SITE <> HOLD/NOHOLD; (Default site is where command;
e.g., HOUSTON HOLD default print status is HOLD.)

Possible errors

E0003A 'SITE' IS NOT A VALID PRINT SITE
(If this error occurs, question 3 is reissued.)

E0003B 'STATUS' IS NOT A VALID HOLD STATUS
PLEASE ENTER HOLD OR NOHOLD

Figure 2-2.- Continued.

24

Question 4. WILL YOUR EOD-LARSYS JOB BE USING A MSS DATA TAPE? (YES OR NO)

Possible responses

A. YES (Default)
The above question will not be asked if the user is going to run batch since a tape is required for batch. However, the following line will be typed for either interactive or batch mode.

TYPE IN MSS DATA TAPE NUMBER

Possible response

NNNN;
e.g., 1267

B. NO

Implies a MSS data tape is not needed for this job.

Question 5. DO YOU WISH TO SAVE ANY INTERMEDIATE RESULTS PRODUCED BY EOD-LARSYS, OR USE ANY PREVIOUSLY SAVED ONES? (YES OR NO)

Possible responses

A. NO (Default)
Implies intermediate results will be written to a temporary disk and will not be available after the job finishes execution.

B. YES
The following question will then be asked about each of the intermediate results.

CLASSIFICATION MAP (NO, SAVE, OR USE)

Possible responses

A. NO (Default)
Implies you do not wish to save the classification map or use a previously generated map.

B. SAVE <>FILENAME
Implies you want the classification map stored on your permanent disk.

Figure 2-2.- Continued.

C. USE <>FILENAME
Implies you want a classification map that is stored on your permanent disk to be used.

N-DIMENSIONAL HISTOGRAM (NO, SAVE, OR USE)

TRAINING STATISTICS (NO, SAVE, OR USE)

HISTOGRAM (NO, SAVE, OR USE)

DOT DATA (NO, SAVE, OR USE)

STATISTICS (NO, SAVE, OR USE)

TRANSFORMATION MATRIX (NO, SAVE, OR USE)

THE FOLLOWING INTERMEDIATE RESULTS MAY BE SAVED ON OR USED FROM EITHER YOUR PERMANENT DISK OR A TAPE.

SCATTER PLOT (NO, SAVE, OR USE)

Possible responses

A. NO (Default)

B. SAVE <>FILENAME
Implies the scatter plot will be saved on your permanent disk.

C. SAVE <>NNNN
Implies the scatter plot will be saved on tape NNNN file MM, where the default file number is 1.

D. USE <>FILENAME
Implies the previously saved scatter plot is on your permanent disk.

E. USE <>NNNN
Implies the previously saved scatter plot is on tape NNNN file MM, where the default file number is 1.

TRANSFORMED MSS DATA (NO, SAVE, OR USE)

CLUSTER MAP (NO, SAVE, OR USE)

Figure 2-2.- Continued.

26

If the user is going to run his or her job batch, the following message is printed and only the following questions will be asked since intermediate results may only be saved on or used from tape.

THE FOLLOWING INTERMEDIATE RESULTS MAY BE USED FROM AND/OR SAVED ON TAPE.

SCATTER PLOT (NO, SAVE, OR USE)

TRANSFORMED MSS DATA (NO, SAVE, OR USE)

CLUSTER MAP (NO, SAVE, OR USE)

The only valid responses for batch processing are A, C, and E listed under "SCATTER PLOT" above.

Possible errors

E0005A TYPE IS NOT PERMITTED FOR THIS DATA SET

Possible errors for USE <>FILENAME

E0005B 'FILENAME' DOES NOT EXIST ON YOUR A-DISK

Possible errors for SAVE <>FILENAME

E0005C SAVE IS NOT A VALID OPTION SINCE YOU DO NOT HAVE AN A-DISK THAT CAN BE WRITTEN ON.

E0005D THERE IS NOT ENOUGH SPACE ON YOUR A-DISK TO SAVE 'FILENAME'. TYPE A SERIES OF CMS COMMANDS TO MAKE ROOM. TYPE READY WHEN DONE OR TYPE SKIP IF YOU DO NOT WANT TO MAKE ROOM.

Since files are not saved until after the EOD-LARSYS job is completed, error E0005D will not appear until then.

Question 6. DO YOU WANT TO RUN ANOTHER JOB? (YES OR NO)

Possible responses

A. YES (Default)
Then question 1 is asked again and so on.

B. NO
An exit to CMS is taken.

Figure 2-2.- Continued.

After the job has been sent to the batch machine, the following message is typed.

YOUR JOB HAS BEEN SENT TO THE ___________ BATCH MACHINE

Then, question 6 is asked of the user, and so on.

Figure 2-2.- Concluded.

3. INPUT

3.1 IMAGE TAPES

Every processor except SELECT, TRSTAT, SCTRPL, and LABEL uses an MSS data file (DATAPE). The assignment defaults to logical unit 11, but the user may assign any unit available by input of the DATA control card. For details, see the unit assignment chart in section 4 and the control card section for each processor. The file may be in LARSYS II/III or Universal format (defined in appendixes B and C, respectively) or in certain Landsat formats (defined in references 5 and 6).

The control card DATA allows the user to communicate the number of the MSS data file to be processed and the logical unit assignment. This is optional input to every processor that requires MSS data. The first file of the tape will be processed unless otherwise specified by the DATA control card. In executing the same and/or different processors consecutively, the DATA control card may be input only to the first processor executed if the same file and logical unit are to be used throughout the execution. For example,

```
$HIST
DATA       UNIT=11,FILE=2
(Other control cards)
*END
(Field definition)
$END
$GRAYMAP
CHANNELS 5,6
*END
(Field definition)
$END
```

[File 2 of the MSS data tape (DATAPE) assigned to unit 11 is processed by GRAYMAP as well as HIST.]

3.2 CARDS AND CARD IMAGES

Card input to the system must be one of the types discussed below. It should be noted that card image files are normally used in remote processing applications. In the discussion following, "card image" should be understood for "card."

3.2.1 PROCESSOR CARDS

A processor card identifies the processor that is to be executed. The system monitor routine calls the appropriate processor, which initiates the loading of all routines used by the processor. The processor card is always a $ symbol followed by the processor name and must always be punched left justified beginning in column 1. No blanks are allowed. The $ symbol and the first five characters are the unique processor identification used by the system monitor routine.

Below is a list of all processor cards recognized by the system, along with the section in which each processor is described.

$HIST	Section 6
$GRAYMAP	Section 7
$STAT	Section 8
$ISOCLS	Section 9
$SELECT	Section 10
$CLASSIFY	Section 11
$DISPLAY	Section 12
$DATATR	Section 13
$TRSTAT	Section 14
$NDHIST	Section 15
$SCTRPL	Section 16
$DOTDATA	Section 17
$LABEL	Section 18
$DAMRG	Section 20
$GTDDM	Section 21

30

$GTTCN	Section 22
$TESTSP	Section 23
$EXIT	Execution terminates when this card is encountered.

3.2.2 CONTROL CARDS

Each processor has its own set of control cards which allow the user to exercise various options in the particular processor or to change the default value assigned to certain parameters in the system. These cards must immediately follow a processor card. The control cards are identified by a keyword in columns 1 through 10 of the card. Only the first four characters are used for testing. In columns 11 through 72, the parameter values or options are indicated. These columns are free form, blanks are ignored (unless of legitimate parameter value), and multiparameter values or options are separated by commas. Columns 73 through 80 of the card are not used. With the exceptions of the FORMAT, *END, $END, and in some cases the STATFILE cards, control cards may occur in any order. (The STATFILE control card exception is noted in the section for the appropriate processor.) If the list of parameter values for a given keyword is too long for one card, the remaining values can be input on another card with the same keyword. (The continuation of a CATEGORY control card is slightly different; see section 11, table 11-1.) In every processor, the *END control card indicates the end of a set of control cards, and the $END card indicates the end of the field definition card input. The FORMAT card defines the format of the MSS data file by a 1 or a U (Universal), a 2 or an L (LARSYS II/III), a 3 or a B (Landsat bulk), or a 4 or an E (EROS*) in column 11 (or subsequent columns). This card precedes all others in

*Earth Resources Observation Systems Data Center at Sioux Falls, South Dakota.

a job setup. The user should ensure that all files written in the run are consistent in format.

Ancillary control cards, common to all processors, can be used to control the titles on printout sheets. As shown in table 3-1, defaults present standard titles and the current date. An optional COMMENT card can be used to furnish further identification on output sheets.

3.2.3 CLASS, SUBCLASS, AND FIELD DEFINITIONS

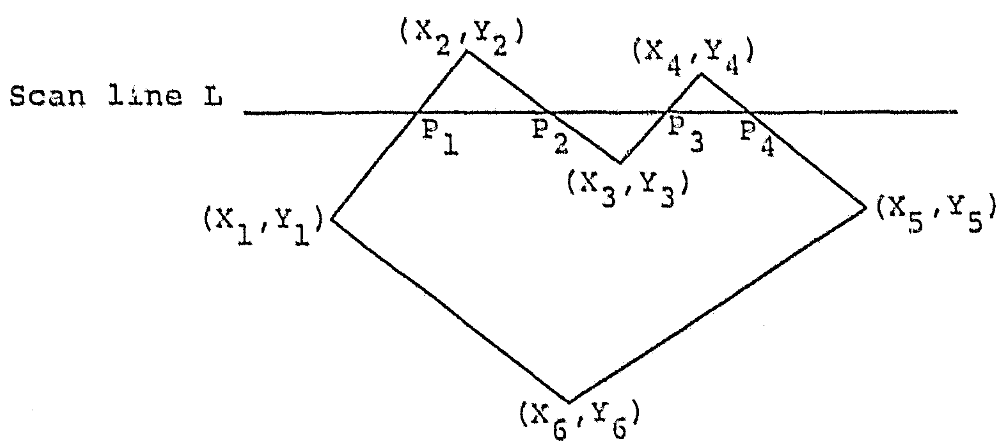
A field is a specific block of data to be extracted from the input MSS data file and processed. It is defined by a sample increment, a line increment, and from 1 to 10 vertices. Optionally, the user may associate a name with each field. The alphanumeric field description is located in columns 1 through 6. In columns 11 through 72, sample and line increments are separated by a comma and enclosed together in parentheses. A comma separates the increments and each of the following vertices. The vertices must be arranged in clockwise order. Sample and line numbers which describe a vertex are separated by a comma and enclosed together in parentheses. The sample number must be given first for each vertex. More than one card may be used to describe a field. An asterisk occurring after a vertex indicates a continuation card is to be read beginning in column 11. A vertex must be completed on a card and cannot be split between two cards. The numbers which describe the increments and vertices must be integers.

It is the user's responsibility to ascertain that all defined fields are within the bounds of the MSS image. In determining which pixels belong to a particular field, the EOD-LARSYS examines the pixel intercepts of each scan line with each side of the field. The pixel intercept X, with the scan line L and the side defined by vertices (X_1,Y_1) and (X_2,Y_2), is calculated by the equation

$$X = \frac{(L - Y_1)(X_2 - X_1)}{(Y_2 - Y_1)} + X_1 \qquad (3\text{-}1)$$

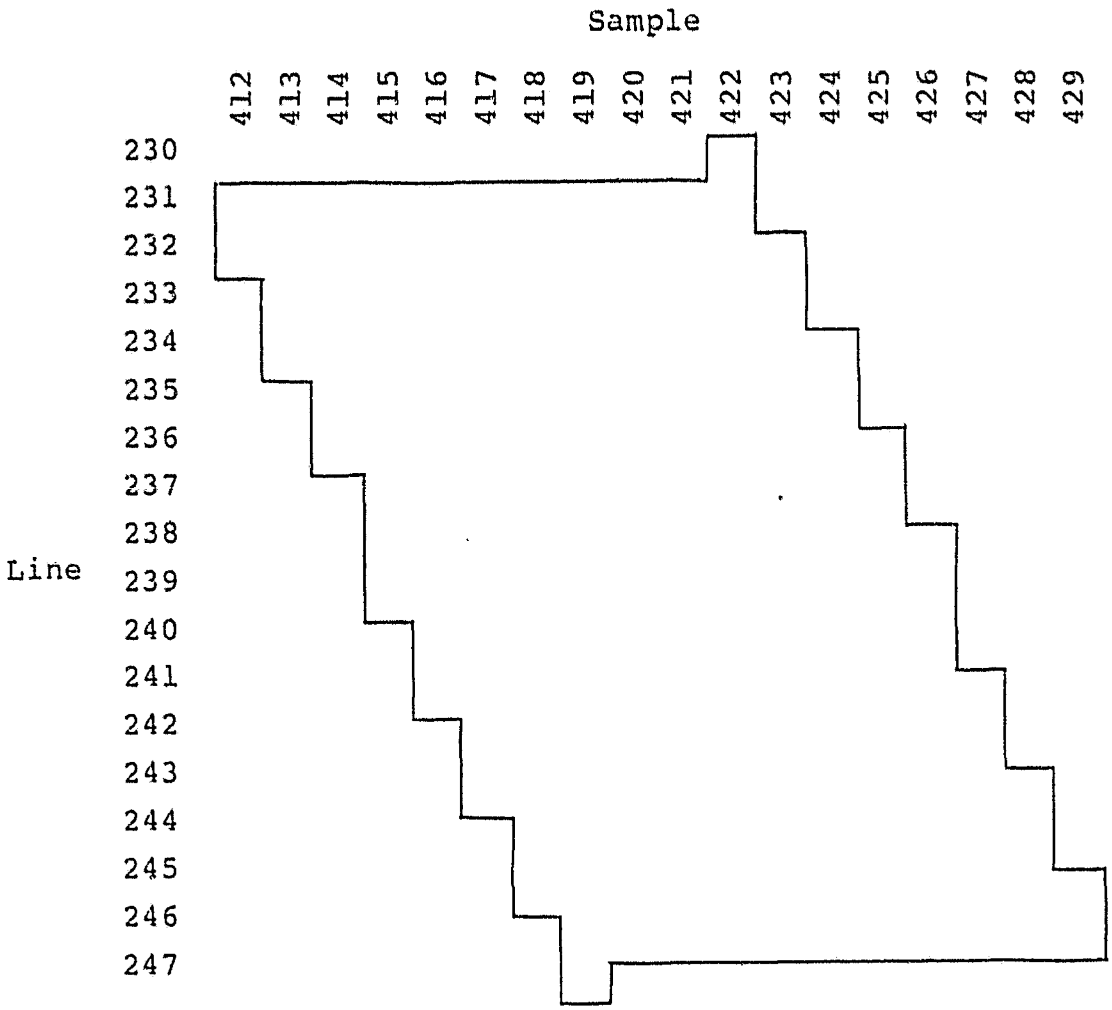
The value of X is computed as a floating-point number; however, the actual pixel intercept must be an integer number. Therefore, if the fractional part of X is greater than one-half, the pixel intercept is the next higher integer number. If the fractional part of X is less than one-half, the pixel intercept will be the next lower integer number. When the fractional part of X is exactly one-half, the integer pixel intercept depends on the direction of movement from the point (X_1,Y_1) to (X_2,Y_2). If Y_1 is less than Y_2, the pixel intercept is the next higher integer. If Y_1 is greater than Y_2, the pixel intercept is the next lower integer number.

After all pixel intercepts for a given scan line have been determined, the intercepts are taken in pairs and all pixels between and including the pair of intercepts are included in the field. In the following example for scan line L, all pixels between and including P_1 and P_2 are included, and all pixels between and including P_3 and P_4 are included.



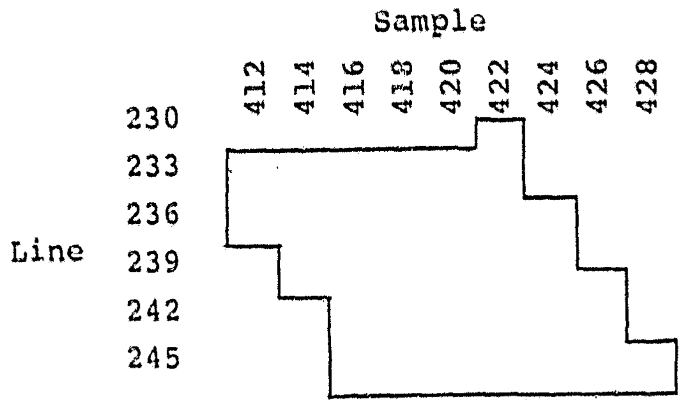
The following three examples describe field definition cards and the fields they define. In example 1, the sample and line increments are equal to 1 for field F1, and there are four vertices.

F1 (1,1),(412,231),(422,230),(429,246),(419,147)



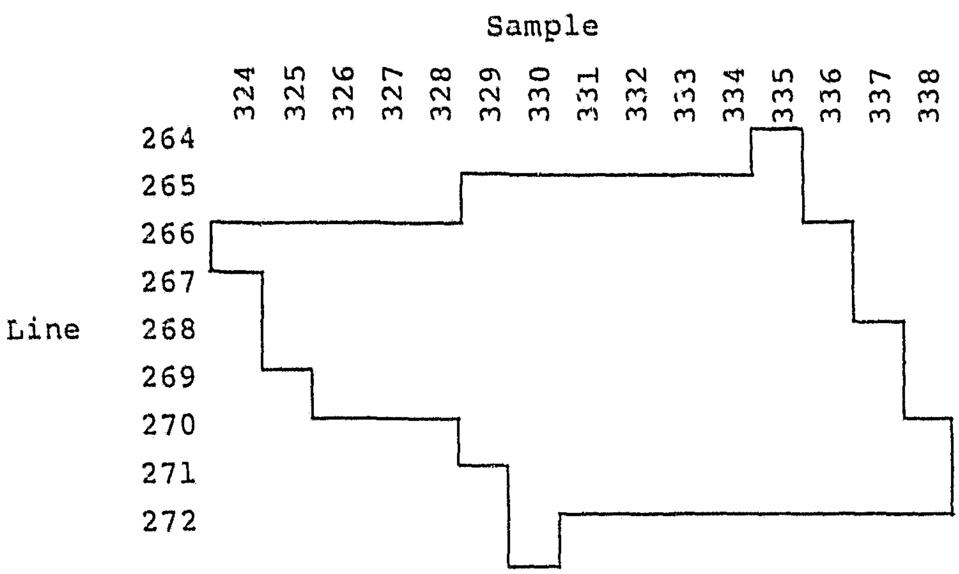
In example 2, the field F2 has the same vertices as F1; however, the sample increment is 2 and the line increment is 3.

F2 (2,3),(412,231),(422,230),(429,246),(419,247)



In example 3, the sample and line increments for field W187 are equal to 1, and there are six vertices.

W187 (1,1),(324,266),(355,264),(338,271),
(330,272),(329,269),(326,269)



Except for SELECT, TRSTAT, and scatter plot (SCTRPL), every processor accepts the input of field definition cards. Field definitions are always input between the *END and $END control cards.

The exact way of using these cards differs from processor to processor, as is shown in table 3-2. In the STAT and ISOCLS processors, fields must be associated with a class or subclass name. In the DISPLAY processor, fields may be test fields or designated fields. In the NDHIST processor, fields are associated with class or subclass and may be test, training, or any user-defined field.

The fields defined in STAT and ISOCLS are called training fields, and the data within these fields are used for computing statistics. Training fields are grouped into subclasses, and subclasses are further grouped into classes, using the STAT processor. In ISOCLS, training fields are grouped into classes, and the clustering procedure breaks the class data into subclasses (clusters). To allow for these groupings, cards bearing a class name and a subclass name are necessary.

A class name card has the keyword CLASS beginning in column 1 and the alphanumeric name of the class left justified in columns 11 through 14 of the card. Blanks should not be embedded in the class or subclass names.

A subclass name card has the keyword SUBCLASS beginning in column 1 and the alphanumeric name of the class left justified in columns 11 through 14 of the card.

In STAT, a CLASS card must immediately follow the *END control card. The CLASS card is followed by one or more SUBCLASS cards, each of which must be followed by one or more field definition cards. See the example for STAT (section 8.4.4).

In ISOCLS, a CLASS card must immediately follow the *END control card. The CLASS card is immediately followed by one or more field definition cards. The data from the fields associated with a given class name are clustered as one data set. The class is broken into subclasses (clusters) which do not have field boundaries. So, even though statistics are computed on a subclass level, training fields cannot be associated with subclasses in ISOCLS. See the example for ISOCLS (section 9.5.4).

In DISPLAY, test fields (if input) must be identified by a previously defined class or subclass name. When associated with classes, a CLASS card should immediately follow the *END control card. Test fields for that class should follow immediately. When associated with subclasses, a SUBCLASS card should immediately follow the *END card, followed by the test fields for that subclass.

Designated fields are the other type of field input to DISPLAY. Fields may be DU or DO. For input of designated fields, a card with the keyword DESIGNATE beginning in column 1 and the keyword OTHER or UNIDENTIFIABLE beginning in column 11 must precede the field definition cards. See section 12.4.4 for sample input of test and designated fields.

3.2.4 SPECIAL SYSTEM FILES

The card image files described in this section are special files normally output from one processor to be used at some future time as input to another processor. However, if the user can obtain all of the information needed for any of the special card image files from some other source, such information may be input directly to the processor if the formats described in this section are followed.

These files are always referred to in the job setup by the control cards for the particular processor. The first card image of each file acts as a keyword which initiates the input of the file. It is not necessary to input the same file to more than one processor in the same run.

3.2.4.1 Module STAT File

The module STAT file is optional output from the STAT, ISOCLS, and TRSTAT processors. It contains the statistics (mean vectors and covariance matrices) for all the subclasses input to STAT, the statistics for the clusters computed by ISOCLS, or the transformed statistics for all subclasses or clusters input to TRSTAT. These statistics are needed in the computation of the probability density functions in CLASSIFY and the computation of separability measures in SELECT.

This file also contains all the training field boundaries, the class and subclass numbers to which the training fields belong, the class and subclass names, the number of subclasses in each class, and the number of points in each subclass or cluster. By defining the required training fields in STAT, the user has absolute control over the data samples that will define a subclass from the MSS data file. Every data sample occurring in any one of the training fields defined by a particular subclass is used in computing the mean vector and covariance matrix for that subclass.

In the clustering processor ISOCLS, the user has no control over the specific samples that comprise a cluster. The processor determines which data samples are used in computing the mean vector and covariance matrix for each cluster. Because it is desirable to use these cluster statistics in other processors, the ISOCLS processor creates a file in the same format as that used in the STAT processor. The file may be punched if desired. Training fields are associated with classes rather than subclasses. Clusters are given a six-character name. The first three characters are the first three characters of the class name associated with the cluster, and the last three characters are digits. The digits for the subclasses are in sequential order.

When the module STAT file is input to the CLASSIFY or SELECT processor, the user may request subsets of the statistics to be used for classification or channel selection via the SUBCLASS and CHANNELS control cards in both processors. Subclasses are numbered as they were input to STAT, and clusters are numbered as they were created in ISOCLS. The channels are numbered as they occur on the MSS data file. To select a subset of the statistics in the module STAT file, the user should indicate by number the subclasses and/or channels he or she wishes to use. (Unless the user has previous knowledge of the number of clusters in the module STAT file, he or she cannot accurately select a subset of the clusters when executing ISOCLS in conjunction with another processor.)

The first card in the module STAT file acts as a control card, with the keyword MODULE starting the input of the remainder of the file. The entire file is composed of the card image types listed below. All integers should be right justified in the specified field, and alphanumeric characters should be left justified in the specified field.

- Card type 1 — Keyword MODULE in columns 1 through 6.

- Card type 2 — Number of classes, subclasses, channels, fields, and vertices for training fields.

Columns	Format	Definition
7-10	I4	Number of training classes from STAT or ISOCLS
19-20	I2	Number of training subclasses from STAT (clusters from ISOCLS)

Columns	Format	Definition
29-30	I2	Number of channels used in computing statistics
38-40	I3	Number of training fields input to STAT or ISOCLS
49-52	I4	Number of vertices in all the training fields

- Card type 3 — Actual channels used in computing statistics.

Columns	Format	Definition
11-12	I2	Channel 1
13-14	I2	Channel 2
15-16	I2	Channel 3
⋮	⋮	⋮
69-70	I2	Channel 30

- Card type 4 — Training field information: The first card of the set.* Names should not exceed four characters.

Columns	Format	Definition
1-6	A4	Field name
11-12	I2	Number of the class associated with this field

*Card types 4 and 5 define a training field. To complete the set of information for one training field, one card of type 4 and one or two cards of type 5 are required. The number of card sets is determined by the number of training fields.

40

Columns	Format	Definition
21-22	I2	Number of the subclass associated with this field input to STAT (Since ISOCLS associates fields with classes, ISOCLS dummies this information by setting it equal to 0.)
31-32	I2	Number of vertices for this field, including closure point

- Card type 5 — Vertices for the training field: Up to 10 vertices plus the closure point are allowable for each training field, 7 vertices per card with coordinates ordered (sample, line). The coordinates are listed in a clockwise manner, with the coordinate having the smallest sample number listed first.*

Columns	Format	Definition
11-15	I5	Sample number of first vertex
16-20	I5	Line number of first vertex
21-25	I5	Sample number of second vertex
26-30	I5	Line number of second vertex
⋮	⋮	⋮
76-80	I5	Line number of seventh vertex

- Card type 6 — Class names, nine names per card, left justified in field: The number of cards is determined by the number of classes.

Columns	Format	Definition
11-16	A4	Four-character class name for first class
19-24	A4	Four-character class name for second class

Columns	Format	Definition
27-32	A4	Four-character class name for third class
⋮	⋮	⋮
75-80	A4	Four-character class name for ninth class

- Card type 7 — Number of subclasses in each class, 24 per card: The number of cards is determined by the number of subclasses.

Columns	Format	Definition
9-10	I2	Number of subclasses in first class
12-13	I2	Number of subclasses in second class
15-16	I2	Number of subclasses in third class
⋮	⋮	⋮
78-79	I2	Number of subclasses in 24^{th} class

- Card type 8 — Subclass names, 10 per card: The number of cards is determined by the number of subclasses.

Columns	Format	Definition
9-14	A4	Four-character subclass name for first subclass
16-21	A4	Four-character subclass name for second subclass
23-28	A4	Four-character subclass name for third subclass
⋮	⋮	⋮
72-77	A4	Four-character subclass name for 10^{th} subclass

42

To complete the set of statistics for one subclass, the following three types of cards are grouped together. The number of sets of cards is determined by the number of subclasses.

- Card type 9 — Number of points in this subclass.

Columns	Format	Definition
13-20	I8	Number of points in this subclass

- Card type 10 — Mean vector for this subclass, five values per card: The number of cards is determined by the number of channels.

Columns	Format	Definition
6-20	E15.8	Mean for first channel for this subclass
21-35	E15.8	Mean for second channel for this subclass
36-50	E15.8	Mean for third channel for this subclass
51-65	E15.8	Mean for fourth channel for this subclass
66-80	E15.8	Mean for fifth channel for this subclass

- Card type 11 — Covariance matrix for this subclass: Only the lower triangular portion of the matrix is output; the number of values input for this matrix is equal to (number of channels) × (number of channels + 1)/2. Five values are written on each card image in the order indicated.

1						
2	3					
4	5	6				
7	8	9	10			
—	—	—	—	—		
—	—	—	—	—	—	
—	—	—	—	—	—	—

Columns	Format	Definition
6-20	E15.8	Element 1 of matrix
21-35	E15.8	Element 2 of matrix
36-50	E15.8	Element 3 of matrix
51-65	E15.8	Element 4 of matrix
66-80	E15.8	Element 5 of matrix

3.2.4.2 B-Matrix File

The B-matrix file is an optional output of the SELECT processor when the Davidon-Fletcher-Powell procedure is used. The file contains a transformation matrix which extremizes a given separability measure for the subclasses being used. The matrix is optimized using the Davidon-Fletcher-Powell procedure. The linear transformation of the original measurements can be used in the CLASSIFY, DATATR, TRSTAT, or SCTRPL processor to reduce the dimensionality of the data and/or statistics.

The B-matrix deck, or corresponding file, is an optional input to SELECT, CLASSIFY, and SCTRPL, and a required input to DATATR and TRSTAT. When input to SELECT, the matrix is used to evaluate a specific separability measure or it is used as a first guess for the Davidon-Fletcher-Powell procedure, depending on the user's request. When input to CLASSIFY, classification is performed using the linear transformation. When input to SCTRPL, the dimension of the data from the MSS data file is reduced to two linear combinations. The DATATR processor uses the matrix to create a new image file with the reduced dimensionality. The TRSTAT processor creates on SAVTAP a new file containing the transformed statistics.

The keyword B-MATRIX on a control card indicates that the B-matrix is being input. Since the matrix may be input on cards or from a disk file, the parameter CARDS or FILE must be specified on the same card in columns 11 through 72. The entire file is defined below by card types.

- Card type 1 — The keyword B-MATRIX in columns 1 through 10 and CARDS or FILE in columns 11 through 72 start the input of the file.

- Card type 2 — One card of this type.

Columns	Format	Definition
6-7	I2	Number of linear combinations
13-14	I2	Number of channels
18-80	30I2	The remainder of this card lists by number the channels for which the matrix was computed (e.g., columns 18 through 19, first channel, etc.), for a maximum of 30 channels right justified in the field.

- Card type 3 — The number of these cards is determined by the size of the matrix. The values are input by column as indicated below, five values per card.

B(k,n), k = linear combinations; n = channels

$$
\begin{array}{cccc}
1 & (k+1) & \cdots & [nk-(k-1)] \\
2 & (k+2) & \cdots & [nk-(k-2)] \\
3 & (k+3) & \cdots & [nk-(k-3)] \\
\vdots & \vdots & & \vdots \\
k & 2k & & nk
\end{array}
$$

Columns	Format	Definition
6-20	E15.8	Element 1 of matrix
21-35	E15.8	Element 2 of matrix
36-50	E15.8	Element 3 of matrix
51-65	E15.8	Element 4 of matrix
66-80	E15.8	Element 5 of matrix (Continued on next card)

3.2.4.3 Cluster Means File

The cluster means file is an optional input to the clustering processor ISOCLS. It may be used to begin the clustering process by estimating cluster centers (means). The means can be taken from the module STAT file (see section 3.2.4.1) created by TRSTAT, STAT, ISOCLS, or the user. Means may be input for up to 30 channels for each cluster center, and a subset of the channels to be used may be indicated on the CHANNELS control card.

The keyword MEANS in the control cards for ISOCLS indicates initial cluster means are being input. Since the means may be input on cards or from a disk file, the parameter CARDS or FILE must be specified on the same card in columns 11 through 72. If on cards, CARDS starts the input of the cluster means deck, which must immediately follow. The format for the entire file is indicated below.

- Card type 1 — Control card keyword MEANS is left justified in columns 1 through 5. The parameter CARDS in columns 11 through 72 starts the input of the card deck.

- Card type 2 — Number of clusters and channels.

116

Columns	Format	Definition
6-10	I5	Number o initial clusters for which means are provided
25-30	I5	Number of channels for which means are provided

- Card type 3 — Actual channels used in computation of means.

Columns	Format	Definition
6-7	I2	Channel 1
8-9	I2	Channel 2
10-11	I2	Channel 3
⋮	⋮	⋮
64-65	I2	Channel 30

- Card type 4 — Mean vectors for the initial clusters: These cards are in the same format as the mean cards (card type 10) in the module STAT file. The first mean for each cluster always begins on a new card. The number of cards depends on the number of channels and the number of clusters. Five values are placed on each card.

Columns	Format	Definition
6-20	E15.8	Mean for channel 1
21-35	E15.8	Mean for channel 2
36-50	E15.8	Mean for channel 3
51-65	E15.8	Mean for channel 4
66-80	E15.8	Mean for channel 5 (Continued on consecutive cards of the same format)

TABLE 3-1.- ANCILLARY CARDS FOR ALL PROCESSORS

Keyword	Parameter and default values	Function
HED1	Any 60 characters beginning in column 11 Default: LYNDON B. JOHNSON SPACE CENTER	Replaces first header line with the indicated characters in the parameter field.
HED2	Any 60 characters beginning in column 11 Default: HOUSTON, TEXAS	Replaces second header line with indicated characters in the parameter field.
DATE	Any 12 characters beginning in column 11 Default: Current date	Prints the indicated 12 characters in the right corner of the heading in place of the current date.
COMMENT	Any 60 characters beginning in column 11 Default: No comments printed	Prints a comment line using the 60 characters found in the parameter field.

45

TABLE 3-2.- FIELDS IN THE PROCESSORS OF EOD-LARSYS

Processor	Required fields	Kinds of fields	Permitted structures		
HIST	Required	Areas to be histogrammed	*END Field cards $END		
GRAYMAP	Required	Areas to be displayed as gray-scale maps	*END Field cards $END		
STAT	Required	Training fields	*END CLASS SUBCLASS Field cards CLASS (etc.) $END		
[a]ISOCLS	Required	Training fields	*END CLASS Field cards CLASS (etc.) $END	*END [b]DESIGNATE OTHER or DESIGNATE UNIDE Field cards $END	
SELECT	None				
CLASSIFY	Required	Areas to be classified	*END Field cards $END		
DISPLAY	Optional automatic display of classified fields	[c]Test fields or designated fields	[d]*END CLASS Field cards CLASS (etc.) $END	[e]*END SUBCLASS Field cards SUBCLASS (etc.) $END	*END DESIGNATE OTHER Field cards DESIGNATE UNIDE Field cards $END
DATATR	Required	Areas to be transformed	*END Field cards $END		
TRSTAT	None				
NDHIST	Required	Areas to be histogrammed	OPTION SUBCLS *END CLASS SUBCLASS Field cards SUBCLASS (etc.) $END	OPTION CLASS *END CLASS Field cards CLASS (etc.) *END	*END Field cards $END
SCTRPL	None				
DOTDATA	Required	Dots are treated as point fields	*END TYPE 1 CLASS Field cards CLASS (etc.) Type 2 CLASS Field cards CLASS (etc.) $END	*END [f]LACIE-formatted dot cards $END	
LABEL	Optional	Areas of the unconditional cluster map if a procedure is selected	*END Field cards $END		
DAMRG	Only one field required unless making a mosaic	Rectangular fields only	*END Field cards $END		

[a]ISOCLS generates clusters, which are subclasses.

[b]Designated other or designated unidentifiable.

[c]Can include both.

[d]This type of field deck can be used profitably if statistics were generated by STAT or ISOCLS.

[e]This deck should be used only if statistics were generated by STAT.

[f]LACIE is the acronym for the Large Area Crop Inventory Experiment, conducted at JSC from 1975 through 1978.

4. FILES

The files described in this section are used internally by the system to pass information between processors. It is the user's responsibility to assign the necessary files for his or her particular job. In the discussion that follows, the names of the units on which the files are written are also used to identify the files.

4.1 STATISTICS FILE (SAVTAP, UNIT 20)

The SAVTAP file must be assigned either to disk or to tape, normally the former, whenever one or more of the processors STAT, ISOCLS, SELECT, CLASSIFY, DATATR, TRSTAT, NDHIST, SCTRPL, or LABEL are executed. One file is written on unit 20 for each execution of STAT, TRSTAT, or ISOCLS or for input of a module STAT file to some other processor. The file contains the same information as itemized in section 3.2.4.1 for the module STAT file.

Multiple files may be written on a single unit, usually disk, and may be accessed by using the STATFILE control card. This control card communicates the file number for positioning the unit and the logical unit assignment. The first file is always assumed unless otherwise specified by the user, and the unit assignment assumes Fortran unit 20 unless otherwise specified by the STATFILE control card. In executing several processors consecutively and referencing the same file, only one STATFILE control card need be submitted. If different files are to be referenced during one execution, then the file number may be changed from one processor to the next by input of the STATFILE control card to each processor. For example,

```
$STAT
STATFILE   UNIT=20,FILE=2
(Other control cards)
```

```
*END
(Class, subclass, and field definitions)
$END
$CLASSIFY
*END
(Fields to be classified)
$END
```

The STAT processor will write the training statistics for this run on file 2 of the SAVTAP file (unit 20). (The system files, their logical units, and assignments are set out in table 4-1.) CLASSIFY will use all of the statistics on file 2 of the tape for classification.

The following example shows assignments for back-to-back execution of STAT, ISOCLS, and SELECT.

```
$STAT
STATFILE   UNIT=20,FILE=2
(Other control cards)
*END
(Class, subclass, and field definitions)
$END
$ISOCLS
STATFILE   UNIT=20,FILE=3
(Other control cards)
*END
(Class and field definitions)
$END
$SELECT
STATFILE   UNIT=20,FILE=2
BEST       4
*END
$END
```

STAT will write on file 2 of unit 20, ISOCLS will write on file 3 of the same unit, and SELECT will go back to file 2 of unit 20 for the statistics computed by STAT.

4.2 B-MATRIX FILE (BMFILE, UNIT 10)

The file written to the BMFILE unit contains the transformation matrix which corresponds to the B-matrix file (section 3.2.4.2). The file is an optional input to SELECT, CLASSIFY, and SCTRPL and a required input to DATATR and TRSTAT. When the card deck or image file is input to any of these processors, this file is automatically written. The B-matrix is computed by the SELECT processor and is automatically output to the file when the Davidon-Fletcher-Powell procedure is executed.

4.3 ONE-DIMENSIONAL HISTOGRAM FILE (HISFIL, UNIT 13)

On logical unit 13, the HIST processor creates the HISFIL, which is used by the GRAYMAP processor.

4.4 CLASSIFICATION MAP FILE (MAPTAP, UNIT 2)

The MAPTAP file (appendix D), which is output by CLASSIFY, contains the statistics actually used in the classification, the training field information, and all of the classification results.

4.5 N-DIMENSIONAL HISTOGRAM FILE (NHSTUN, UNIT 4)

The NDHIST processor writes a file to the NHSTUN to be used as an interface to the SCTRPL processor. The default assignment is unit 4, but the user may assign any available unit. The NHSTUN file format is defined in appendix E. In earlier documentation of EOD-LARSYS, the file written on NHSTUN, unit 4, is referred to as the NDIM file.

4.6 TRANSFORMED STATISTICS FILE (SAVTAP, UNIT 20)

The TRSTAT processor writes the transformed statistics on the SAVTAP file. (See section 4.1 for further information.)

4.7 DOT DATA FILE (DOTUNT, UNIT 19)

The DOTDATA processor writes unformatted files on the DOTUNT. The files contain information extracted from the MSS data file, using all or a subset of 209 possible grid points (dots). The files created on the DOTUNT are used in processors ISOCLS, DISPLAY, and LABEL. The format of the dot data file is defined in appendix F.

TABLE 4-1.- DEFAULT FILE ASSIGNMENTS

Fortran unit	EOD-LARSYS file
2	MAPTAP (classification map)
4	NHSTUN (n-dimensional histogram)
6	PRTUNT (printer results)
9	SAVTAP (output from TRSTAT)
10	BMFILE (transformation matrix if on separate file)
11	DATAPE (MSS data tape)
12	SCTRUN (scatter plots)
13	HISFIL (one-dimensional histogram)
14	TRFORM (transformed MSS data)
16	MAPUNT (mixed or conditional cluster map, ISOCLS cluster map, or final classification map)
19	DOTUNT (dot data)
20	SAVTAP (statistics)
21	CRDUNT (control card images)
22	RANDIO (direct access file for pixel storage)
23	GTRWU (LACIE-formatted dot card images)
None	GTRDU (ground truth images — Universal format)

NOTE: Unit 16 is fixed at this time.

54

5. OUTPUT

5.1 CLUSTER MAP FILE (MAPUNT, UNIT 16)

On logical unit 16, the DISPLAY processor optionally outputs a multifile data tape (MAPUNT) containing the subclass number to which each pixel was assigned during classification by CLASSIFY. Also on logical unit 16, the ISOCLS processor outputs a file containing either the cluster number (OPTION CLUSTER control card) or the mean vector to which each pixel was assigned during clustering. A key containing the color code for each cluster is given for the mean vectors. The color codes may be ordered according to the cluster number or to greenness (OPTION ORDER control card). (See section 9.5.3, table 9-1, for ISOCLS control cards.)

The results of the classification/clustering may be displayed on a suitable display device. The necessary tape must be mounted on a nine-track tape drive compatible with the device and may be output in either LARSYS II/III or Universal format. The display may be made without color keys (appendixes B and C) or with color keys (see appendix G for tape format). To exercise this option, see FORMAT control card (table 9-1) for the ISOCLS processor and section 12 (table 12-1) for the DISPLAY processor.*

One file is written on the output tape for each field classified or clustered. In earlier documentation of EOD-LARSYS, the file written to the MAPUNT is referred to as MAPFIL.

*These data are available as input to NDHIST via seven- or nine-track tape or disk.

5.2 SCATTER PLOT DATA FILE (SCTRUN, UNIT 12)

The SCTRPL processor outputs two-axis color-coded spectral plots on a multifile Universal-formatted tape. The default assignment is unit 12, but the user may assign any available unit. (See file assignment chart, table 4-1.) The relevant tape format is defined in appendix H. In earlier documentation of EOD-LARSYS, the file written to the SCTRUN is referred to as PLOTAP.

5.3 TRANSFORMED DATA FILE (TRFORM, UNIT 14)

The DATATR processor outputs a multifile image tape of transformed data. The image file may be produced in either LARSYS II/III or Universal format, defined in appendixes B and C, respectively. The output is usually assigned to logical unit 14.

5.4 MERGED MSS DATA FILE (DATAPE, UNIT 11)

The DAMRG processor outputs a merged image composed from up to six input images. Merging may be by channel, spatial relation, or user-chosen lines.

5.5 LACIE-FORMATTED DOT FILES (GTRWU, UNIT 23)

The GTTCN and GTDDM processors produce labeled dots on the basis of ground truth input.

Table 5-1 summarizes the input and output of EOD-LARSYS files.

TABLE 5-1.- OVERVIEW OF EOD-LARSYS INPUT/OUTPUT

Processor	Possible inputs	Output
HIST	DATAPE Field cards	HISFIL
GRAYMAP	DATAPE HISFIL Field cards	
STAT	DATAPE Field cards	SAVTAP
ISOCLS (TESTSP)	DATAPE BMFILE SAVTAP DOTUNT Field cards	SAVTAP MAPUNT
SELECT	SAVTAP BMFILE	Best channels through common block INFORM BMFILE
CLASSIFY	DATAPE BMFILE SAVTAP Field cards	MAPTAP
DISPLAY	DATAPE DOTUNT MAPTAP Field card images — [a]DO/DU or test fields	MAPUNT
DATATR	DATAPE BMFILE Field cards	TRFORM
TRSTAT	SAVTAP BMFILE	SAVTAP
NDHIST	DATAPE SAVTAP MAPUNT Field cards	NHSTUN (histogrammed by class, subclass, or field)
SCTRPL	NHSTUN SAVTAP BMFILE	SCTRUN
DOTDATA	DATAPE Field card images for dots LACIE-formatted dot card images	DOTUNT
LABEL	SAVTAP DOTUNT MAPUNT Field cards, all-of-a-kind option	MAPUNT (conditional or mixed cluster map for DISPLAY) MAPTAP SAVTAP DOTUNT
DAMRG	MSS data file	MSS data file
GTDDM	MAPUNT	LACIE-formatted dot labels
GTTCN	Ground truth images	MAPUNT

[a]Designated other (DO) or designated unidentifiable (DU).

6. ONE-DIMENSIONAL HISTOGRAM PROCESSOR — HIST

The processor HIST computes individual field histograms and a total histogram for all the fields and channels defined by the user. An individual statistics report is printed for every field histogrammed. The report contains a field description and the data range, mean, standard deviation, and normalized range (mean ±3 standard deviations) for that field.

A cumulative histogram of all the fields is calculated and written on an internal file to be read later by the GRAYMAP processor. As for the field histograms, a statistics report is printed for the combined fields.

The input DISPLAY control card allows the user to obtain a line-printer plot of the histograms. A histogram for each channel on the DISPLAY card (described in table 6-1) is displayed for each field, along with a cumulative histogram for all the fields.

6.1 INPUT FILES

An MSS data file (DATAPE) must be input. The assignment defaults to logical unit 11; but, by input of the DATA control card, the user may assign any available logical unit. (See table 4-1 for file assignments and section 3.1, Image Tapes, for further information on format.)

6.2 OUTPUT FILES

The HIST processor writes a file for the GRAYMAP processor on logical unit 13. This file (HISFIL) contains the histogram data for each channel requested.

6.3 SCRATCH FILES

The HIST processor does not require an additional scratch file.

6.4 CARD INPUT

The formats for all system card input are defined in section 3.2.

6.4.1 PROCESSOR CARD

The processor keyword is left justified starting in column 1, thus,

$HIST

This card directs the system monitor routine to select the HIST processor and causes all the routines used by the HIST processor to be loaded into the system.

6.4.2 SPECIAL SYSTEM FILES

The HIST processor does not use any special input files.

6.4.3 CONTROL CARDS

Control cards allow the user to input various options. These cards are identified by a keyword left justified in columns 1 through 10 of the card, with the parameter values or additional keywords in columns 11 through 72 (beginning in any column after column 10). These control cards may be in any order, but they must be the first cards after the processor card $HIST.
Table 6-1 lists all available options, along with their default values.

6.4.4 FIELD DEFINITIONS

Fields to be histogrammed are input immediately following the *END control card. The card column format for field definitions is given in section 3.2.2. Input of field definition cards is terminated by the $END control card.

6.5 CARD OUTPUT

This processor does not output any card files.

6.6 RESTRICTIONS

a. The maximum number of channels is 30.

b. The number of histograms requested to be plotted may be limited if internal dimensions are too small for all user requests. (For example, if the user requests 30 channels to be histogrammed, only 14 of those histograms may be plotted; however, all 30 will be histogrammed.)

   This limitation is a function of the number of channels requested on the CHANNELS control card. If too many channels are indicated on the DISPLAY control card, a diagnostic is printed but execution continues.

c. The channels on the DISPLAY card must be subset of the channels on the CHANNELS card.

d. The data for all channels for one scan line are unpacked into an array dimensioned 12 000. If the number of channels times [(sample end - sample begin)/sample increment] exceeds 12 000, a diagnostic message is printed. Sample end is reset to fit the dimensions, and execution continues.

6.7 DIAGNOSTIC MESSAGES

The diagnostic messages and the subroutines in which they appear are listed in appendix I.

TABLE 6-1.- CONTROL CARDS FOR HIST

Keyword (a)	Parameter and default values (b)	Function
	Required cards	
CHANNELS	$C_1, C_2, \cdots, C_k$ $k \leq 30$	Channels to be histogrammed, $C_1, C_2, \cdots, C_k$, should be integer numbers separated by commas.
*END	Blank	Signals the end of the control cards.
$END	Blank	Signals the end of all card input for this processor.
	Optional cards	
DATA	UNIT=n,FILE=m Default: n=11,m=1	n is the number of the Fortran logical unit to which the MSS data tape (DATAPE) has been assigned; m-1 is the number of files to be skipped on the unit. For back-to-back execution of more than one processor, if using the same file number, only one DATA control card need be submitted.

[a]The keyword must be left justified in card columns 1 through 10.

[b]The parameter values are in card columns 11 through 72 (beginning in any column past 10).

61

TABLE 6-1.- Concluded.

Keyword	Parameter and default values	Function
SIZE	XHIGH=K $0 < K \leq 255$ Default: XHIGH=255	K is an integer which sets the maximum radiance value which will be histogrammed. XHIGH becomes x_{max} on the x-axis of the histogram plot.[c]
SIZE	XLOW=J $0 < J < XHIGH$ Default: XLOW=0	J is an integer which sets the minimum radiance value which will be histogrammed. XLOW becomes x_{min} on the x-axis of the histogram plot.[c]
SIZE	YSIZ=L $0 < L \leq f(x)_{max}$ Default: YSIZ=15	L is an integer which sets the height of the y-axis (number of print lines). Using the input YSIZ, the y-axis scale for the histogram plot will be determined by the processor to be $f(x)_{max}$ + (YSIZ-1)/YSIZ [$f(x)_{max}$ denotes the largest count in the histogram].
DISPLAY	$C_1, C_2, \cdots, C_k$ $k \leq 30$ Default: No plots	Channels for which histograms will be plotted. $C_1, C_2, \cdots, C_k$ must be subset of the CHANNELS card.

Ancillary cards

HED1, HED2, DATE, COMMENT (see table 3-1)

[c]The difference between XHIGH and XLOW must be at least 100.

7. GRAY MAP PROCESSOR — GRAYMAP

The chief purpose of GRAYMAP is to produce alphanumeric pictorial printouts of digitized MSS data. Each data sample is eight bits, providing 256 possible gray levels for each MSS data channel. To allow a meaningful distinction in gray-scale tones, GRAYMAP assigns each of the 256 levels to 1 of as many as 16 possible symbols. These symbols may be preassigned or arbitrarily assigned for each run. The specifications for the bin edges for each symbol may be assigned arbitrarily by the user for each run or computed from the histogram data in order to result in equal activity for each of the symbols. In any case, the data are subsequently output in terms of symbols, and each symbol represents a range of data values in which the corresponding data points fall.

7.1 INPUT FILES

An MSS data file must be input to the GRAYMAP processor. The assignment defaults to logical unit 11; however, by input of the DATA control card, the user may assign any available logical unit. (See table 4-1 for file assignments and section 3.1, Image Tapes, for further information on format.)

The GRAYMAP processor requires the bin levels to be input on a control card or computed from the histograms output by the HIST processor on the HISFIL. When the bin levels are to be computed, logical unit 13 may be assigned to either disk or tape or allowed to default to disk (no assignment). If the HIST processor has not been executed before running GRAYMAP and bin levels have not been input, a default histogram of every 10th line for 500 lines and every 10th sample for 200 samples is computed, and HISFIL is created on logical unit 13.

Figure 7-1 shows the interaction of the HIST and GRAYMAP processors.

7.2 OUTPUT FILES

No files are output by the GRAYMAP processor.

7.3 SCRATCH FILES

The GRAYMAP processor does not require additional scratch files.

7.4 CARD INPUT

All system card input formats referred to in this section are defined in section 3.2.

7.4.1 PROCESSOR CARD

The processor keyword is left justified starting in column 1, thus,

$GRAYMAP

This card directs the system monitor routine to select the GRAYMAP processor and initiates loading of routines used by GRAYMAP.

7.4.2 SPECIAL SYSTEM FILES

None of the special system files are required for this processor.

7.4.3 CONTROL CARDS

Table 7-1 lists all available options and control cards recognized by GRAYMAP, along with their default values.

7.4.4 FIELD DEFINITIONS

Fields for which gray-scale maps are desired must follow the *END control card. See section 3.2.3 for the format of field definition cards. Field definition input is terminated by the $END control card.

7.5 CARD OUTPUT

The GRAYMAP processor produces no card output.

7.6 RESTRICTIONS

The system-related restrictions in section 24 apply to this processor.

The maximum number of channels allowed is 30, and the maximum number of bin levels is 16.

7.7 DIAGNOSTIC MESSAGES

The diagnostic messages for this processor are listed by subroutine in appendix I.

TABLE 7-1.- CONTROL CARDS FOR GRAYMAP

Keyword (a)	Parameter and default values (b)	Function
	Required cards	
*END	Blank	Signals the end of the control cards.
$END	Blank	Signals the end of all card input for this processor.
	Optional cards	
DATA	UNIT=n,FILE=m Default: n=11,m=1	n is the number of the Fortran logical unit to which the image data tape has been assigned; m-1 is the number of files to be skipped on the unit. For back-to-back execution of more than one processor, if using the same file number, only one DATA control card need be submitted.
CHANNELS	$C_1, C_2, \cdots, C_k$ $k \leq 30$ Default: Graymap for all channels on HISFIL (created by a previous execution of HIST)	Provides pictorial printout for requested channels.

[a]The keyword must be left justified in card columns 1 through 10.
[b]The parameter values are in card columns 11 through 72 (beginning in any column past 10).

64

TABLE 7-1.- Concluded.

Keyword	Parameter and default values	Function
BINLEVEL	$N_1, N_2, \cdots, N_k$ $k \leq 16$ Default: Histograms used to set bin levels	Upper bin edges for gray-scale levels with a range of 0 to 255 and a maximum of 16 levels; the last bin level should always be 255.
SYMBOLS	$S_1, S_2, \cdots, S_k$ $k \leq 16$ Default: Two sets of 10 symbols over-printed, resulting in one of $,X,Θ,O,*,=,.,-,/,b	Set of characters separated by commas, with a maximum of 16 symbols per SYMBOLS card. If two sets are input, the second overprints the first. The number of symbols input on one card determines the number of bin levels when using the histograms to set the levels. Blank is a legitimate character.

Ancillary cards

HED1, HED2, DATE, COMMENT (see table 3-1)

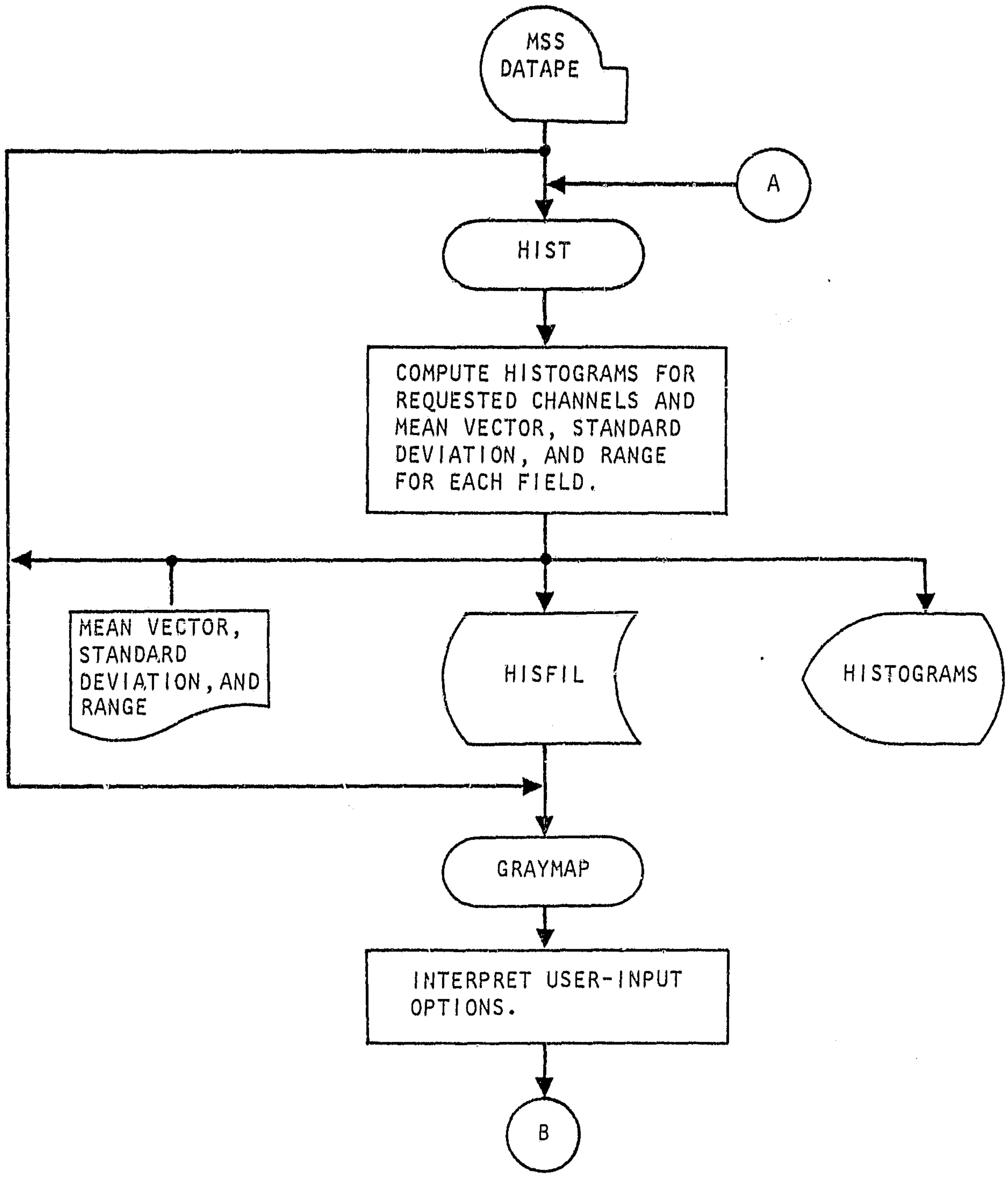


Figure 7-1.— Functional diagram showing interaction of the HIST and GRAYMAP processors.

68

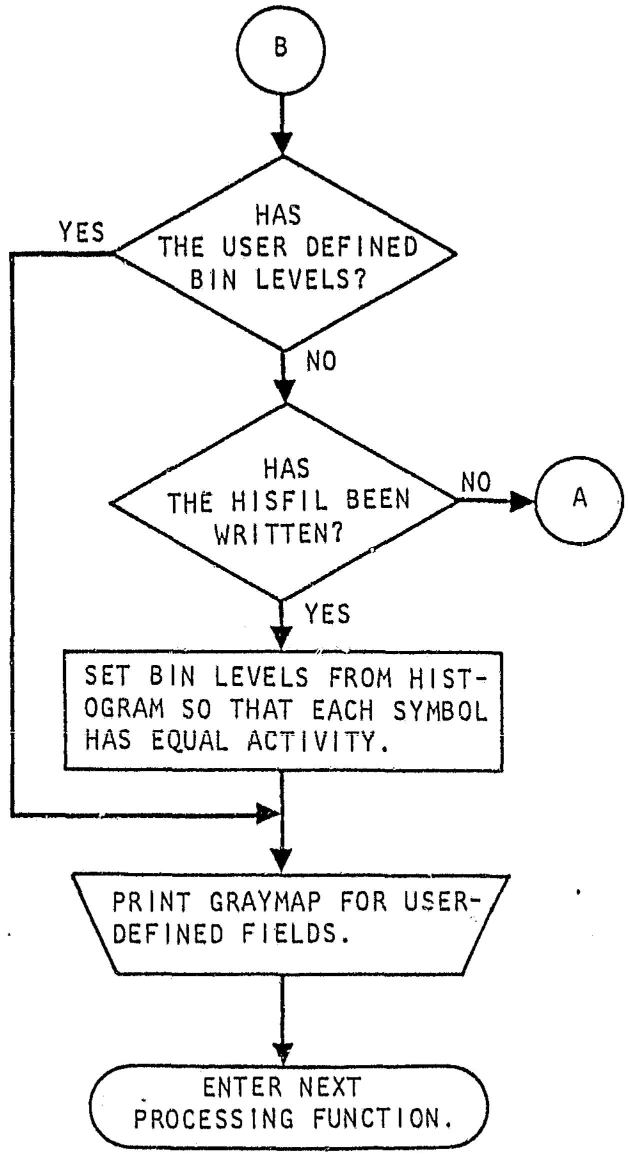


Figure 7-1.— Concluded.

69

8. STATISTICS PROCESSOR — STAT

The statistics processor, STAT, computes the multichannel mean, standard deviation, covariance matrix, and correlation coefficient for each training field and each training subclass defined through user input. In addition, at the user's option, histograms and spectral plots may be computed for fields and subclasses.

The STAT processor requires user input of both card images and an MSS data file. Card image input consists of an optional number of cards from the set of control cards defined in table 8-1 and the training field definitions described in section 3.2.3. The required input MSS data must encompass the areas specified by the training field definitions. The processor is activated by a specific processor card defined in section 8.4.1. Optional functions are directed by means of the input control cards or the built-in default for any control card that is not input.

In addition to the optional printouts under the direction of the control cards, the STAT processor creates the output file SAVTAP, which contains the computed statistics (mean vector and covariance matrix) for each training subclass. The training subclass statistics optionally may be output on punched cards (the module STAT deck). Both the output statistics file SAVTAP and the output module STAT file are in a format acceptable as input to other processors in EOD-LARSYS. Figure 8-1 gives the functional flow of the STAT processor.

The mean vector for the i^{th} subclass is computed as follows:

$$\mu_i = \overline{X}_{1i}, \overline{X}_{2i}, \cdots, \overline{X}_{pi}, \cdots, \overline{X}_{Pi} \qquad (8\text{-}1)$$

70

where

$$\overline{X}_{pi} = \frac{1}{N_i} \sum_{j=1}^{N_i} X_{pj} = \text{average value in channel } p \text{ for subclass } i$$

p = channel number

P = largest channel number

N_i = number of samples in all training fields for subclass i

X_{pj} = the j^{th} sample of the MSS data for channel p (a value between 0 and 255)

The covariance matrix for the i^{th} subclass is computed as follows:

$$K_1 = \begin{bmatrix} k_{11i} & k_{12i} & \cdots & k_{1Pi} \\ k_{21i} & k_{22i} & \cdots & \\ \vdots & & k_{pqi} & \vdots \\ & & \vdots & \\ k_{P1i} & & \cdots & k_{PPi} \end{bmatrix}$$

where

$$k_{pqi} = \frac{1}{N_i - 1}\left(\sum_1^{N_i} X_p X_q - \frac{1}{N_i}\sum_1^{N_i} X_p \sum_1^{N_i} X_q\right)$$

q = channel number

Closely related statistics are the standard deviation and correlation coefficient for the i^{th} subclass, which are computed as follows:

$$\left.\begin{aligned} \sigma_{pi} &= (k_{ppi})^{1/2} \\ \rho_{pqi} &= \frac{k_{pqi}}{(k_{ppi}k_{qqi})^{1/2}} \end{aligned}\right\} \qquad (8-3)$$

where

σ_{pi} = standard deviation in channel p for subclass i; p = q

ρ_{pqi} = correlation coefficient between channels p and q for subclass i

k_{pqi} = element of the covariance matrix for subclass i; the variance between channels p and q

8.1 INPUT FILES

An MSS data file must be input to the STAT processor. The assignment defaults to logical unit 11; however, by input of the DATA control card, the user may assign any available logical unit. (See table 4-1 for file assignments and section 3.1, Image Tapes, for further information on format.)

8.2 OUTPUT FILES

The STAT processor always outputs the statistics on the SAVTAP file and, optionally (by means of the OPTION PUNCH control card), provides the module STAT file on cards. (See section 3.2.4.1 for format of card file.)

The required output file SAVTAP will contain the class names, subclass names, and the subclass statistics (mean vector and covariance matrix) computed by the STAT processor for every subclass defined. The output statistics file must be assigned to either disk or tape. The assignment defaults to logical unit 20; however, by input of the STATFILE control card, the user may assign any available unit. (See table 4-1 for file assignments and table 8-1 for control card description.)

If the STATFILE control card is used, the statistics from more than one execution of STAT may be saved on the same tape.

8.3 SCRATCH FILES

The STAT processor does not require the use of a separate scratch file.

8.4 CARD INPUT

The specific card column formats for the information to be input on the processor and control cards are given in sections 3.2.1 and 3.2.2. Table 8-1 describes the complete set of keywords and option parameters recognized and acted upon by the STAT processor.

If possible, each keyword and its option parameters are to be completely contained on one control card. However, if more parameters are required than can be contained on one card, the control card may be repeated and the parameters continued on the next control card. The parameters for a control card of a given type will be cumulative over all cards of that type.

The control cards follow the $STAT processor card. All options available on the STAT processor have a default setting which is used by the processor for those option parameters not input via control card. The control card *END must be input to signify the end of the set of control cards. Immediately following the *END card, a set or sets of CLASS, SUBCLASS, and training field definition cards must be input. See section 8.4.4 for further details on training field definitions.

8.4.1 PROCESSOR CARD

The processor keyword is left justified starting in column 1, thus,

$STAT

This card directs the system monitor routine to select the STAT processor and indicates the input of STAT processor control card images.

8.4.2 SPECIAL SYSTEM FILES

The processor does not expect input of any of the system-generated files described in section 3.2.4.

8.4.3 CONTROL CARDS

Table 8-1 gives the complete set of control cards which the user may input to direct the STAT processor functions and the default functions performed by the processor. With the exception of the *END and $END control cards, the sequence of the control cards is optional. The *END card must immediately follow the last control card, if any; the CLASS, SUBCLASS, and training field definition cards must immediately follow the *END card; and the $END card must immediately follow the last card of the input training field definitions.

8.4.4 CLASS, SUBCLASS, AND FIELD DEFINITIONS

All CLASS, SUBCLASS, and field definition cards occur between the *END and $END control cards. The formats for these cards are given in section 3.2.3. Training fields are grouped into statistically similar subclasses, and subclasses are further grouped into classes.

A training class is defined to the processor by one card containing the keyword CLASS in columns 1 through 10. The user-determined alphanumeric name to be assigned to the class begins in column 11 and may contain a maximum of four characters (through column 14). At least one CLASS card must be input.

A CLASS card must be followed by at least one subclass grouping. A subclass grouping is on a SUBCLASS card followed immediately by one or more field definition cards. All fields defined by field definition cards following the SUBCLASS card will contribute a cumulative sample set from which the training subclass statistics will be computed for the named subclass. The set of cards — one SUBCLASS card followed by one or more field definition cards — generates the statistics for one training subclass. The number of training fields to be defined for one given subclass is not restricted. The following example shows the grouping of subclasses into classes.

Classes	Subclasses	Card	
		$STAT	
		. . . (Control cards)	
		$END	
1	1	CLASS	CLASS1
		SUBCLASS	SUB 11 (of CLASS1)
		. . . (Fields)	
	2	SUBCLASS	SUB 12 (of CLASS1)
		. . . (Fields)	
2	3	CLASS	CLASS2
		SUBCLASS	SUB 21 (of CLASS2)
		. . . (Fields)	
3	4	CLASS	CLASS3
		SUBCLASS	SUB 31 (of CLASS3)
		. . . (Fields)	
	5	SUBCLASS	SUB 32 (of CLASS3)
		. . . (Fields)	
	6	SUBCLASS	SUB 33 (of CLASS3)
		. . . (Fields)	
		$END	

8.5 CARD OUTPUT

The STAT processor will optionally output on punched cards the module STAT file (section 3.2.4.1). The module STAT file contains the training field vertices, the subclass names, the subclass numbers, the association of training fields with each subclass and class, and the computed statistics for each training subclass.

The module STAT file is output by the processor on cards only when specified by user input of the OPTION PUNCH control card.

8.6 RESTRICTIONS

The system-related restrictions in section 24 apply to the STAT processor.

In addition, a core storage limitation is associated with the total storage required by the training subclass statistics and the various options (i.e., producing histograms and spectral plots). The upper limit on core storage available for all requirements generated by the input to the STAT processor is 10 600 locations. Each subclass covariance matrix requires approximately 1/2(number of channels)^2 locations; each subclass mean vector requires locations equal to the number of channels; and each training field requires seven locations. If a large number of subclasses, channels, and training fields in combination with one or more of the options available by means of the OPTION control cards causes the core storage limits to be exceeded, the STAT processor prints a diagnostic message requesting the user to decrease options, after which it halts (see section 8.7).

76

The following formulas determine the maximum number of fields that can be input for a case without any histograms (eq. 8-4) and another case with subclass histograms (eq. 8-5).

$$\text{NOFLD} = \frac{10\ 600 - \left[5\text{NOSPEC} + 7\text{MAXSUB} + \left(\frac{4 + 2\text{MAXSUB} + 5}{2}\text{NOFEAT} + 1\right)\text{NOFEAT} + 40\right]}{32} \qquad (8\text{-}4)$$

where

NOSPEC = number of subclasses grouped together (maximum of 20)

MAXSUB = maximum number of subclasses

NOFEAT = number of channels

$$\text{NOFLD} = \frac{10\ 600 - \left[5\text{NOSPEC} + 7\text{MAXSUB} + \left(\frac{4 + 2\text{MAXSUB} + 5}{2}\text{NOFEAT} + 1\right)\text{NOFEAT} + 40 + \text{XSIZ}(\text{NOHIST} + 1)\right]}{32} \qquad (8\text{-}5)$$

If fields and subclasses need to be histogrammed, the term XSIZ(NOHIST + 1) should be added to the numerator of equation (8-4), where XSIZ = range of histogram (maximum of 101) and NOHIST = number of channels histogrammed.

8.7 DIAGNOSTIC MESSAGES

The diagnostic messages provided by the STAT processor are listed in appendix I, along with the probable cause and remedy of the condition which prompted the message. During statistical computations, other messages also may be output by utility routines common to both STAT and other processors in EOD-LARSYS. See the system-related messages in appendix I for additional messages obtained from a STAT execution.

TABLE 8-1.- CONTROL CARDS FOR STAT

Keyword (a)	Parameter and default values (b)	Function
		Required cards
CHANNELS	$N_1, N_2, \cdots, N_k$ $1 \leq k \leq 30$ Default: k=30; however, unless the MSS data file (DATAPE) has exactly 30 channels, the default must not be taken.	N's are the integer channel numbers used by the processor in computing training subclass and training field statistics. The channel numbers must be from the set of channels available on the MSS DATAPE file.
*END	Blank	Signals the end of the control cards.
$END	Blank	Signals the end of all card input for this processor.
		Optional cards
DATA	UNIT=n,FILE=m Default: n=11,m=1	n is the number of the Fortran logical unit to which the MSS data tape has been assigned; m-1 is the number of files to be skipped on the unit. For back-to-back execution of several processors, if the same file number is used, only one DATA control card need be input.

[a]The keyword must be left justified in card columns 1 through 10.

[b]The parameter values are in card columns 11 through 72 (beginning in any column past 10).

TABLE 8-1.- Continued.

Keyword	Parameter and default values	Function
OPTION	MAXSUB=N Default: MAXSUB=15	Informs the processor as to the maximum number of subclasses which will be input. The parameter value is used for dimensioning purposes. This parameter must be set by the user if the number of subclasses he or she is about to define will exceed the default. It is advisable to use this option when a large number of training fields are to be processed or when histograms have been requested.
HISTO	$N_1, N_2, \cdots, N_k$ $1 \leq k \leq 30$ Default: k=30	N's are integer channel numbers for use in the histogram option. The channel numbers must be from the set designated on the CHANNELS control card. (NOTE: This control card does not initiate the histogram option.)

'79.

TABLE 8-1.- Continued.

Keyword	Parameter and default values	Function
OPTION	HIST Default: No histograms	A histogram showing frequency distribution of pixels (resolution elements or radiance values) is printed out for every training field and every training subclass defined in the input training field definition deck. For each subclass (or field), a histogram is provided for every channel designated on the HISTO control card.
OPTION	HIST=C Default: No histograms	A histogram printout is provided for every training subclass defined in the input training field definition deck.
OPTION	HIST=F Default: No histograms	A histogram printout is provided for every training field defined in the input training field definition deck.
SIZE	XHIGH=K $0 < K \leq 255$ Default: XHIGH=220	K is an integer which sets the maximum radiance value which will be histogrammed. XHIGH becomes x_{max} on the x-axis of the histogram plot.

TABLE 8-1.- Continued.

Keyword	Parameter and default values	Function
SIZE	XLOW=L $0 \leq L \leq$ HIGH Default: XLOW=120	L is an integer which sets the minimum radiance value which will be histogrammed. XLOW becomes x_{min} on the x-axis of the histogram plot.
SIZE	YSIZ=J $0 < J \leq f(x)_{max}$ Default: YSIZ=14	J is an integer which sets the number of increments on the y-axis of the histogram plot; therefore, it is the height (number of print lines) of the y-axis. Using the input YSIZ, the processor will determine the y-axis scale for the histogram plot to be $f(x)_{max}$ + (YSIZ-1)/YSIZ.
SIZE	XSIZ=K Default: XHIGH - XLOW	Sets the range which will be histogrammed; maximum range is 101.
OPTION	COVAR Default: Statistics are not printed.	The multichannel mean, standard deviation, and covariance matrix (lower triangular portion) are printed out for each training subclass and training field defined in the input training field definition deck.

TABLE 8-1.- Continued.

Keyword	Parameter and default values	Function
OPTION	COVAR=C Default: Statistics are not printed.	The multichannel mean, standard deviation, and covariance matrix (lower triangular portion) are printed out for each training subclass defined in the input training field definition deck.
OPTION	COVAR=F Default: Statistics are not printed.	The multichannel mean, standard deviation, and covariance matrix (lower triangular portion) are printed out for each training field defined in the input training field definition deck.
OPTION	NOCOVAR	No training subclass or training field statistics are printed out.

C-2

82

TABLE 8-1.- Continued.

Keyword	Parameter and default values	Function
SPECTRAL	M_1, M_2, M_3, M_4 $1 \leq M_i \leq 30$ Default: 4 subclasses per spectral plot; subclasses 1, 2, 3, and 4 on the first plot; 5, 6, 7, and 8 on the second plot; etc.	M's are the integer numbers of up to four subclasses that are to be plotted on one single composite spectral plot. The subclass numbers must be obtained from the set of subclasses defined in the input training field definition deck. Subclass 1 is the first subclass defined in the deck, and subsequent subclass numbers are obtained by sequentially numbering the subclasses as they occur in the training field definition deck.
OPTION	SPECTRAL	A spectral plot is printed out for every training subclass and training field defined in the input training field definition deck. The plot consists of the subclass (or field) mean radiance value, mean standard deviation (σ), and mean $\pm 3\sigma$ plotted versus the channel (spectral band) for every channel designated on the CHANNELS control card.

TABLE 8-1.- Continued.

Keyword	Parameter and default values	Function
OPTION	SPECTRAL=C	A spectral plot will be printed out for every subclass defined in the input training field definition deck.
OPTION	SPECTRAL=F Default: Spectral plots for subclasses	A spectral plot will be printed out for every field defined in the input training field definition deck.
SIZE	SPECBAS=I $0 \leq I \leq 105$ Default: SPECBAS=75	I is an integer which sets the minimum radiance value on the y-axis of the spectral plot (i.e., y_{min}). The processor has a fixed y-axis increment (3) and a fixed number of y-axis values (50). Using SPECBAS, the processor determines the y-axis range to be from y_{min}=SPECBAS to y_{max}=SPECBAS +150.
STATFILE	UNIT=n,FILE=m Default: n=20,m=1	n is the number of the Fortran logical unit to which the SAVTAP file has been assigned; m-1 is the number of files to be skipped on the unit before writing SAVTAP.

TABLE 8-1.- Concluded.

Keyword	Parameter and default values	Function
OPTION	PUNCH Default: The module STAT file is not punched, in which case statistics are output on the SAVTAP file only.	The subclass mean vector and covariance matrix for every subclass defined by user input will be punched on cards in a format acceptable as input to other processors in the system. This punched card deck is called the module STAT file.

Ancillary cards

HED1, HED2, DATE, COMMENT (see table 3-1)

85

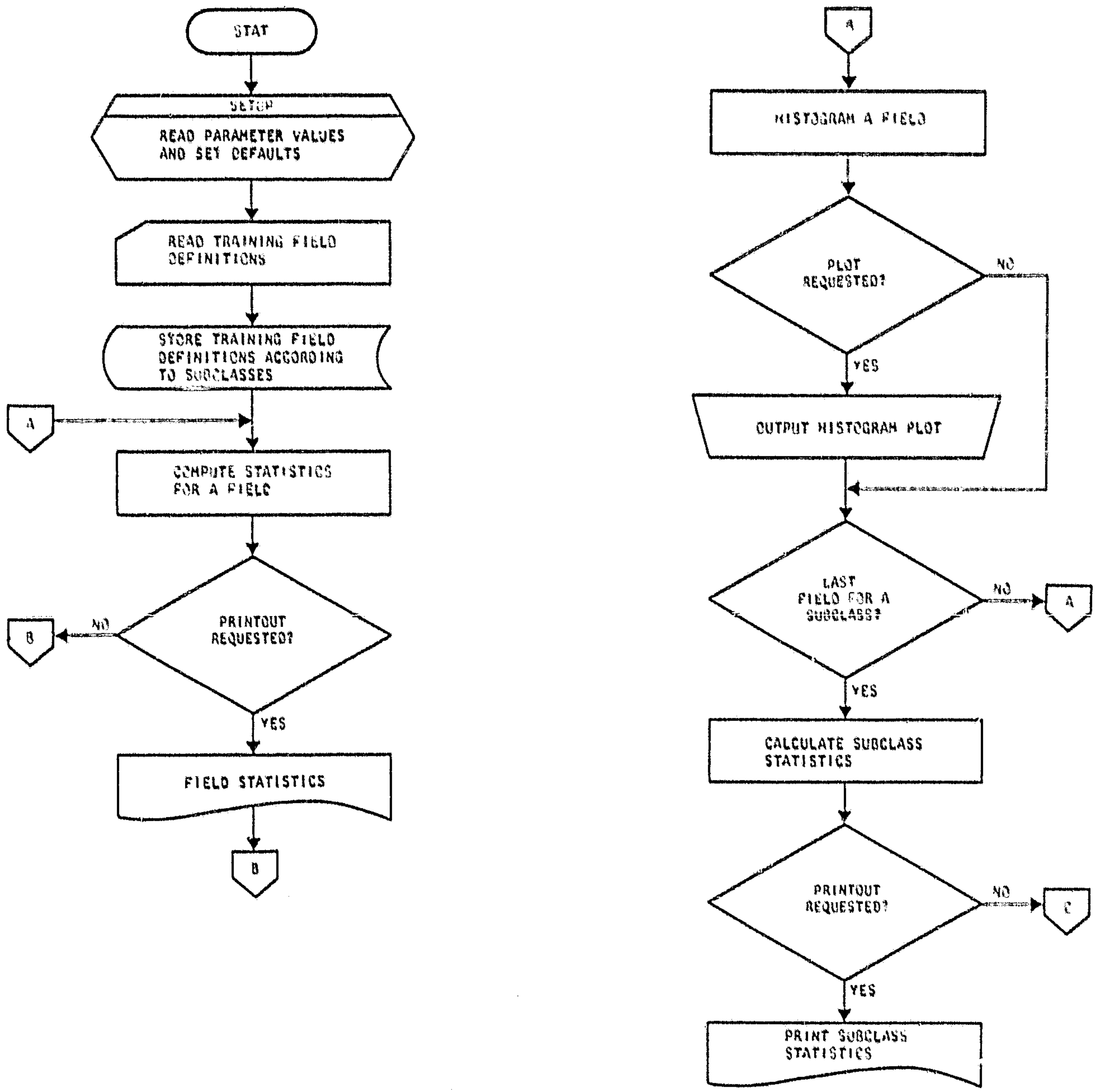


Figure 8-1.– Functional flow chart for the STAT processor.

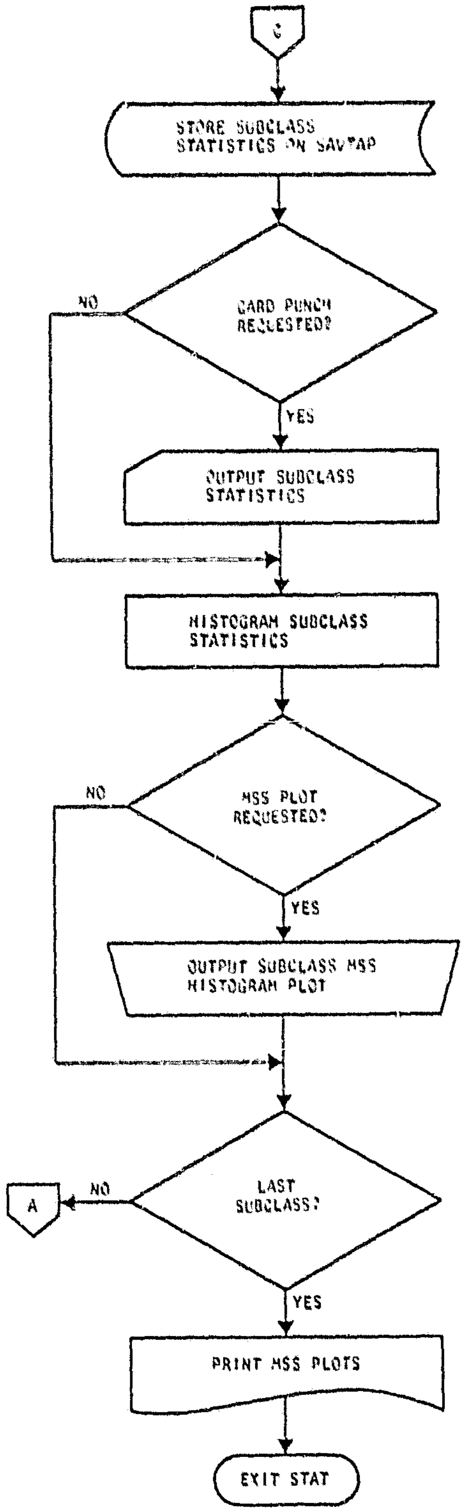


Figure 8-1.— Concluded.

9. ITERATIVE SELF-ORGANIZING CLUSTERING SYSTEM PROCESSOR — ISOCLS

A data set to be clustered by ISOCLS consists of spectral values from one or more user-specified fields in a certain class. With one entry into the ISOCLS processor, any number of classes may be defined and clustered as individual data sets. The user can determine the maximum number of clusters allowed per class by means of the CLUSTERS control card. However, the procedure may find fewer clusters than the maximum allowed. If the user plans to use the statistics generated from the clusters in later CLASSIFY or SELECT runs, he or she must control the number of clusters. The SAVTAP file may contain statistics for up to 75 clusters (or subclasses), but only 60 can be used for processing at any one time in CLASSIFY or SELECT. (Control cards are set out in table 9-1.)

The clustering procedure used in ISOCLS (ref. 2) is an iterative procedure which assigns each MSS data sample to that cluster whose center is the nearest (in absolute distance). At the end of each iteration (i.e., when every sample has been assigned to a cluster), new cluster centers are defined by computing the mean vector for the data samples actually assigned to the cluster. After this initial split sequence, the iterative procedure continues with a user-specified sequence of splits and combinations. See the SEQUEN control card. Using the STDMAX and DLMIN control cards, the user specifies criteria for splitting a cluster or combining clusters.

After the final iteration, the covariance matrix for each cluster is computed and, at the user's option, is printed. All cluster statistics for the class are saved on the scratch file until all classes have been clustered, at which time the SAVTAP file is written. The chaining of clusters for the final map printout is performed, if the user has requested this option (see CHAIN control card, table 9-1). Statistics for the chained clusters are not computed.

The KRN and MAP control cards allow users to control the amount of line-printer output they receive. A final map of the clustered data is always output along with a statistical summary of the clusters, which includes mean and standard deviation vectors, total points assigned to each cluster, and intercluster distances.

The user may either input initial cluster centers to hasten the clustering process or allow the program to start the process by assigning all the data to one cluster, obtaining the mean and standard deviation, and then splitting (unsupervised mode). Initial means may be input (1) by cards (see control card MEANS and cluster means deck, section 3.2.4.3); (2) by the SAVTAP file (see control card STATFILE); or (3) in Procedure 1, from a dot file. Input of the initial means causes a scratch file to be written so that the means can be used repeatedly. Successive classes may or may not use the same means to establish cluster centers for a new class. The control card MEANS allows the user to request cluster centers from the last class to be read from the scratch file and used as initial centers for a new class. Input of a new set of initial means will cause the scratch file to be overwritten with new cluster centers.

Several additions have been made in support of Procedure 1 requirements. A general description of these additions follows.

a. Optionally, starting dots (pixels) from the dot data file, or DOTUNT, may be used to begin clustering.

b. The analyst may identify DO and DU pixel sets (fields) by field card input. The pixels in these fields are not included as inputs to the clustering algorithm. They are assigned special cluster numbers and mean vectors for display purposes.

c. Using the Sun angle correction table, the pixel radiance values may be modified. (The correction table is built in.) The radiance value modification applies only for clustering purposes. The user may input the Sun angles by cards or request that these angles be extracted from the header record of a Universal-formatted MSS data tape.

The clustering algorithm in ISOCLS is detailed step by step in the following subsections. This entire procedure is repeated for each class (or data set).

See the functional flow chart for ISOCLS (fig. 9-1).

9.1 PROCEDURES

9.1.1 NOTATION

Symbol	Fortran name	Definition
CLD_{ij}	CLD(I,J)	Intercluster distance between clusters i and j.
$d\left[X_k, \mu^{(i)}\right]$	DIST	Distance from the data point k to the center of cluster i.
DLMIN	DLMIN	Threshold value for combining clusters.
ISTOP	ISTOP	Maximum number of initial split iterations.
CHNTHS	CHNTHS	Chaining threshold value.
LNCAT	LNCAT or INCAT	Number of existing clusters at a given time.
N(i)	N(I)	Total number of data points assinged to cluster i.

Symbol	Fortran name	Definition
SEQUEN	SEQUEN	User-specified sequence of split and combine iterations.
NMIN	NMIN	Minimum number of data points allowed per cluster both for the initial split iterations and for the first through (NOSEQ-1)th SEQUEN iterations.
PMIN	PMIN	Just before exiting ISOCLS, clusters having fewer than PMIN+NOFEAT pixels are deleted.
NOSEQ	NOSEQ	Maximum number of SEQUEN iterations.
ISEQ	ISEQ	Number of SEQUEN iterations at a given time.
NOFEAT	NOFEAT	Number of coordinates (channels) in a data vector.
STDMAX	STDMAX	Threshold value for splitting clusters.
X_k	C(I,K)	Data vector k, C(I,K)= $(X_1,X_2,\cdots,X_{NOFEAT})_k$
$\mu_j^{(i)}$	MEANS(J,I)	Mean of the jth channel of the ith cluster.
$\gamma_j^{(i)}$	AVP(J,I)	Temporary summing variable for the calculation of the standard deviation of the jth channel of the ith cluster.
$M_j^{(i)}$	AMN(J,I)	Summing variable for computation of new means. After all data have been assigned to clusters on any one iteration, AMN(J,I) is the new mean of the jth channel of the ith cluster.
$\sigma_j^{(i)}$	STDEV(J,I)	Standard deviation of the jth channel of the ith cluster.

9.1.2 ESTABLISHING THRESHOLD VALUES

Establish threshold values for splitting clusters (STDMAX), combining clusters (DLMIN), and deleting clusters (NMIN and PMIN). Then begin the following iterative procedure.

9.1.3 ITERATIVE PROCEDURE

9.1.3.1 Classify and Calculate New Statistics

Assign each data point to a cluster and at the same time collect the means, standard deviations, and point counts of the newly developing clusters. If there are no clusters, set i = 1 and go to iteration b. If clusters exist, begin with iteration a.

a. Assign the data point $X_k = (X_1, X_2, \cdots, X_{NOFEAT})_k$ to the i^{th} cluster if $d\left[X_k, \mu^{(i)}\right] \leq d\left[X_k, \mu^{(j)}\right]$ for all $j \neq i$, where $d\left[X_k, \mu^{(i)}\right]$ is defined as

$$d\left[X_k, \mu^{(i)}\right] = \sum_{j=1}^{NOFEAT} \left| X_{jk} - \mu_j^{(i)} \right| \qquad (9\text{-}1)$$

b.
$$N(i) = N(i) + 1 \qquad (9\text{-}2)$$

c.
$$M_j^{(i)} = \frac{N(i) - 1}{N(i)} M_j^{(i)} + \frac{1}{N(i)} X_{jk} \qquad (9\text{-}3)$$

d.
$$\gamma_j^{(i)} = \frac{N(i) - 1}{N(i)} \gamma_j^{(i)} + \frac{1}{N(i)} X_{jk}^2 \qquad (9\text{-}4)$$

e.
$$\sigma_j^{(i)} = \left\{ \gamma_j^{(i)} - \left[\mu_j^{(i)}\right]^2 \right\}^{1/2} \quad ; \quad j = 1, NOFEAT \qquad (9\text{-}5)$$

Return to step a and repeat iterations a through e until all data points have been classified.

9.1.3.2 Delete Clusters

For the initial split iterations and for the first through (NOSEQ-1)th user-specified SEQUEN iterations, delete all clusters that have fewer than NMIN members. For the NOSEQth user-specified iteration (last user-input sequence), delete all clusters that have fewer than PMIN members. A cluster is deleted simply by removing the statistics for that cluster and reducing the number of clusters (specified by LNCAT) accordingly.

9.1.3.3 Test for Completion

If this is not the last iteration, proceed to 9.1.3.4. If this is the last iteration and no clusters were deleted, the procedure is finished. If one or more clusters were deleted for having less than PMIN members, go back to 9.1.3.1 and reassign the data to the clusters obtained from iteration (NOSEQ-1).

9.1.3.4 Determine Type of Iteration

Determine whether this is to be a split iteration or a combine iteration and proceed to the appropriate step.

The sequence of iterations will be as follows:

SSSS	CCSCSC
ISTOP and/or PERCENT	SEQUEN

where

S = split iteration

C = combine iteration

The beginning sequence of split iterations is terminated either (1) when the percentage of stabilized clusters is greater than or equal to the user-input percentage (see PERCENT control card, table 9-1) or (2) when ISTOP iterations have been reached. After that point, the type of iteration (split or combine) and number of iterations (NOSEQ) are determined by the SEQUEN parameter.

The initial split iterations are primarily intended for the automatic establishment of cluster centers in the event they are not input. The sequence is shortened considerably if initial cluster centers are input.

9.1.3.5 Split Clusters

A cluster is split along the j^{th} channel (1) if the j^{th} channel has the maximum standard deviation for the cluster, (2) if the standard deviation along the j^{th} channel is greater than the STDMAX threshold parameter, and (3) if the cluster has more than 2(NMIN+1) data points.

If conditions (1) through (3) are met, two new clusters are created and the parent cluster is deleted. A cluster is created merely by defining its centers (means) for each channel. If the i^{th} cluster is split in the j^{th} channel, the two new clusters will have centers at $\left[\mu_1^{(i)}, \mu_2^{(i)}, \cdots, \mu_j^{(i)} \pm \alpha, \cdots, \mu_{NOFEAT}^{(i)}\right]$, where α will normally be $\sigma_j^{(i)}$ but can be a user-input constant (see SEP control card). On a given split iteration, if the maximum number of clusters (CLUSTERS) has not been reached, all clusters having a standard deviation greater than the STDMAX parameter will be split. To ensure that the clusters with the largest standard deviations receive the highest priority for splitting, when $2 \times$ LNCAT is greater than CLUSTERS, the standard deviations are ordered along the j^{th} channel in descending order. Return to 9.1.3.1 after splitting clusters.

9.1.3.6 Combine Clusters

Two clusters are combined if the distance between them is less than the DLMIN threshold parameter. The distance between clusters i and j is calculated as

$$CLD_{ij} = \left[\sum_{k=1}^{NOFEAT} \frac{\mu_k^{(i)} - \mu_k^{(j)}}{\alpha_k^{(i)} \alpha_k^{(j)}} \right]^{1/2} \qquad (9\text{-}6)$$

If CLD_{ij} < DLMIN and CLD_{ij} = MIN(CLD_{ij}) for all i = 1,LNCAT and j = 1,LNCAT for all i ≠ j, clusters i and j will be merged to form a new cluster L with means

$$\mu_k^{(L)} = \frac{N(i)\mu_k^{(i)} + N(j)\mu_k^{(j)}}{N(i) + N(j)} \quad ; \quad k = 1,NOFEAT \qquad (9\text{-}7)$$

The clusters i and j are deleted. The new cluster L is not considered as a candidate for merging with any other cluster on the iteration in which it was formed. Return to 9.1.3.1 after combining clusters.

9.1.4 CHAINING

A final optional step in the clustering procedure groups all clusters which have intercluster distances less than the chaining threshold (CHNTHS) to form one cluster. The chaining procedure was adopted because the minimum variance criterion used in the iterative procedure above tends to group the data into spherical (or ellipsoidal) groupings with Gaussian distributions. This type of grouping is certainly a natural grouping and would quite often be completely satisfactory. However, some natural groupings of the data are odd shaped and cannot be approximated by Gaussian distributions. Two examples are given in figure 9-2.

95

At the end of the sequence of split and combine iterations, groupings of the type in figure 9-2 are likely to be separated into subclusters as illustrated in figure 9-3. The chaining algorithm allows the user to group subclusters 1, 2, and 3 (fig. 9-3) into one composite cluster; likewise, subclusters 4, 5, 6, and 7 can be grouped together to form one cluster.

The algorithm scans the intercluster distance table (CLD) and begins a chain with the first appearance of two clusters within a distance of CHNTHS units. Once a subcluster is in the chain, all clusters that are within CHNTHS units of the subcluster are added to the chain. See figure 9-4.

The statistics (mean vector and covariance matrix) of a cluster resulting from chaining are not calculated by the program because, in many cases, the chained cluster cannot be represented by a Gaussian distribution.

There are, of course, instances where one can safely combine those subclusters that are chained by the program into one composite (Gaussian) cluster. For example, subclusters 1, 2, and 3 in figure 9-5 can judiciously be combined into one final cluster. This is indicated by the fact that, pairwise, these three subclusters are all close to each other. In this case, the following formulas (ref. 2) can be used iteratively to compute the composite statistics.

Assuming that two clusters (n_1,m_1,C_1) and (n_2,m_2,C_2) are to be considered as one cluster (n,m,C), where all n, m, and C are the

96

numbers of points, the mean vectors, and the covariance matrices, respectively, and m^T is the transpose of m then

$$\left.\begin{aligned}
n &= n_1 + n_2 \\
m &= \left(\frac{n_1}{n_1 + n_2}\right)m_1 + \left(\frac{n_2}{n_1 + n_2}\right)m_2 \\
C &= \left(\frac{n_1}{n_1 + n_2}\right)C_1 + \left(\frac{n_2}{n_1 + n_2}\right)C_2 + \left(\frac{n_1}{n_1 + n_2}\right)m_1 m_1^T \\
&\quad + \left(\frac{n_2}{n_1 + n_2}\right)m_2 m_2^T - mm^T
\end{aligned}\right\} \qquad (9\text{-}8)$$

9.2 INPUT FILES

An MSS data file must be input to the ISOCLS processor. The assignment defaults to Fortran unit 11; however, by input of the DATA control card, the user may assign any available logical unit. (See table 4-1 for file assignments and section 3.1, Image Tapes, for further information on format.)

In support of Procedure 1, the starting dot information from the DOTUNT file must be input to begin the cluster processing; i.e., furnish starting cluster centers.

9.3 OUTPUT FILES

Statistics are output by ISOCLS to the SAVTAP file (section 4.1). The file assignment defaults to logical unit 20; but, by input of the STATFILE control card, the user may assign any available logical unit. (See table 4-1 for file assignments and STATFILE control card, table 9-1, for further information.)

A cluster map tape (MAPUNT) may be generated optionally for displaying the results of the clustering on a suitable display device. The FORMAT control card requests the option and names the desired format of the tape. Logical unit 16 should be assigned to a nine-track tape drive when this option is exercised (see section 5.1).

A printout of the cluster results consists of the following data items by class: cluster numbers and symbols; cluster mean vectors (by channel); cluster standard deviations by channel; intercluster distances; number of pixels per cluster; number of clusters; and cluster map by field.

9.4 SCRATCH FILES

The program dynamically assigns random access disk storage for scratch files. ISOCLS uses the disk storage for temporary storage of cluster statistics, the data to be clustered, and the classification of each pixel.

9.5 CARD INPUT

9.5.1 PROCESSOR CARD

The processor keyword is left justified starting in column 1, thus,

$ISOCLS

This card directs the monitor routine to call the ISOCLS processor and causes all routines used by the ISOCLS processor to be loaded into the system.

9.5.2 SYSTEM CARD FILES

The processor will read a cluster means deck in the format defined in section 3.2.4.3. The deck may be used to establish cluster centers for the clustering procedure.

9.5.3 CONTROL CARDS

Control cards allow the user to input various options. They are identified by a keyword that is left justified in columns 1 through 10 of the card, with parameter values or additional keywords in columns 11 through 72. These control cards may be in any order, but they must be the first cards after the processor card $ISOCLS. Table 9-1 lists all available options, along with their parameter values.

9.5.4 CLASS AND FIELD DEFINITIONS

A CLASS card, followed by at least one field definition card, must immediately follow the *END control card. The formats for these cards are defined in section 3.2.3.

The pixels from all fields for one class are extracted from the sequential MSS data file (tape or disk) and stored on a direct-access disk file. The data from all fields for one class are clustered as one data set. The statistics for all clusters in that class are saved on a scratch file, and the next class is clustered. When all classes have been clustered, the statistics are written on the SAVTAP file. The input for definition of classes and fields is explained below.

	*END	
Wheat clustered as one data set.	CLASS	WHT
	FLD 1	
	FLD 2	
	FLD 3	
Nonwheat clustered as one data set.	CLASS	NWHT
	FLD 4	
	FLD 5	
	$END	

Note that actual names may not exceed four characters.

ISOCLS recognizes DO/DU fields. All the DO/DU field cards (for all classes) must be input before the fields to be clustered. These fields must immediately follow the *END card. The CLASS card follows the last DO/DU field card.

Examples:

- If DO/DU fields are being defined:

```
*END
DESIGNATED     OTHER
OTHER          (1,1),(1,1),(40,1),(40,20),(1,20)
DESIGNATED     UNIDENTIFIABLE
UNIDE          (1,1),(5,7),(8,7),(8,10),(5,10)
CLASS          WHT
WHT1           (1,1),(1,1),(196,1),(196,117),(1,117)
$END
```

- If no DO/DU fields are being defined:

```
*END
CLASS          WHT
WHT1           (1,1),(1,1),(196,1),(196,117),(1,117)
$END
```

9.6 CARD OUTPUT

A module STAT file (see section 3.2.4.1) may be punched and used as an interface between ISOCLS and SELECT or CLASSIFY. This option is exercised via the OPTION PUNCH control card. (Normally, however, an unformatted statistics file, SAVTAP, is written to disk for subsequent use.)

9.7 RESTRICTIONS

The ISOCLS processor uses disk for a temporary scratch file. There are approximately 750 000 words of storage available on this file. The data to be clustered for one class are stored on this file, along with other information. To compute the maximum number of pixels per class, use the following formula.

$$\text{Maximum pixels} = \frac{750\,000 - 30\left\{\begin{matrix}\text{number of}\\ \text{classes}\end{matrix}\left[\left(\begin{matrix}\text{number of}\\ \text{channels}\end{matrix}\right)^2 + 3\left(\begin{matrix}\text{number of}\\ \text{channels}\end{matrix} + 2\right)\right] - 1800\right\}}{\text{number of channels} + 1} \tag{9-9}$$

The maximum number of clusters per class is 60, and the maximum number of channels is 30. The covariance matrices for all clusters in one class must be stored in core at one time. They are stored in an array dimensioned 11 500. The following formula may be used to see if enough storage is available for the covariances.

$$11\,500 \geq \begin{matrix}\text{number of}\\ \text{clusters}\end{matrix}\left[\begin{matrix}\text{number of}\\ \text{channels}\end{matrix}\left(\frac{\begin{matrix}\text{number of}\\ \text{channels}\end{matrix} + 1}{2}\right)\right] \tag{9-10}$$

9.8 DIAGNOSTIC MESSAGES

Diagnostic messages for the ISOCLS processor are presented in appendix I.

TABLE 9-1.- CONTROL CARDS FOR ISOCLS

Keyword (a)	Parameter and default values (b)	Function
	Required cards	
CHANNELS	DATA=$C_1, C_2, \cdots, C_k$, STAT=$A_1, A_2, \cdots, A_k$ $k \leq$ number of channels on SAVTAP ≤ 30	C's are integer channel numbers that (1) will be used in clustering and (2) refer to the MSS data file. A's are integer channel numbers that (1) will be the starting vectors (initial means) of clusters, (2) refer to the SAVTAP file, and (3) must be a subset of the channels on the SAVTAP file. The same channels must be used throughout one execution of ISOCLS. If a cluster means card file is input, the channels on the CHANNELS card must be a subset of the channels in the means card file.
*END	Blank	Signals the end of the control cards.

[a]The keyword must be left justified in card columns 1 through 10.

[b]The parameter values are in card columns 11 through 72 (beginning in any column past 10).

TABLE 9-1.- Continued.

Keyword	Parameter and default values	Function
$END	Blank	Signals the end of all classes to be clustered for this set of control cards. The SETUP routine will be reentered to read new control cards for the next class until all classes have been clustered.
	Optional cards	
DATA	UNIT=n,FILE=m Default: n=11,m=1	n is the number of the Fortran logical unit to which the MSS data file has been assigned; m-1 is the number of files to be skipped on the unit. For back-to-back execution, if the same data file is to be processed throughout the execution, only one DATA card need be submitted.
CLASSES	N Default: 1	Number of classes to be clustered.
CLUSTERS	N Default: 60	Maximum number of clusters per class; N must be $\leq$ 60.

TABLE 9-1.- Continued.

Keyword	Parameter and default values	Function
OPTION	CLUSTER Default: The output cluster MAPUNT file will contain the mean vector of the cluster to which the corresponding pixel was assigned.	The output cluster MAPUNT file will contain the cluster number to which the corresponding pixel was assigned. When selecting this option, the FORMAT control card must be input also.
MEANS	CARDS Default: Clustering procedure is automatically begun if this card or means file is not input.	Initializes input of the cluster means deck. This deck is used to establish cluster centers for the clustering procedure.
MEANS	FILE Default: Cluster centers are automatically established if this card or the means file is not input.	Indicates that means for initial clusters have been input previously from cards and stored on file. The same initial means are to be used again to repeat the process for a new data set.

TABLE 9-1.- Continued.

Keyword	Parameter and default values	Function
STATFILE	INPUT/UNIT=n,FILE=m, OUTPUT/UNIT=ℓ,FILE=s Default: No defaults for INPUT; ℓ=20,s=1 for OUTPUT	n is the number of the Fortran logical unit to which the SAVTAP file containing the initial means has been assigned; m-1 is the number of files to be skipped on the unit; ℓ is the number of the Fortran logical unit to which the SAVTAP file containing the generated statistics will be output; s-1 is the number of files to skip before writing the cluster statistics.
SUBCLASSES	$C_1, C_2, \cdots, C_k$ $k \leq 60$ Default: All subclasses/clusters on SAVTAP file will be used in starting the clustering.	C's are integer subclass or cluster numbers that (1) will be used in computing the initial means, (2) refer to the SAVTAP file, and (3) must be a subset of the subclasses or clusters on the SAVTAP file.
MODULE	Blank	Causes the reading of the module STAT deck that immediately follows this card.
ISTOP	N Default: 10	A maximum of N iterations is performed in the initial split sequence.

TABLE 9-1.- Continued.

Keyword	Parameter and default values	Function
SEQUEN	AA···A Default: SC	The A's represent the sequence of S and C characters used for iteration control after the initial split sequence. A maximum of 19 characters may be input.
STDMAX	X Default: 4.5	On a split iteration, splits any cluster whose maximum standard deviation is greater than X units.
SEP	X Default: Maximum of the channel standard deviations in the cluster	When splitting a cluster, separates the new clusters by a distance of X units.
PERCENT	N Default: 80	If the number of clusters flagged for splitting divided by the total number of clusters is less than or equal to $\frac{100-N}{100}$, then splitting stops.
DLMIN	X Default: 3.2	On a combine iteration, combines any two clusters whose means are closer than X units.

106

TABLE 9-1.- Continued.

Keyword	Parameter and default values	Function
NMIN	N	Deletes any cluster with fewer then N members on the first through next-to-last iteration.
PMIN	N	Deletes any cluster with fewer than N members on the last iteration.
OPTION	STATS	Prints the covariance matrix for each cluster.
OPTION	ERCOMP	Prints a measure of the scattering of the clusters after each iteration.
CHAIN	CHNTHS Default: Chaining not performed	Chains all clusters within CHNTHS units of each other to form one cluster. Chaining affects only the final map printout and MAPUNT file.
KRN	N Default: 20	Prints out a summary of the clusters at every N^{th} iteration.
MAP	N Default: 20	Prints out a map of the cluster data along with the summary for every N^{th} iteration. A final cluster map is printed regardless of this parameter.

109

TABLE 9-1.- Continued.

<table>
<tr><th>Keyword</th><th>Parameter and default values</th><th>Function</th></tr>
<tr><td>FORMAT</td><td>UNIVERSAL
Default: Output MAPUNT file is not generated.</td><td>Generates the output cluster MAPUNT file in Universal format.</td></tr>
<tr><td>FORMAT</td><td>LARSYS
Default: Output MAPUNT file is not generated.</td><td>Generates the output cluster MAPUNT file in LARSYS format.</td></tr>
<tr><td>SYMBOLS</td><td>$S_1,S_2,S_3,\cdots$
Default: 1,2,···,9, A,B,···,Z,—,|,¬,/,-, *,+,$,",=,0,blank,=, +,),(,:,&,>,;,?,H, comma,period,blank</td><td>Symbols used to identify clusters in the printout.</td></tr>
<tr><td>OPTION</td><td>ORDER
Default: The color keys will be ordered according to cluster number.</td><td>The color keys on the MAPUNT file will be ordered according to greenness.</td></tr>
<tr><td>OPTION</td><td>PUNCH</td><td>Punches the means and covariance matrix for each cluster in the module STAT deck format.</td></tr>
<tr><td colspan="3">Special cards, for Procedure 1</td></tr>
<tr><td>DOTFIL</td><td>INPUT/UNIT=n,FILE=m
Default: Self-starting</td><td>Defines the Fortran unit number n and file number m of the dot data file (DOTUNT) which contains the starting vectors.</td></tr>
</table>

108

TABLE 9-1.- Concluded.

Keyword	Parameter and default values	Function
DOTS	$n_1, n_2, \cdots, n_{60}$ Default: Dots will not be used for starting vectors.	These integer numbers separated by commas specify the dots to be used as starting vectors.
SUNANG	TAPE Default: No Sun angle correction applied	Sun angles are extracted from the Universal-formatted MSS data tape.
SUNANG	$n_1, n_2, \cdots, n_j$ $\{n_j\}$ are integer numbers, $j \leq 8$ Default: No Sun angle correction applied	$\{n_j\}$ are the Sun angles to be used in computing the Sun angle corrections for use in the clustering algorithm. A Sun angle must be input for each acquisition whose channels appear on the CHANNELS control card.

Ancillary cards

HED1, HED2, DATE, COMMENT (see table 3-1)

109

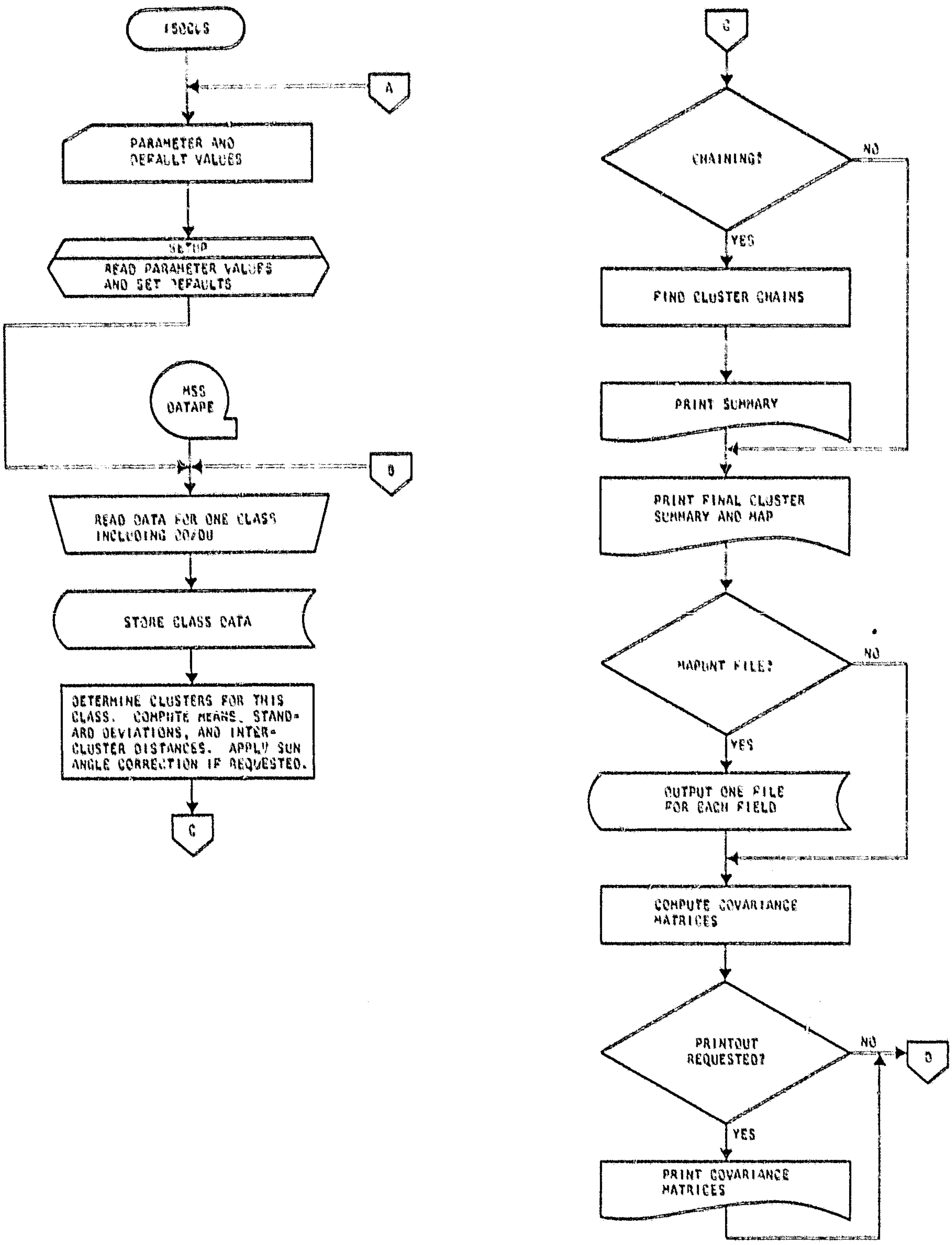


Figure 9-1.— Functional flow chart for the ISOCLS processor.

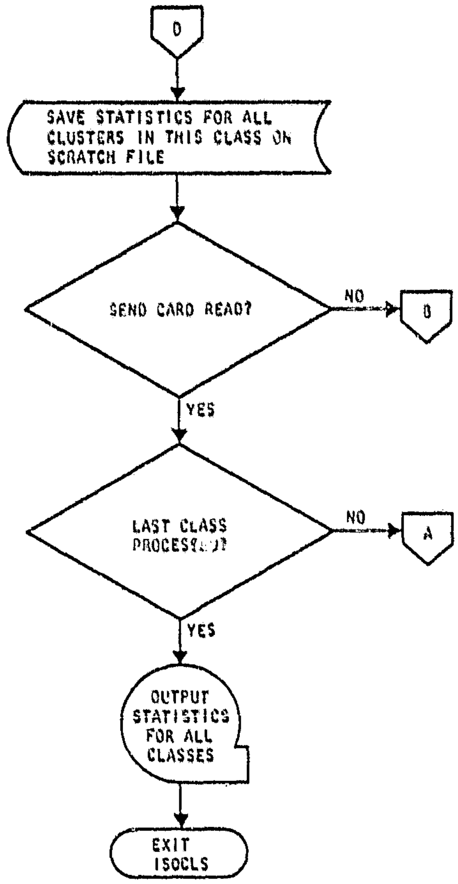
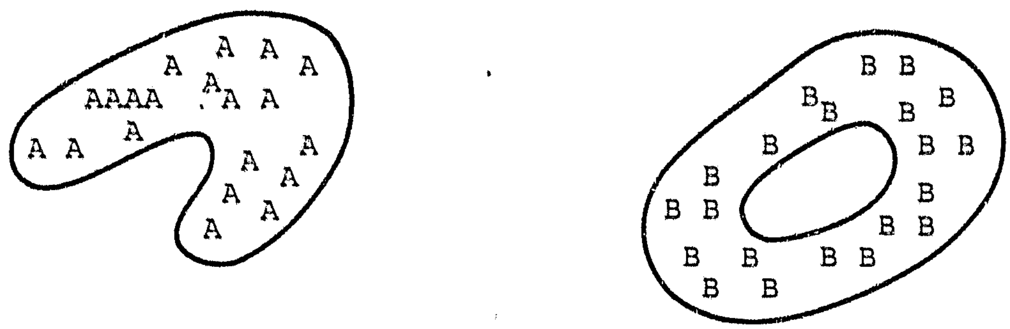


Figure 9-1.— Concluded.



(a) The boomerang-shaped cluster.

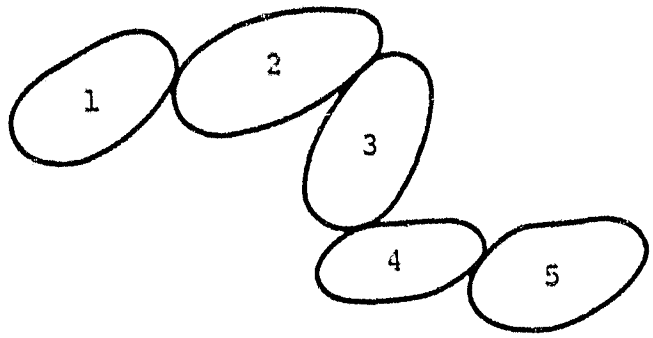
(b) The donut-shaped cluster.

Figure 9-2.— Odd-shaped clusters.

(a) Subclustering of the boomerang-shaped cluster.

(b) Subclustering of the donut-shaped cluster.

Figure 9-3.— Breaking up of the clusters into subclusters.



(a) Cluster structure.

j \ i	1	2	3	4	5	
1	0.0	7.5	6.2	3.2	11.8	
2	7.5	0.0	3.1	5.6	3.0	
3	6.2	3.1	0.0	3.1	6.3	CHNTHS = 3.2
4	3.2	5.6	3.1	0.0	9.7	
5	11.8	3.0	6.3	9.7	0.0	

(b) Intercluster distance table.

Figure 9-4.— Example of chaining.

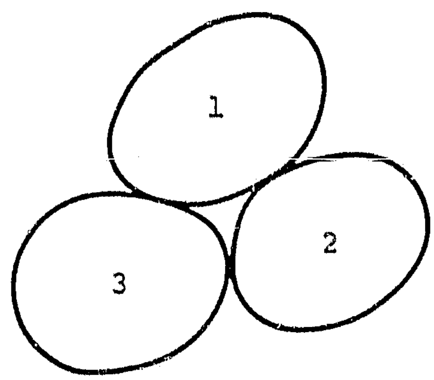


Figure 9-5.— Example in which chained subclusters can judiciously be combined into one composite cluster.

10. FEATURE SELECTION PROCESSOR — SELECT

The feature selection processor, SELECT, provides a means of measuring the relative importance of the individual channels and obtaining the set of channels that provides the best discrimination between subclasses. The processor allows the user to choose one of the following three criteria for measuring the separability of the subclasses for a set of channels or for linear combinations of the channels.

- Weighted average interclass divergence
- Weighted average transformed divergence
- Weighted average Bhattacharyya distance

Either the exhaustive search or the without replacement procedure can be used with one of the above criteria to select a "best" set of channels. The exhaustive search procedure determines the best set of k out of n channels by computing the separability measure for every possible combination of k channels. This results in $n!/k!(n - k)!$ computations of the separability measure. The computer time required for this procedure is prohibitive for a large n. In such cases, the without replacement procedure can be used.

The without replacement procedure determines the best k out of n channels in the following manner. First, the single channel that extremizes the separability measure is selected. Each of the remaining (n - 1) channels is paired with the best single channel to select the best pair of channels. The best triplet is determined by combining the remaining (n - 2) channels with the best pair. The process continues until the best set of k channels has been selected. The number of times the separability measure must be computed is $n + (n - 1) + (n - 2) + \cdots (n - k + 1)$.

A third procedure, the Davidon-Fletcher-Powell procedure, is a powerful iterative descent method for finding a local minimum of a function of several variables. The procedure is discussed in reference 7. How the procedure applies to the problem of channel selection or dimensionality reduction is discussed in reference 8. In SELECT, the Davidon-Fletcher-Powell procedure computes a k-by-n linear transformation matrix that extremizes a given separability measure. This matrix, referred to as the B-matrix, is saved on the BMFILE (section 4.2) and optionally is punched on cards (B-matrix deck, section 3.2.4.2) for later input into the CLASSIFY, DATATR, TRSTAT, or SCTRPL processors.

An initial guess for the B-matrix must be provided for the Davidon-Fletcher-Powell routines, and this guess may be input via the B-matrix on card deck or BMFILE. If the initial guess is not provided by the user, SELECT will execute the without replacement procedure first to obtain a best set of channels, which it will use as a first-guess B-matrix for the Davidon-Fletcher-Powell procedure.

In addition to selecting a best set of channels and/or linear combinations, the processor will evaluate any one of the three separability measures for a specified linear combination of the channels. The linear combination must be input via the B-matrix deck or the BMFILE. This option is the fourth procedure referred to by the PROCEDURE control card.

The processor can also evaluate any of the separability measures for specified sets of channels. This request is made using the EVALUATE and PROCEDURE control cards. This is the fifth option referred to by the PROCEDURE control card.

The best subset of passes (Landsat acquisitions) from a set of passes can also be determined using the sixth procedure option.

For Procedure 1 applications, the SELECT processor provides an option, CLSWT, to weight subclass pairs automatically. The processor determines the class-subclass correspondence (after any grouping of subclasses by use of the GROUP control card) and assigns a weight of 1.0 to each subclass pair representing two different classes. Pairs of subclasses from within one class are given a weight of 0.0. The APRIORI control card can be used to modify these weights with factors computed from the number of pixels in each cluster.

See the functional flow chart for the SELECT processor (fig. 10-1).

10.1 INPUT FILES

The SELECT processor requires the statistics output from either STAT or ISOCLS. Both STAT and ISOCLS write the SAVTAP file and optionally punch the module STAT file on cards (see section 3.2.4.1 for format) which may be used as input to SELECT.

10.2 OUTPUT FILES

The BMFILE is output by SELECT when the Davidon-Fletcher-Powell procedure is used (see appendix J for sample execution). The file is written on logical unit 10.

The corresponding B-matrix card file is punched if the OPTION PUNCH control card is included in the deck setup.

10.3 SCRATCH FILES

Disk files are used as scratch files in SELECT. No assignment is necessary.

10.4 CARD INPUT

All system card input formats referred to in this section are defined in section 3.

10.4.1 PROCESSOR CARD

The processor keyword is left justified starting in column 1, thus,

$SELECT

This card directs the system monitor routine to execute SELECT and initiates loading of routines used by SELECT.

10.4.2 SYSTEM CARD FILES

The processor will read and process the module STAT file and the B-matrix file.

10.4.3 CONTROL CARDS

Table 10-1 lists the control cards that are recognized by SELECT.

10.4.4 FIELD DEFINITIONS

Field definitions do not apply to the SELECT processor.

10.5 CARD OUTPUT

SELECT outputs the B-matrix file on cards as an option (see control card OPTION PUNCH). This is optional output only when the Davidon-Fletcher-Powell procedure is executed.

10.6 RESTRICTIONS

The system-related restrictions in section 24 apply to the SELECT processor.

Two large arrays are dimensioned in SELECT and used for the variable dimensioning of several smaller arrays. Storage requirements of one array are a function of the number of subclasses and channels requested. That is,

$$\left[\text{number of channels}\left(\frac{\text{number of channels} + 3}{2}\right)\right]\text{number of subclasses} + \left[\text{number of best}\left(\text{number of best} + 3\right)\right]\text{number of subclasses}$$
$$+ \left(\text{number of subclasses}\right)\left(\text{number of subclasses} + 3\right) \leq 10\ 600 \qquad (10\text{-}1)$$

Storage requirements of the other array are dependent on the procedure and criterion being used. The Davidon-Fletcher-Powell procedure requires much more storage than the other procedures, and the weighted average interclass divergence requires more storage than the other criteria.

10.7 DIAGNOSTIC MESSAGES

Diagnostic messages for the SELECT processor are listed by subroutine in appendix I.

TABLE 10-1.- CONTROL CARDS FOR SELECT

Keyword (a)	Parameter and default values (b)	Function
		Required cards
BEST	$N_1, N_2, \cdots, N_k$	Find the best set of $N_1, N_2, \cdots, N_k$ channels if the exhaustive search or the without replacement procedure is indicated. If the Davidon-Fletcher-Powell procedure is indicated, the best $N_1, N_2, \cdots, N_k$ linear combinations of the channels are found. $N_1, N_2, \cdots, N_k$ are integers separated by commas. A request can be made for a maximum of 10 best in one call to SELECT.
*END	Blank	Signals the end of the control cards.
$END	Blank	Signals the end of all card input for this processor.

[a]The keyword must be left justified in card columns 1 through 10.

[b]The parameter values are in card columns 11 through 72 (beginning in any column past 10).

120

TABLE 10-1.- Continued.

Keyword	Parameter and default values	Function
	Optional cards	
CHANNELS	$C_1, C_2, \cdots, C_k$ $k \leq$ number of channels on SAVTAP $\leq$ 30 Default: All channels on the SAVTAP file	Selects the best set of channels from those indicated on this card. Must be a subset of the channels for which statistics are input via the SAVTAP file or the module STAT file. $C_1, C_2, \cdots, C_k$ are integers separated by commas.
STATFILE	UNIT=n,FILE=m Default: n=20,m=1	n is the number of the Fortran logical unit to which the SAVTAP file has been assigned; m is the number of the file from which the training statistics are to be retrieved. (NOTE: If the module STAT file is input, m is the number of the file on which the statistics are to be stored. If $m \neq 1$, this control card must precede the module STAT file in the control card deck setup.)
MODULE	Blank	Indicates that the module STAT deck immediately follows. The SAVTAP file will be written as this card file is read.

TABLE 10-1.- Continued.

Keyword	Parameter and default values	Function
PROCEDURE	N Default: N=2	N=1: the exhaustive search procedure is used; N=2: the without replacement procedure is used; N=3: the Davidon-Fletcher-Powell procedure is used; N=4: the user-input B-matrix is used to evaluate the separability measure; N=5: the evaluate channels procedure is used; N=6: selects best subset of passes.
INCLUDE	$C_1, C_2, \cdots, C_k$ Default: None	Includes channels $C_1, C_2, \cdots, C_k$ in the best set; meaningful only for the without replacement procedure. $C_1, C_2, \cdots, C_k$ must be a subset of the channels on the CHANNELS card.
ICOUNT	N Default: N=300	Number of iterations for the Davidon-Fletcher-Powell procedure.
BSPASS	N Default: None	N is the number of passes (Landsat acquisitions) to be included in the best set.
NCPASS	N Default: N=4	Number of channels per pass.

122

TABLE 10-1.- Continued.

Keyword	Parameter and default values	Function
CRITERION	N Default: N=1	The indicated criterion is used to measure the separability between subclasses. N=1 for weighted average divergence; N=2 for weighted transformed divergence; and N=3 for weighted average Bhattacharyya distance.
EVALUATE	$C_1, C_2, \cdots, C_k$ Default: None	Evaluates the separability measure indicated on the CRITERION card for channels $C_1, C_2, \cdots, C_k$. The set of channels to be evaluated must be (1) a subset of the channels on the CHANNELS card and (2) must be on one card. Several sets of channels may be input by using more than one EVALUATE card.
B-MATRIX	CARDS Default: None	Indicates that the B-matrix card deck immediately follows; results in the evaluation of the separability measure using the linear combinations defined by the B-matrix if the fourth procedure is indicated. If the Davidon-Fletcher-Powell procedure is indicated, the B-matrix will be used as a first guess.

TABLE 10-1.- Continued.

Keyword	Parameter and default values	Function
B-MATRIX	FILE Default: None	Indicates that a BMFILE has been written. Depending on the PROCEDURE card, the B-matrix on file will be used as an initial guess for the Davidon-Fletcher-Powell procedure or in evaluating the separability measure.
SUBCLASSES	$K_1, K_2, \cdots, K_i$ $i \leq$ number of subclasses on SAVTAP $\leq$ 60 Default: All subclasses on the SAVTAP file	Provides for use of only subclass $K_1, K_2, \cdots, K_i$ statistics for computation of the separability measure. $K_1, K_2, \cdots, K_i$ are integers representing the subclass numbers as they occur in the SAVTAP file.
GROUP	NAME,I,J,··· Default: No grouping; individual subclasses are used.	Groups the training subclasses I,J,···; pools their statistics; and assigns NAME as the group name. NAME may be any six characters. Integers I,J,··· must correspond to the subclasses as they occur in the module STAT deck or the SAVTAP file.

TABLE 10-1.- Continued.

Keyword	Parameter and default values	Function
WEIGHTS	C1=XX,(C1,C2)=YY, OTHERS=ZZ Default: All weights set to 1.0 for criteria 2 and 3. For criterion 1, weights for subclass pair (i,j) are $W_{ij} = e^{-D_{ij}/16}$, where D_{ij} is the divergence for subclass pair (i,j).	Sets to XX weights for all subclass pairs involving subclass C1; then sets to YY the weight for subclass pair (C1,C2); and finally sets to ZZ the weights for all subclass pairs not already weighted. Subclass names C1, C2, etc., must match a subclass name from the module STAT file or from the SAVTAP file, or a GROUP NAME.[c] See OPTION CLSWT for constraint.

[c]Consider the problem of selecting the channels that best separate wheat from nonwheat classes, where wheat is divided into the subclasses W1, W2, and W3, and nonwheat is divided into the subclasses NW1, NW2, NW3, and NW4. It is desirable to set all weights for pairs of subclasses in the same class to 0, whereas wheat/nonwheat subclass pair weights are set to 1. This can be accomplished by the following WEIGHTS control cards: W1=1., W2=1., W3=1.; (W1,W2)=0., (W1,W3)=0., (W2,W3)=0.; and OTHERS=0.

(1) W1=1. will set weights equal to 1 for the following subclass pairs: (W1,NW1), (W1,NW2), (W1,NW3), (W1,NW4), (W1,W2), and (W1,W3).
(2) W2=1. will set weights equal to 1 for the following subclass pairs: (W2,NW1), (W2,NW2), (W2,NW3), (W2,NW4), (W2,W3), and (W2,W1).
(3) W3=1. will set weights equal to 1 for the following subclass pairs: (W3,NW1), (W3,NW2), (W3,NW3), (W3,NW4), (W3,W1), and (W3,W2).
(4) (W1,W2)=0., (W1,W3)=0., and (W2,W3)=0. will reset these subclass pair weights equal to 0.
(5) OTHERS=0. sets all other subclass pair weights to 0; namely, (NW1,NW2), (NW1,NW3), (NW1,NW4), (NW2,NW3), (NW2,NW4), and (NW3,NW4).

TABLE 10-1.- Continued.

Keyword	Parameter and default values	Function
OPTION	CLSWT	The processor determines the class-subclass correspondence (after any grouping of subclasses by use of the GROUP control card) and assigns a weight of 1.0 to all interclass subclass pairs. Intraclass subclass pairs are given a weight of 0.0. (NOTE: The WEIGHTS control card remains available to allow the user to set weights for specific subclass pairs. If used, the input subclass pair weights override the processor-set subclass pair weights. The WEIGHTS OTHERS capability is not available when this option is exercised. If input, it is ignored by the processor.)
APRIORI	Blank	This card causes the weights assigned to subclass pairs to be modified by factors computed from the number of pixels in each subclass.
OPTION	STATS Default: No statistics printed	Prints a summary of the statistics for the subclasses and channels actually used in SELECT.

TABLE 10-1.- Concluded.

Keyword	Parameter and default values	Function
OPTION	RUN Default: Exit the processor	This card allows the user to complete a run of this processor for the case in which the input SAVTAP file has only one class. Such a run is not applicable when the fourth or fifth PROCEDURE option is being taken.

Ancillary cards

HED1, HED2, DATE, COMMENT (see table 3-1)

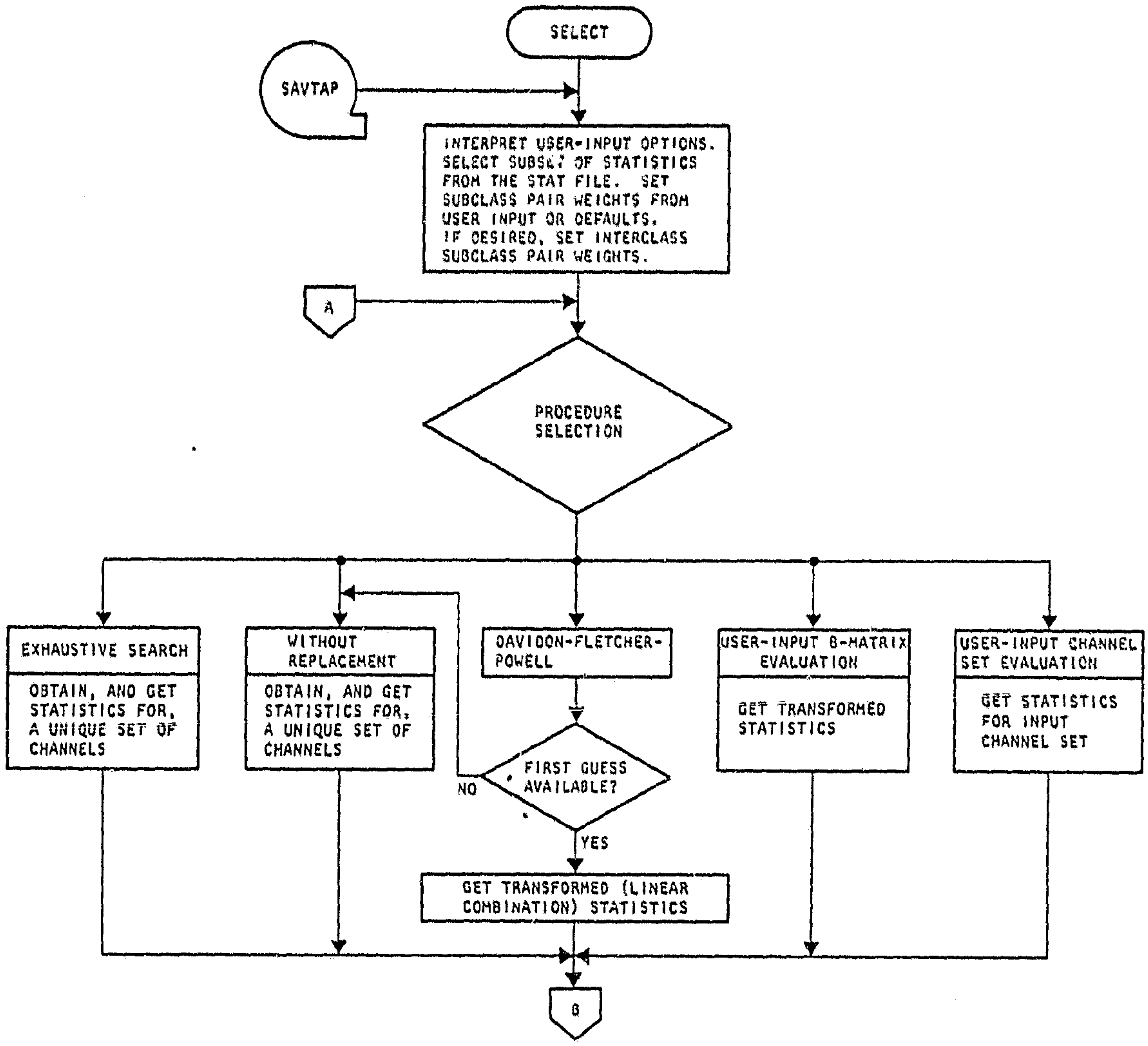


Figure 10-1.— Functional flow chart for the SELECT processor.

128

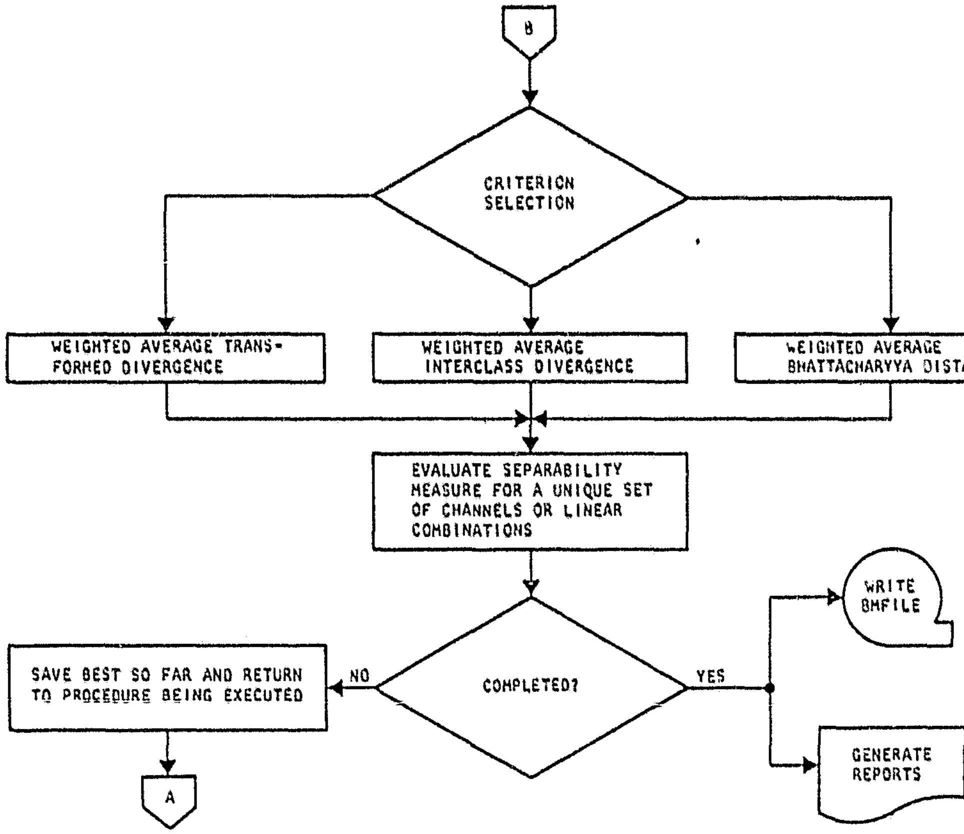


Figure 10-1.– Concluded.

11. CLASSIFICATION PROCESSOR — CLASSIFY

The classification processor, CLASSIFY, classifies MSS image data on the basis of statistics (mean vectors and covariance matrices) computed from training fields or labeled clusters.

11.1 PROCEDURES

Using the statistics for each subclass of interest, the processor assigns each data point within the defined classification field to a subclass. It does so by one of two procedures.

In the first procedure, the user does not define categories in his or her input, and the standard m-class maximum likelihood classification rule is followed. To decrease the number of times the density function must be computed, the classification-by-thresholding procedure proposed by Hallum and Minter (ref. 9) and improved upon and implemented by Feiveson (ref. 10) is used. The standard classification rule (i.e., when no categories are defined by the user) is outlined in section 11.1.1.

In the second procedure, the user defines categories in his or her input, and the sum-of-normal-densities classification rule is followed, as set out in section 11.1.2.

11.1.1 STANDARD M-CLASS CLASSIFICATION

Assuming multivariate normal probability density functions and using the maximum likelihood classification rule, the data vector $X^T = (X_1, X_2, X_3, \cdots, X_N)$ is assigned to subclass i in the following manner.

The assumed joint probability density function of X is given in the following equation.

$$P_i(X) = \frac{a_i}{(2\pi)^{N/2}|K_i|^{1/2}} e^{-0.5Q_i(X)} \tag{11-1}$$

where

a_i = a priori probability for subclass i

N = number of channels used for classification

K_i = covariance matrix for subclass i

Q_i = Gaussian quadratic form = $(X - \mu_i)^T K_i^{-1}(X - \mu_i)$

X = data vector $(X_1, X_2, X_3, \cdots, X_N)$

μ_i = mean vector for subclass i

Because of the exponential form of P_i and because $\ln(P_i)$ is a monotonically increasing function of P_i, for computational purposes it is convenient to define a new function V_i by

$$V_i = \ln(P_i) = \ln(a_i) - \frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln|K_i| - \frac{1}{2}(X - \mu_i)^T K_i^{-1}(X - \mu_i) \tag{11-2}$$

The data vector X is classified as belonging to subclass i if $V_i > V_j$ for all $i \neq j$, where $j = 1,2,3,\cdots,n$ and n = number of subclasses.

The number of times the function V_i must be computed may be reduced by the use of thresholds; i.e., real numbers γ_{ij} (independent of X) such that

$$\left.\begin{array}{l} V_i(X) > \gamma_{ij} \text{ implies } V_i(X) > V_j(X) \\ \text{and} \\ V_j(X) > \gamma_{ij} \text{ implies } V_j(X) > V_i(X) \end{array}\right\} \tag{11-3}$$

where $i,j = 1,2,3,\cdots,n$ and $i \neq j$.

131

The values of γ_{ij} are computed from input statistics as part of the initialization of processor CLASSIFY. The utility of these thresholds is that, if $V_i(X) > \gamma_{ij}$, $V_j(X)$ need not be computed. Once the values for γ_{ij} have been determined, they may be used for each observation vector X.

11.1.2 SUM-OF-NORMAL-DENSITIES CLASSIFICATION

Also assuming multivariate normal probability density functions, the category classifier classifies the data vector $X^T = (X_1, X_2, X_3, \cdots, X_N)$ to category r and subclass i in the following manner.

The probability density function for each category r is computed by the following equation.

$$P_r(X) = \sum_{i=1}^{M_r} \frac{a_i}{(2\pi)^{N/2}|K_i|^{1/2}} e^{-0.5Q_i(X)}$$

$$= \sum_{i=1}^{M_r} P_{ri}(X) \qquad (11\text{-}4)$$

where

- i = subclass number
- r = category number
- M_r = number of subclasses in category r
- a_i = a priori probability for subclass i in category r
- N = number of channels used for classification
- $Q_i(X)$ = Gaussian quadratic form = $(X - \mu_i)^T K_i^{-1}(X - \mu_i)$

Having computed the probability density function for all categories, the data vector X is classified as belonging to category r if $P_r > P_s$, for all $r \neq s$, where $s = 1,2,3,\cdots,S$ and S = number of categories.

The data vector is classified as belonging to subclass i if the probability density function for subclass i in category r is such that $P_{ri} > P_{rm}$ for all $i \neq m$, where $m = 1,2,3,\cdots,M_r$. In the computation of P_r, if the value of the quadratic form $Q_i(X)$ is smaller than -88, the computer cannot store the computed value of $e^{Q_i(X)}$. Thus, $e^{Q_i(X)} = 0$ for $Q_i(X) \leq -88$. In the case of all $P_{si}(X) = 0$ for $s = 1,2,3,\cdots,S$, the data point will not be classified; it will be assigned to a null subclass.

When the line-printer map of the classified data is displayed, each data point is printed with the symbol representing the legitimate subclass to which the data point belongs, and the null subclass is printed with the blank symbol. Figure 11-1 gives the functional flow of the CLASSIFY processor.

11.1.3 PROCEDURE 1

For Procedure 1 applications, the CLASSIFY processor allows the option of obtaining subclass a priori values from subclass population data in the input statistics file, SAVTAP. It also allows the system to use the class names from SAVTAP as the assigned category names.

Both options are in addition to the usual capability of the analyst to input a priori probability values at the subclass, class, or category level via the APRIORI control card and to input category names via the CATEGORY control card.

11.2 INPUT FILES

An MSS data file must be input to the CLASSIFY processor. The assignment defaults to logical unit 11; however, by input of the DATA control card, the user may assign any available logical unit. (See table 4-1 for file assignments and section 3.1, Image Tapes, for further information on format.)

11.3 OUTPUT FILES

The classification results are output on the MAPTAP file (see appendix D), which is defaulted to logical unit 2. In the event of card input of the module STAT file, the statistics will be output on the SAVTAP file (see section 4.1).

11.4 SCRATCH FILES

The processor requires no scratch file.

11.5 CARD INPUT

All required input card types are described below.

11.5.1 PROCESSOR CARD

The processor keyword is left justified beginning in column 1, thus,

$CLASSIFY

This card directs the system monitor routine to load all routines used by the CLASSIFY processor.

11.5.2 SPECIAL SYSTEM FILES

The training statistics may be input by means of the module STAT file. The B-transformation matrix may be input by means of the

B-matrix file. The EOD-LARSYS deck formats are described in section 3.2. The use of card input is an option; normally, card image files are used.

11.5.3 CONTROL CARDS

Table 11-1 describes the complete set of control cards recognized by the CLASSIFY processor. Ordering of the sequence of processor control cards is unnecessary, with the exceptions that (1) the *END card must follow the last processor control card, (2) the $END card must follow the last field definition card for an area to be classified, and (3) the STATFILE control card must precede the input of the module STAT file in some cases.

11.5.4 FIELD DEFINITIONS

An area to be classified is defined for the classification processor by a field definition card (described in section 3.2.3), which contains the scan line and sample coordinates for the area over which classification is to be performed. At least one field definition card must be in the run deck immediately following the *END control card. As many field definition cards as there are areas to be classified may be input. The processor will classify each field in the order in which it is identified. It will print on the line printer the first 110 samples of the classification map. And, for each field classified, it will print any optional output prescribed by the control cards. The scan line and sample coordinates specified on the field definition card must be available on the input MSS data file.

11.6 CARD OUTPUT

The classification processor outputs no punched cards.

135

11.7 RESTRICTIONS

The system-related restrictions described in section 24, along with the following, apply to the CLASSIFY processor.
The category, class, and subclass relationship is as follows:

$$\begin{matrix}\text{number of} \\ \text{categories}\end{matrix} \leq \begin{matrix}\text{number of} \\ \text{classes}\end{matrix} \leq \begin{matrix}\text{number of} \\ \text{subclasses}\end{matrix} \leq 60 \qquad (11\text{-}5)$$

The size of the B-matrix cannot exceed 450 locations:

$$\begin{pmatrix}\text{number of linear} \\ \text{combinations}\end{pmatrix}\begin{pmatrix}\text{number of channels} \\ \text{in B-matrix}\end{pmatrix} \leq 450 \qquad (11\text{-}6)$$

The channels used in computing the B-matrix automatically replace the channels, if any, on the CHANNELS control card.

The difference between the largest sample number of the classification field and the smallest sample number of the classification field cannot exceed 1000.

Beginning with the smallest sample number of the classification field, only the next 110 samples are displayed on the line-printer map output by CLASSIFY, but the entire classified field is displayed on the line-printer map output by DISPLAY.

When applying the category classifier option, 12 500 storage locations are reserved by the data such that

$$(\text{points per scan line})(\text{number of channels}) \leq 12\ 500 \qquad (11\text{-}7)$$

When applying the standard classifier option, the table computed for the class-pair thresholding procedure shares this storage of 12 500 locations reserved for the data such that

$$\left[\left(\begin{matrix}\text{number of}\\ \text{subclasses}\end{matrix} - 1\right)\frac{\left(\begin{matrix}\text{number of}\\ \text{subclasses}\end{matrix} - 2\right)}{2} + \begin{matrix}\text{number of}\\ \text{subclasses}\end{matrix}\right]$$

$$+ (\text{points per scan line})\left(\begin{matrix}\text{number of}\\ \text{channels}\end{matrix}\right) \leq 12\ 500 \qquad (11\text{-}8)$$

11.8 DIAGNOSTIC MESSAGES

The diagnostic messages and the routines in which they appear are listed in appendix I.

TABLE 11-1.- CONTROL CARDS FOR CLASSIFY

Keyword (a)	Parameter and default values (b)	Function
		Required cards
MAPTAP	OUTPUT/UNIT=n,FILE=m (usually n=2,m=1)	Unit and file number of classification map. This must be the first card after $CLASSIFY.
*END	Blank	Signals the end of the control cards.
$END	Blank	Signals the end of all card input for this processor.
		Optional cards
DATA	UNIT=n,FILE=m Default: n=11,m=1	n is the number of the Fortran logical unit to which the MSS data file has been assigned; m-1 is the number of files to be skipped on the unit. For back-to-back execution of several processors, if using the same file number, only one DATA control card need be input.

[a]The keyword must be left justified in card columns 1 through 10.

[b]The parameter values are in card columns 11 through 72 (beginning in any column past 10).

TABLE 11-1.- Continued.

Keyword	Parameter and default values	Function
CHANNELS	STAT=$N_1,N_2,\cdots,N_k$, DATA=$M_1,M_2,\cdots,M_k$ $k \leq 30$ Default: (1) If executed back to back with SELECT, the channels selected by the SELECT processor are used. (2) If a B-matrix is input, the channels resulting from the matrix applicat on are used. (3) Otherwise, all channels in the training statistics are used.	$N_1,N_2,\cdots,N_k$ are the channel numbers (integers) from the SAVTAP file to be used in classification; $M_1,M_2,\cdots,M_k$ are the channel numbers (integers) from the MSS data file (DATAPE). The number of channels selected from SAVTAP and DATAPE must be equal.
STATFILE	UNIT=n,FILE=m Default: n=20,m=1	n is the number of the Fortran logical unit to which the SAVTAP file has been assigned; m-1 is the number of files to be skipped on the unit. (NOTE: If the module STAT file is input, m is the number of the file for storing the statistics. If m ≠ 1, this control card must precede the module STAT deck in the control card file setup.)

- 139

TABLE 11-1.- Continued.

Keyword	Parameter and default values	Function
SUBCLASSES	$K_1, K_2, \cdots, K_i$ $i \leq$ number of subclasses on SAVTAP $\leq$ 60 Default: All subclasses on the SAVTAP file	Provides for use of only subclass $K_1, K_2, \cdots, K_i$ statistics to classify the unknown data points. $K_1, K_2, \cdots, K_i$ are integers representing the subclass numbers as they occur in the SAVTAP file.
MODULE	Blank	Indicates to the processor that the training subclass statistics will be input on cards. The module STAT file must immediately follow this control card.
B-MATRIX	CARDS or FILE Default: No transformation of training subclass covariance matrices	Informs the processor that the B-matrix transformation is to be input and applied to the training subclass statistics before classification. If FILE is placed in the parameter field, the mode of B-matrix input will be from BMFILE; if CARDS is specified, the B-matrix card deck must immediately follow this control card. The channels resulting from the B-matrix transformation will be the channels used by the processor in classification.

TABLE 11-1.- Continued.

Keyword	Parameter and default values	Function
GROUP	SUBNAM,K_1,K_2,···,K_i $i \leq$ number of subclasses on SAVTAP ≤ 60 Default: No grouping of subclasses	K_i's are integer subclass numbers taken from the set of available training subclasses. The processor creates a new training subclass by combining the statistics of the training subclasses listed. The set of training subclasses to be used is renumbered by the processor. SUBNAM may be from one to four characters; it becomes the name of the new training subclass.
APRIORI	$A_1, A_2, \cdots, A_M$ or $N^*A_1, K^*A_{N+1}, \cdots, A_M$ $M \leq 60$ Default: If executing the standard classifier, each subclass is given an equal a priori value. If executing the category classifier, each category is given an equal a priori value which is divided equally among the subclasses in that category.	A priori probability values may be input by subclass, class, or category. N or K is the number of times the value is repeated, and the A_i's are decimal numbers such that $\sum_{i=1}^{M} A_i = 1.0$. M = number of training subclasses, training classes, or categories. If input by class or category, the setup routine will distribute the a priori values among the subclasses in the following manner.

141

TABLE 11-1.- Continued.

Keyword	Parameter and default values	Function
		By class = $\frac{\text{class a priori values}}{\text{number of subclasses in that class}}$ By category = $\frac{\text{category a priori values}}{\text{number of subclasses in that category}}$ The order in which the a priori probability values, A_i, are input must be the order in which the categories, classes, or subclasses were defined.
APRIORI	FILE Default: Subclass a priori values will not be computed from the statistics file, SAVTAP.	The subclass a priori probability values are computed using subclass or cluster point populations from the statistics file, SAVTAP.
CATEGORY	CATNAM/NAME_1,NAME_2,··· Default: If no categories are defined, the standard m-class classifier is applied.	Informs the processor that the CATEGORY classifier option will be taken and defines a category name (CATNAM) and the class names (NAME_i's) included in this category. All subclasses for a class are assigned to this category. CATNAM and

142

TABLE 11-1.- Continued.

Keyword	Parameter and default values	Function
		$NAME_i$ may be up to four characters, and $NAME_i$ must match a class name on the SAVTAP file. A slash (/) separates the category name from the class name.
		(NOTE: (1) Every class will be assigned to a category unless the class was eliminated by omitting all of its subclasses using the SUBCLASSES control card. (2) At least two categories must be defined. (3) Continuation of the list of class names in the category on another card is indicated by an asterisk after the last class name on that card. The next card should continue the list of class names in columns 11-72.)
CATEGORY	FILE Default: No categories are defined, and the standard classifier will be applied.	Initiates the assigning of the category names using the class names from the input statistics file, SAVTAP, and invokes the category classifier.

TABLE 11-1.- Concluded.

Keyword	Parameter and default values	Function
OPTION	STATS Default: No training subclass statistics printout	Training statistics will be printed out for each subclass, reflecting the B-matrix transformation, if any, and the Cholesky factorization of the covariance matrices.

Ancillary cards

HED1, HED2, DATE, COMMENT (see table 3-1)

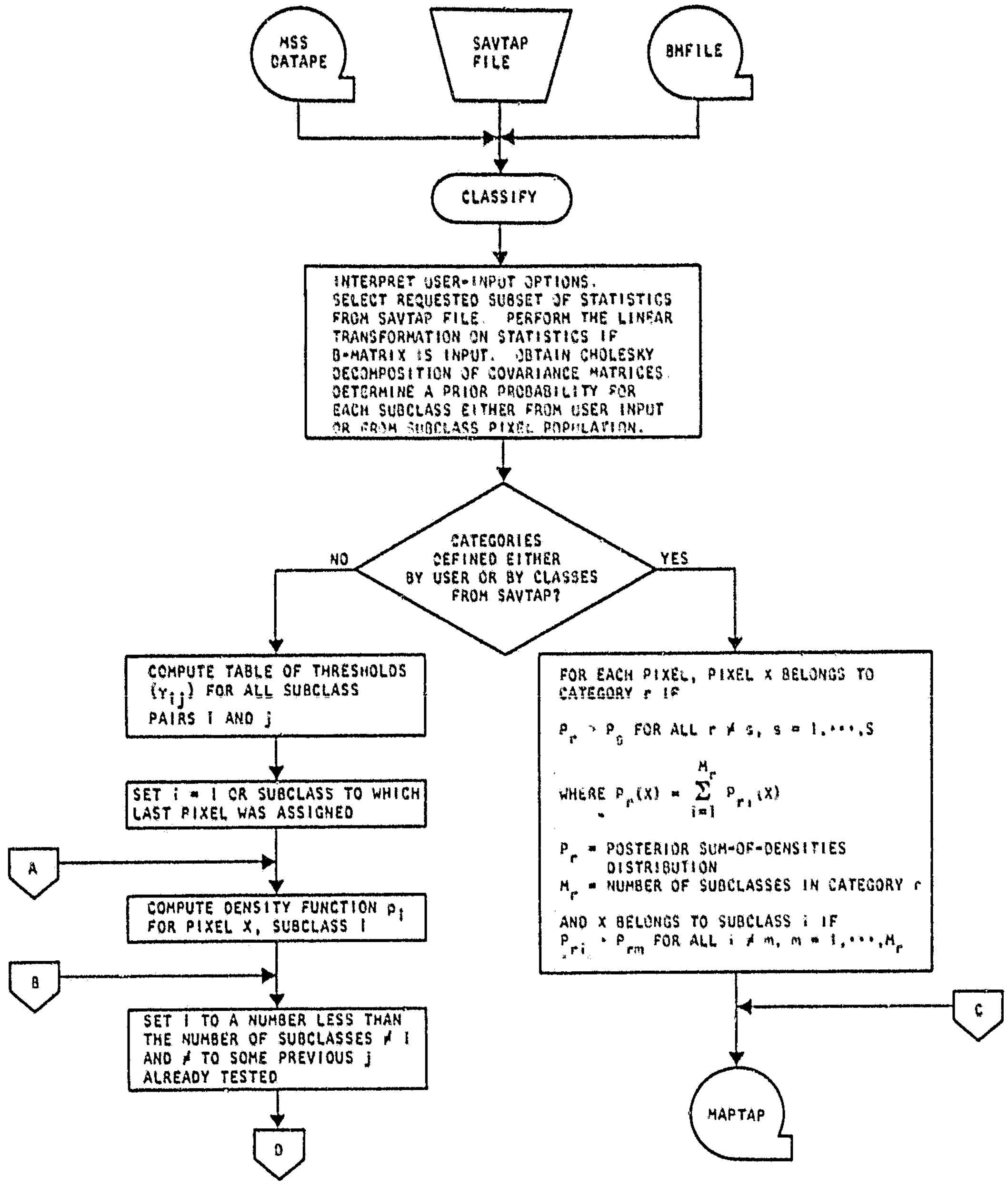


Figure 11-1.— Functional flow chart for the CLASSIFY processor.

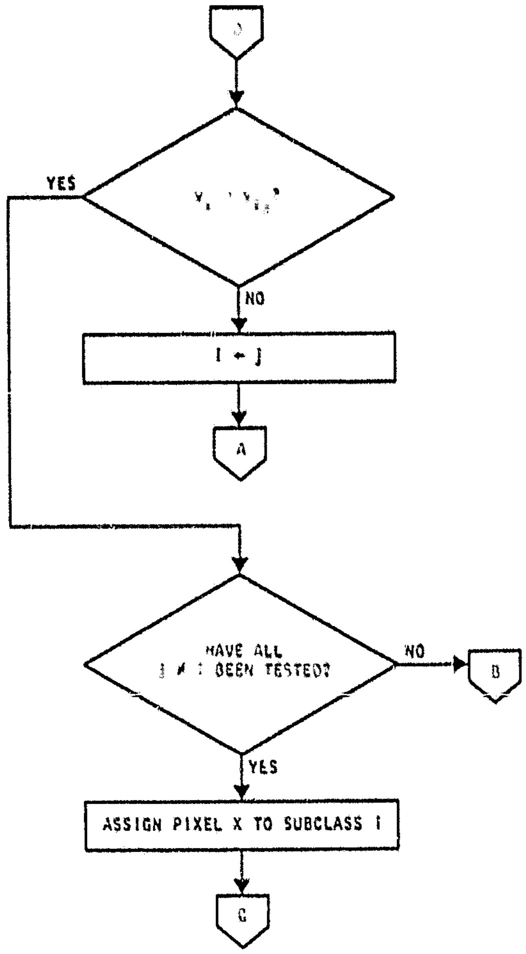


Figure 11-1.– Concluded.

146

12. DISPLAY OF CLASSIFICATION RESULTS PROCESSOR — DISPLAY

The DISPLAY processor reads the MAPTAP file output by CLASSIFY and performs the following functions.

a. Provides a line-printer map of each classified field on MAPTAP. The training and test fields within the classified image are outlined.

b. Produces a classification summary for each classified field, which gives a breakdown of the number of pixels classified into, and the number of pixels thresholded from, each subclass, class, and category.

c. Produces (optionally) an intensive test site (ITS) classification summary for one crop type versus all other crop types; the user-specified crop may be a category, class, or subclass.

d. Allows the user to designate fields to be excluded from the classification summaries. Fields may be designated "unidentifiable" or "other." Pixels within the unidentifiable fields are counted and are not considered in the classification summaries. Pixels within the designated "other" fields are counted as a separate crop type regardless of how they were classified. These pixels are included in category "other" in the ITS report. (See section 12.4.4 for sample input of DO/DU fields.) All pixels within the DO/DU areas are printed with the pound (#) symbol. Also, classification results can be overridden for fields designated as of a certain class by use of a designated [class name] card.

e. Assigns a pixel to the threshold class if thresholding is requested and if $Q_i > t_i$, where

Q_i = the value of the quadratic form $(X - \mu_i)^T K_i^{-1} (X - \mu_i)$ as computed by CLASSIFY (section 11.1.1) for subclass i

μ_i = mean vector for subclass i

PRECEDING PAGE BLANK NOT [illegible]

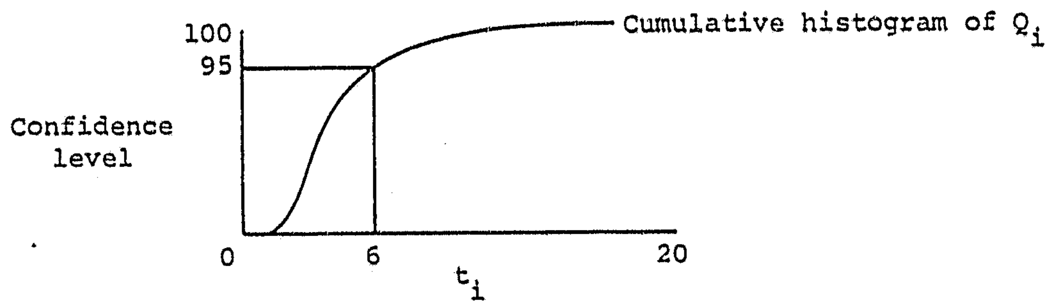
148

K_i = covariance matrix for subclass i

t_i = threshold value for subclass i

f. Allows t_i to be determined in one of four ways:

- User input — The user inputs the exact threshold value. See control cards THRESHOLD and OPTION THRESHOLD VALUE.
- Chi-square option — The user inputs confidence levels for each subclass on the THRESHOLD card and includes the OPTION CHI SQUARE. The program obtains the chi-square threshold value from an internal chi-square functional routine.
- Empirical option — The user inputs confidence levels for each subclass on the THRESHOLD card and includes the OPTION EMPIRICAL card. The program determines the empirical distribution function for each subclass from the cumulative histogram of Q_i for correctly classified pixels in the ground truth areas (i.e., training or test fields), as shown in the following example.



149

From this curve, the user input of a 95-percent confidence level for subclass i would result in a threshold value of 6.0. See reference 11 for more information on the use of empirically computed thresholds.

- Fisher F-distribution option — The user inputs confidence levels for each subclass on the THRESHOLD card and includes the OPTION FISHER card. The program obtains the F-distribution threshold values from an internal routine. If a computational overflow occurs in the routine, the threshold value for that subclass is set equal to 999.999.

g. Produces plots of the empirical distribution function when OPTION PLOT is exercised.

h. Performs (optionally) a four-nearest-neighbors spatial filtering on the classified image. This algorithm takes into consideration that, in many instances, a pixel is probably like its nearest neighbors. When the option is exercised via the OPTION FILTER control card, the four nearest neighbors of each pixel are examined. If all the neighbors are classified the same and the pixel in question is classified differently, then it is assumed that the pixel was classified incorrectly and its classification is changed. In the following example, the pixel classified as X will be changed to C. (See reference 12 for more information on this algorithm.)

Line 1		C	
2	C	X	C
3		C	

i. Outputs (optionally) the classified image (MAPUNT) onto tape or disk in either LARSYS II/III or Universal format via the FORMAT control card.

j. Provides classification performance summaries for ground truth areas within the classified image. The following six performance summaries are available to the user. The fields in these reports are either training fields used in the STAT or ISOCLS processor and transmitted to DISPLAY via the MAPTAP file, or test fields input directly to DISPLAY (see section 12.4.4).

- Field by subclass
- Field by class
- Field by category
- Class by subclass
- Class by class
- Class by category

k. In Procedure 1 applications, DISPLAY is able to

1. Accept a LACIE-formatted dot file (GTUNIT, PPUNIT, or AIUNIT).

2. Provide a dot classification performance summary by dot categories which also includes

   - A tabulation of both the uncorrected proportion and the bias-corrected proportion of each dot category in the total area classified.

   - A confusion table which tabulates proportions for each labeled category of bias correction (type 2) dots. For each category, the number of dots analyst labeled as belonging to that category and machine classified into that and each other possible category is compared to the total number of dots classified into each category.

3. Provide a dot classification performance summary for each dot on the analyst's specified file (GTUNIT, PPUNIT, or AIUNIT).

The functions of the DISPLAY processor are such that the analyst may either exercise the initial processor capabilities (a) through (j) or the LACIE Procedure 1 capabilities (k). The difference between the two capabilities is in the type and format of the classification performance tables output.

Figure 12-1 shows a function flow chart of the DISPLAY processor.

12.1 INPUT FILES

The only input file required for DISPLAY is the MAPTAP file (section 4.4) output by CLASSIFY. This file must be assigned to logical unit 2.

For Procedure 1, the DISPLAY processor accepts as input a LACIE-formatted dot file (GTUNIT, PPUNIT, or AIUNIT) created by the DOTDATA processor. The file is assigned either to logical unit 19 or to a user-specified unit.

12.2 OUTPUT FILES

The DISPLAY processor will optionally generate a tape or disk file of the classified image on the MAPUNT for display. The control card FORMAT allows the user to exercise this option. The tape is usually assigned to a nine-track tape drive for compatibility with display devices. The file assignment must be made to logical unit 16.

For Procedure 1, the DISPLAY processor optionally provides dot data classification performance summaries (instead of the normal output classification summaries described under j above) if the GTUNIT, PPUNIT, or AIUNIT control card is input to the DISPLAY processor.

12.3 SCRATCH FILES

The disk provides random access storage for a scratch file in DISPLAY. No assignment is necessary.

12.4 CARD INPUT

All system formats referred to in this section are defined in section 3.

12.4.1 PROCESSOR CARD

The processor keyword is left justified starting in column 1; the parameter FILE is punched starting in column 11, thus,

$DISPLAY FILE=N

This card directs the system monitor routine to select the DISPLAY processor and initiates the loading of routines used by DISPLAY. Parameter value N is the file number on the MAPTAP file to be processed; if not input, default is to file 1 of MAPTAP.

12.4.2 SPECIAL SYSTEM FILES

No special files are required for the DISPLAY processor.

12.4.3 CONTROL CARDS

Table 12-1 lists the control cards and available options for the DISPLAY processor.

12.4.4 CLASS, SUBCLASS, AND FIELD DEFINITIONS

Both test and designated fields are optional input to DISPLAY. If no test fields are input, the ground truth summaries will be for training fields. When input, test fields must be identified with a previously defined class or subclass. All test class, subclass, and field definitions begin immediately following the *END control card and are terminated by the $END control card.

Formats for the CLASS, SUBCLASS, and field definition cards are defined in section 3.2.3. The following example shows test field input to DISPLAY. Note that test fields are identified by classes; that is, each NAME1, NAME2, NAME3, etc., must match the name of a class defined in either STAT or ISOCLS and used in CLASSIFY.

```
$DISPLAY
(Control cards)
*END
CLASS      NAME1
Field 1    Vertices
Field 2    Vertices
CLASS      NAME2
Field 3    Vertices
Field 4    Vertices
Field 5    Vertices
CLASS      NAME3
Field 6    Vertices
$END
```

In the following example, test fields are identified with subclasses, in which case each NAME1, NAME2, NAME3, etc., must match the name of a subclass used in CLASSIFY.

```
$DISPLAY
(Control cards)
*END
SUBCLASS  NAME1
(Test fields for subclass NAME1)
SUBCLASS  NAME2
(Test fields for subclass NAME2)
$END
```

Note that actual name must not exceed four characters.

Designated fields are large areas within the classified area which are unidentifiable or can be specifically identified as being other than the crop type of interest. This type of field input is meaningful only when the ITS summary report is being generated for one specific crop type. Pixels within unidentifiable areas are removed from the summaries altogether. Pixels within the designated "other" areas are counted as other regardless of how they were classified.

An example of input designated fields follows.

```
$DISPLAY
(Control cards)
*END
DESIGNATE UNIDE
(Field definitions)
DESIGNATE OTHER
(Field definitions)
$END
```

If Procedure 1 is to be carried out, the only kinds of fields that should be input are DO/DU fields. Test fields should not be input; the counterpart of test fields are the bias correction dots on the LACIE-formatted dot file. The format of the DO/DU field cards and the method of input are given in section 3.2.3.

Another usage of designated fields is the specification of a class name on the DESIGNATE card image. This causes the classification of pixels in subsequent field definitions to be overridden as far as performance summaries are concerned.

The user may input any combination of the three types of designated fields.

155

12.5 CARD OUTPUT

No cards are output by the DISPLAY processor.

12.6 RESTRICTIONS

The system-related restrictions given in section 24 apply to this processor.

12.7 DIAGNOSTIC MESSAGES

Diagnostic messages for this processor are presented in appendix I.

TABLE 12-1.- CONTROL CARDS FOR DISPLAY

<table>
<tr><th>Keyword (a)</th><th>Parameter and default values (b)</th><th>Function</th></tr>
<tr><td colspan="3">Required cards</td></tr>
<tr><td>MAPTAP</td><td>INPUT/UNIT=n,FILE=m
(usually n=2,m=1)</td><td>Unit and file number of classification map.</td></tr>
<tr><td>*END</td><td></td><td>Signals the end of the control cards.</td></tr>
<tr><td>$END</td><td></td><td>Signals the end of all card input for this processor.</td></tr>
<tr><td colspan="3">Optional cards</td></tr>
<tr><td>SYMBOLS</td><td>$S_1,S_2,\cdots,S_k$
k = number of sub-classes on MAPTAP
Default: 1,2,···,9,
A,B,C,D,···,Z,−,|,¬,
/,−,*,+,$,",=,0,blank,
=,+,),(,:,&,>,;,?,H,
comma,period,blank,/</td><td>Assigns symbols $S_1,S_2,\cdots,S_k$ to subclasses 1,2,···,k, respectively.</td></tr>
<tr><td>OPTION</td><td>OUTLINE
Default: Training fields are not outlined.</td><td>Outlines training fields with asterisks; has no effect on test fields. (Test fields are always outlined with "+" symbol.)</td></tr>
<tr><td>OPTION</td><td>NOMAP
Default: Map printed</td><td>Instructs the processor not to print a map of the data; only a performance summary is printed.</td></tr>
</table>

[a]The keyword must be left justified in card columns 1 through 10.

[b]The parameter values are in card columns 11 through 72 (beginning in any column past 10).

157

TABLE 12-1.- Continued.

Keyword	Parameter and default values	Function
OPTION	STAT Default: No statistics printed	Prints statistics for subclasses used in the previous CLASSIFY run. These statistics are saved on MAPTAP.
CROP	NAME Default: No ITS report	Initiates the option of printing the ITS summary report for the crop indicated. Name must match a category, class, or subclass name used in CLASSIFY.
ACREAGE	TOTAL=X,CROP=Y, OTHER=Z	The total acreage in the ITS is X; the acreage of the crop named on the CROP control card is Y; and the acreage of all other crop types in the ITS is Z. X, Y, and Z are floating-point numbers. This input is meaningful only if the CROP control card is input.
SITE	Any 24 characters Default: Blanks	Name of the ITS; used in printing the heading for the ITS summary report.
PROCEDURE	Any 60 characters Default: Blanks	Procedure used in classification of ITS; printed in the heading for the ITS summary report.

TABLE 12-1.- Continued.

Keyword	Parameter and default values	Function
ANALYST	Any 18 characters Default: Blanks	Name of the data analyst; printed in the heading for the ITS summary report.
THRESHOLD	$T_1, T_2, \cdots, T_k$ k = number of subclasses on MAPTAP Default: No thresholding	Uses the threshold values $T_1, T_2, \cdots, T_k$ for subclasses $1, 2, \cdots, k$, respectively; thresholds must be positive floating-point numbers. One value must be specified for each subclass on the MAPTAP file. Thresholds may also be specified in the following format: $N_1{*}T_1, N_2{*}T_2, \cdots$ where N_1 and N_2 are integers which specify how many consecutive times the corresponding thresholds should be used. For the THRESHOLD VALUE option, the numbers input on the THRESHOLD cards are the actual values to be used for thresholding (i.e., T_1=10.02 means that the threshold value for subclass 1 is 10.02).

159

TABLE 12-1.- Continued.

Keyword	Parameter and default values	Function
		For the CHI SQUARE and FISHER options, the numbers are input on the confidence levels (i.e., T_1=0.99 means that the user wants to retain 99% or reject 1%). For the EMPIRICAL option, the numbers input on these cards are percentages.
OPTION	THRESHOLD VALUE Default:[c]	Uses the numbers input on the THRESHOLD control card for the actual threshold values.
OPTION	CHI SQUARE Default:[c]	Computes thresholds from the chi-square distribution using the confidence levels input on the THRESHOLD control card.
OPTION	EMPIRICAL Default:[c]	Computes the empirical threshold values using the percentages input on the THRESHOLD control card.

[c]If the THRESHOLD control card is input, one of the four options (CHI SQUARE, FISHER, EMPIRICAL, or THRESHOLD VALUE) should be input also. If the OPTION card is omitted and the THRESHOLD card is input, chi square is assumed. If more than one THRESHOLD option is input, only the last one read will be performed.

TABLE 12-1.- Continued.

Keyword	Parameter and default values	Function
OPTION	FISHER Default:[c]	Computes thresholds from the Fisher F-distribution using the confidence levels input on the THRESHOLD control card.
OPTION	PLOT	Plots the empirical distribution functions obtained from the cumulative histograms of Q_i for each subclass.
OPTION	FILTER Default: Spatial filtering is not performed.	Performs four-nearest-neighbors spatial filtering on the classified image.
OPTION	PCT Default: Performance summary printed for classes only	Prints a performance summary on a per-field as well as a per-class basis for ground truth fields (i.e., training or test fields within the classified image).

[c]If the THRESHOLD control card is input, one of the four options (CHI SQUARE, FISHER, EMPIRICAL, or THRESHOLD VALUE) should be input also. If the OPTION card is omitted and the THRESHOLD card is input, chi square is assumed. If more than one THRESHOLD option is input, only the last one read will be performed.

TABLE 12-1.- Continued.

Keyword	Parameter and default values	Function
FORMAT	NAME Default: No output classification map is generated by DISPLAY.	If NAME=UNIVERSAL, the output classification tape (MAPUNT) will be generated in the Universal format. If NAME=LARSYS, the MAPUNT tape will be generated in the LARSYS II/III format.
	Special cards, for Procedure 1 or LIST (Label Identification from Statistical Tabulation)	
GTUNIT	UNIT=n,FILE=m (usually n=23,m=1) Default: None	Unit and file number of ground truth dot file.
PPUNIT	UNIT=n,FILE=m (usually n=27,m=1) Default: None	Unit and file number of dot file labeled by the Patterson-Pitt-Thadani discrimination procedure or any LACIE-formatted labeled dot file.

TABLE 12-1.- Concluded.

Keyword	Parameter and default values	Function
AIUNIT	UNIT=n,FILE=m (usually n=28,m=1) Default: None	Unit and file number of AI-labeled dot file. (NOTE: Inclusion of any of the GTUNIT, PPUNIT, or AIUNIT cards initiates dot label comparison with machine classification results. In normal processing, the AIUNIT card would be present. This card replaces the DOTFIL card previously used in the DISPLAY processor.)
NAME	One-character name of selected category Default: S	Category name of the category of interest if LIST-type reports are needed.
ALPHA	F_1,F_2 Default: 0.5,0.5	Floating-point weighting factors used in bias correction reports.

Ancillary cards

HED1, HED2, DATE, COMMENT (see table 3-1)

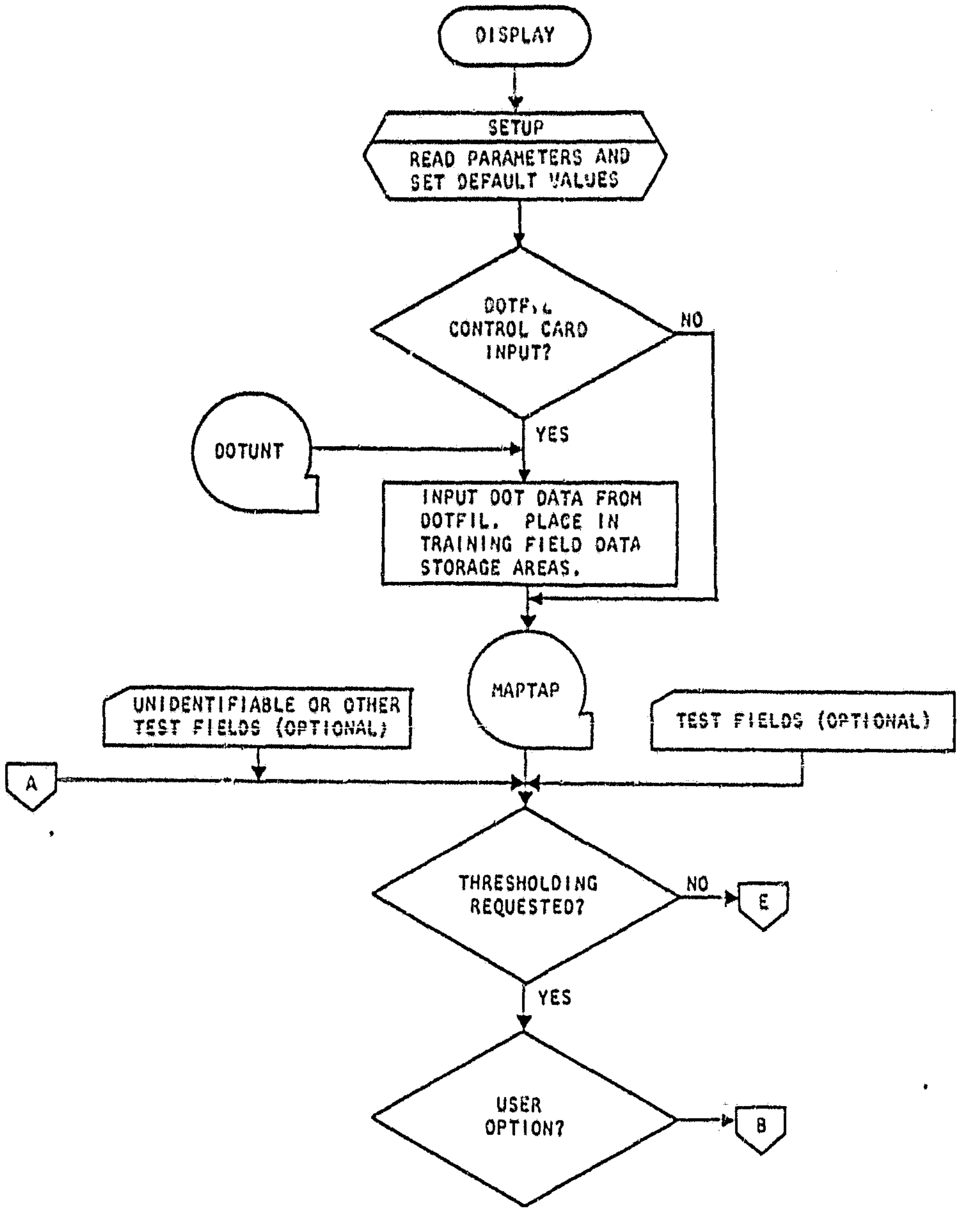


Figure 12-1.— Functional flow chart for the DISPLAY processor.

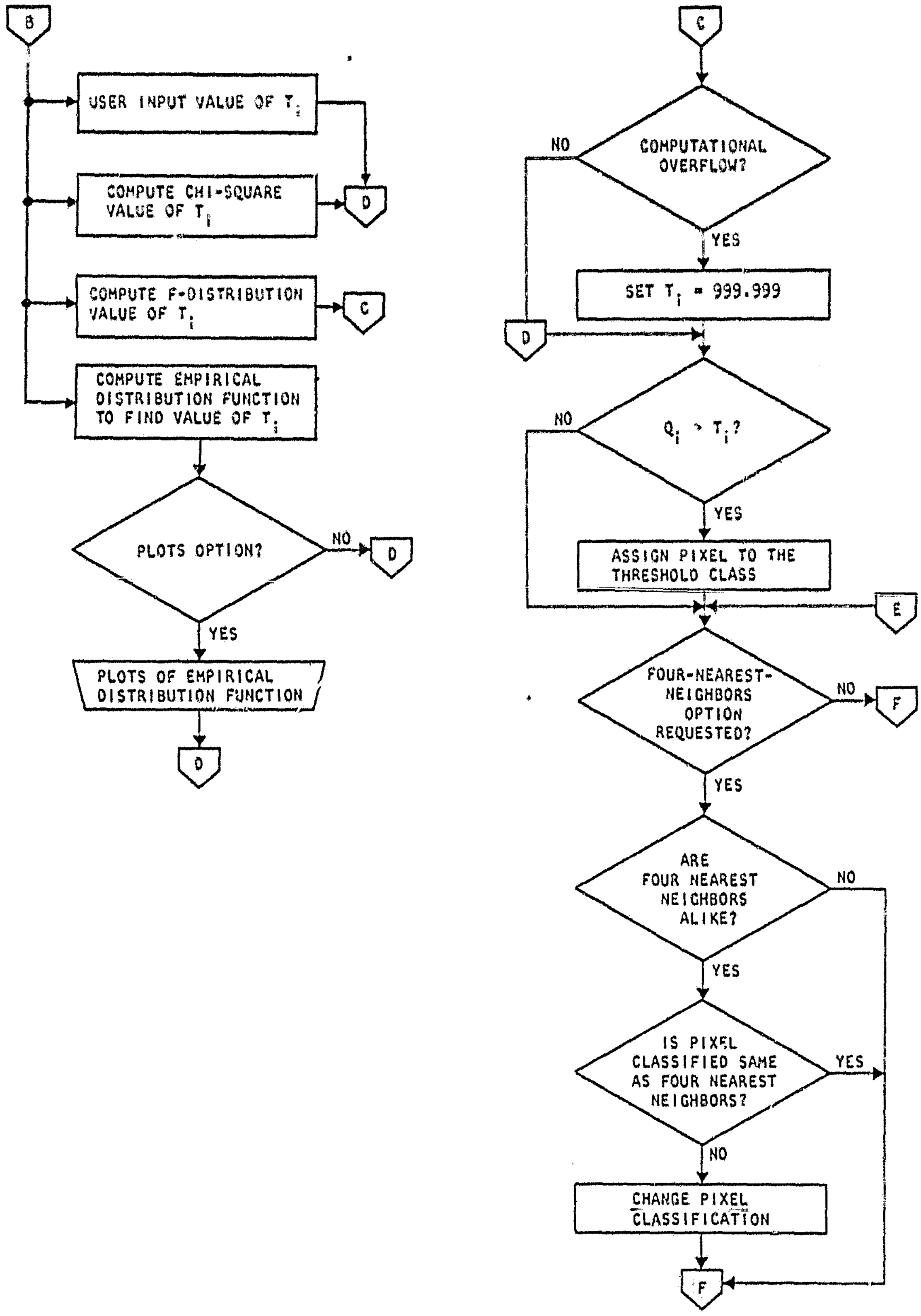


Figure 12-1.– Continued.

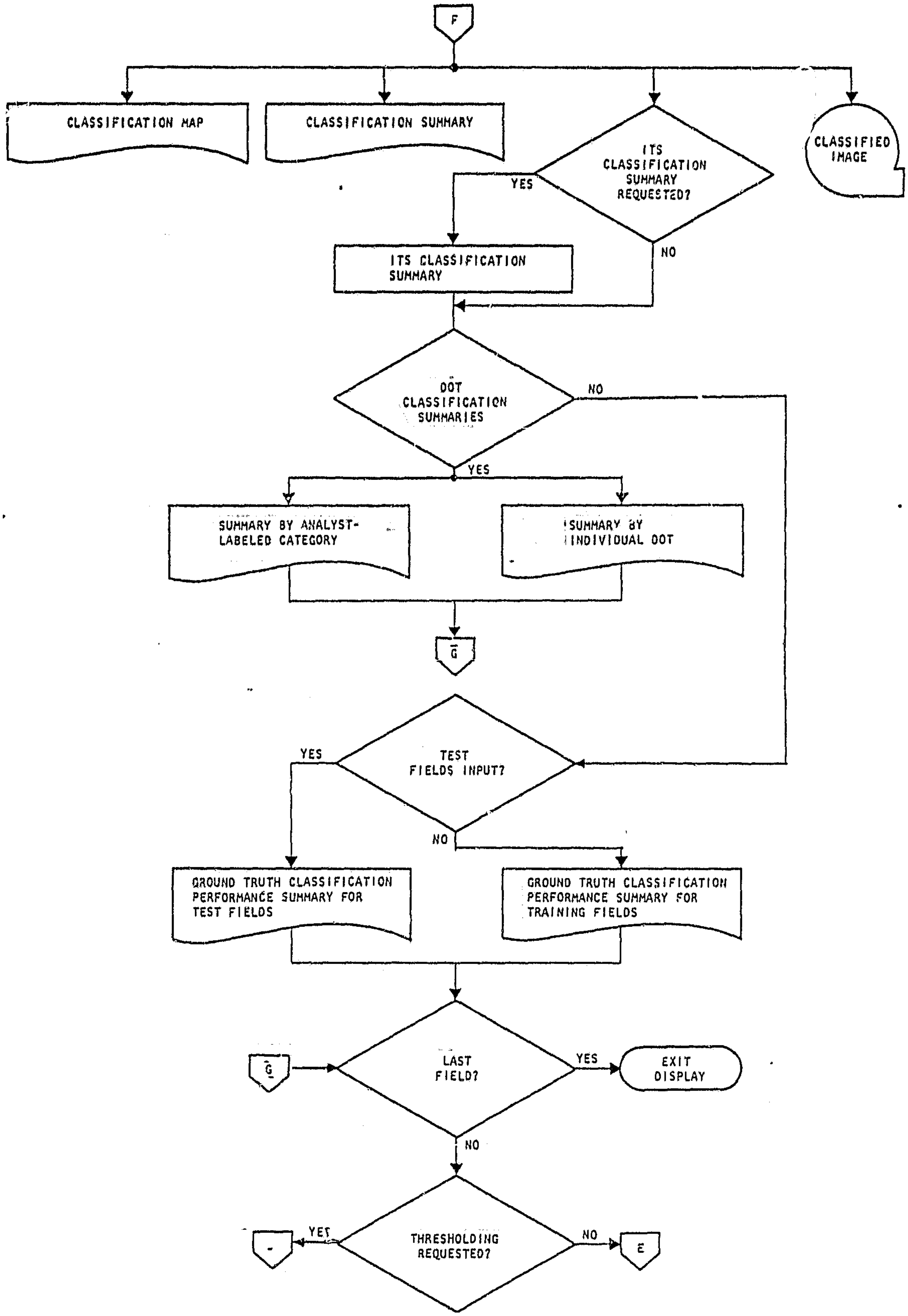


Figure 12-1.— Concluded.

166

13. DATA TRANSFORMATION PROCESSOR — DATATR

13.1 PROCEDURES

The DATATR processor transforms images from the MSS data file. The affine transformation is performed on user-defined fields according to the following formula:

$$\vec{z} = B\vec{x} + \vec{b} \qquad (13\text{-}1)$$

where

$\vec{z}$ = a k-by-1 transformed data vector

$k \leq 16$

B = a k-by-n input transformation matrix (see section 3.2.4.2)

$n \leq 30$

$\vec{x}$ = an n-by-1 data vector

$\vec{b}$ = a k-by-1 bias vector (see BIAS control card, table 13-1)

To determine the type of rescaling to be done on the transformed data, the user must provide the RESCALE control card. (Control cards are listed in table 13-1.) Otherwise, the transformed data will be clipped at 0 and 255. To rescale the data, the following equation is used.

$$Y_i = \frac{255}{R_i} \times |MIN_i - Z_i| \qquad (13\text{-}2)$$

where

Y_i = rescaled transformed data point for channel i

R_i = range of component i $[MAX_i - MIN_i]$

MAX_i = maximum value for component i

MIN_i = minimum value for component i

Z_i = transformed data point for channel i

The user may obtain the parameters R_i and MIN_i in one of three ways: the histogram method, the statistical method, or user input. Figure 13-1 shows the functional flow of the DATATR processor. The method and control cards associated with each method are defined below.

13.1.1 HISTOGRAM (DEFAULT) METHOD

A histogram of a segment of the transformed image is performed to find the R_i and MIN_i for each component of the transformed data. If the user-defined field is smaller than 2000 pixels, all pixels are used in the histogram; otherwise, the following formula is used to determine the line and sample increments needed to obtain 2000 points for the histogram.

$$\alpha = \left(\frac{MN}{2000}\right)^{1/2} \qquad (13\text{-}3)$$

where

α = increment (integer)

M = number of samples

N = number of lines

In deriving an approximate range for the transformed data, the user may specify a percentage of points to be excluded from the upper and lower tails of the histogram by using the PEROUT control card. If not so specified, 2.5 percent of the points on each end of the distribution are excluded when determining the MAX_i and MIN_i values of the central 95 percent of the transformed data distribution.

Optionally, the user may specify the maximum expected data value for each channel n of the input data vector $\vec{x}$. Otherwise, the maximum data value for each channel is set at 255.

13.1.2 STATISTICAL METHOD

Activated by the RESCALE and MODULE FILE (or MODULE CARDS) control cards, the statistical method is applied to derive an approximate MAX_i and MIN_i value for each component i. Using the subclass statistics, an approximate R_i is computed using equations (13-4) and (13-5).

Let
$$\alpha_i = MAX_j\left(\hat{\beta}_i^j + k\hat{\sigma}_i^j\right) \quad (13\text{-}4)$$

and
$$\delta_i = MIN_j\left(\hat{\beta}_i^j - k\hat{\sigma}_i^j\right) \quad (13\text{-}5)$$

where

i = 1,···,m components of z

j = 1,···,w subclasses

$\hat{\beta}_i^j$ = transformed mean of the i^{th} component of subclass j

k = an integer specified by the user (see LAM control card)

$\hat{\sigma}_i^j$ = standard deviation of the i^{th} component of subclass j computed from the transformed covariance matrix for subclass j

The approximate range of each component will be

$$R_i = \alpha_i - \delta_i \quad ; \quad i = 1,\cdots,m \quad (13\text{-}6)$$

and
$$MIN_i = \delta_i \quad (13\text{-}7)$$

The complete transformation, including rescaling, to be performed on each pixel of the original image is

$$Y_i = \frac{255}{R_i} \times \left|\delta_i - (B\vec{x} + b_i)\right| \quad (13\text{-}8)$$

where

b_i = i^{th} element of the bias vector $\vec{b}$

Optional control cards that may be used in conjunction with the statistical method are SUBCLASS, LAM, PEROUT, OPTION ORIG, and OPTION TRANSF. Their functions, as well as those of other control cards, are described in table 13-1.

13.1.3 USER-INPUT METHOD

The user may input his or her own scaling parameters by means of the OPTION SCAFAC control card or use input from a previous execution of DATATR in which the computed scaling parameters $\left[CON_i\left(=\frac{255}{R_i}\right), MIN_i\right]$ were punched on cards using the OPTION PUNCH control card.

The transformed or transformed and rescaled data are output in either Universal or LARSYS II/III format. The option is controlled by the FORMAT control card.

A line-printer plot of the histogram (frequency distribution of the transformed/rescaled data) is printed. If applicable, the MAX_i and MIN_i are printed.

13.2 INPUT FILES

An MSS data file must be input to the DATATR processor. The assignment defaults to logical unit 11; however, by input of the DATA control card, the user may assign any available logical unit. (See table 4-1 for file assignments and section 3.1, Image Tapes, for further information on format.)

13.3 OUTPUT FILES

The transformed or transformed and rescaled data are output on the TRFORM unit (default: 14) in either Universal or LARSYS II/III format.

13.4 SCRATCH FILES

The DATATR processor does not use scratch files.

13.5 CARD INPUT

All system card file input formats referred to in this section are defined in section 3.

13.5.1 PROCESSOR CARD

The keyword for the processor card is left justified beginning in column 1, thus,

$DATATR

This keyword directs the system monitor routine to select the DATATR processor and initiates loading of routines used by DATATR.

13.5.2 SPECIAL SYSTEM FILES

The B-matrix file discussed in section 3.2.4.2 must be input to this processor. The deck may be obtained from a previous execution of SELECT. If a statistics file is to be used, it can be either a SAVTAP file or a module STAT deck.

13.5.3 CONTROL CARDS

Table 13-1 lists the control cards and available options for the DATATR processor.

13.5.4 FIELD DEFINITIONS

See section 3.2.3 for the format of field definition cards. At least one field definition card must immediately follow the *END control card. An output file is created for each field definition input and is written on unit 14. Each of these fields consists of a rectangular field which encompasses the vertices of the input field. All pixels within the rectangular output field but outside the input field are set equal to zero. The lines and samples will be numbered consecutively from 1.

13.6 CARD OUTPUT

The DATATR processor, via the OPTION PUNCH control card, outputs the computed scaling parameters on cards. Two pairs of scaling parameters are punched on each card; i.e., each punched card contains the scaling parameters for two components of the transformed data. The cards must be used in the same order as punched. Their formats and definitions are as follows. The number of cards is determined by the number of components.

Columns	Format	Definition
1-6	A6	OPTION
11-17	A7	SCAFAC=
18-27	A1,F9.3,F9.3,A1	(CON_1,MIN_1) where CON_1=255/R_1, R_1 is the range of component 1, and MIN_1 is the minimum value for component 1. Parentheses must be input.
28-37	A1,F9.3,F9.3,A1	(CON_2,MIN_2) where CON_2=255/R_2, R_2 is the range of component 2, and MIN_2 is the minimum value for component 2. Additional pairs are recorded on succeeding cards.

13.7 RESTRICTIONS

The system-related restrictions in section 24 apply to this processor.

The maximum number of channels allowed is 30, and the maximum number of components in the transformed vector is 16.

13.8 DIAGNOSTIC MESSAGES

Diagnostic messages for the DATATR processor are listed by subroutine in appendix I.

TABLE 13-1.- CONTROL CARDS FOR DATATR

Keyword (a)	Parameter and default values (b)	Function
	Required cards	
B-MATRIX	CARDS or FILE	CARDS indicates that the B-matrix is on cards immediately following. FILE indicates that the B-matrix is on file and initiates input of the BMFILE.
*END	Blank	Signals the end of the control cards.
$END	Blank	Signals the end of all card input for this processor.
	Optional cards	
DATA	UNIT=n,FILE=n Default: n=11,m=1	n is the number of the Fortran logical unit to which the MSS data file has been assigned; m-1 is the number of files to be skipped on the unit.

[a]The keyword must be left justified in card columns 1 through 10.

[b]The parameter values are in card columns 11 through 72 (beginning in any column past 10).

174

TABLE 13-1.- Continued.

Keyword	Parameter and default values	Function
BIAS	$b_1, b_2, \cdots, b_M$ $N^*b_1, Kb_{N+1}, \cdots, b_M$ $M \leq 16$ Default: b_i=0.0	All b_i's are decimal (floating-point) numbers, separated by commas; they comprise the bias vector to be applied to the transformation of the input data set: $\vec{z} = B\vec{x} + \vec{b}$ N or K is the number of times a certain b_i is repeated. M is the number of components in the transformed data set.
RESCALE	Blank	Initiates rescaling of the transformed data to the range of 0 to 255. If not present, the data are simply clipped at 0 and 255.
MAXPT	$M_1, M_2, \cdots, M_k$ $k \leq 30$ Default: 255,255,...	Maximum expected value of MSS data for each channel. M's are integers used in deriving an approximate range (MIN_i, MAX_i) of the transformed data set for the histogram method of rescaling.

TABLE 13-1.- Continued.

Keyword	Parameter and default values	Function
PEROUT	N Default: N=5	An integer which specifies the percentage of points to be deleted from the extremes of the transformed data distribution in computing an approximate range for rescaling. For the histogram method of rescaling, N/2 percent is deleted from each of the tails of the histogram, leaving the central 95 percent. For the statistical method of rescaling, N percent is deleted from each of the tails of the distribution, leaving the central 90 percent.
MODULE	CARDS Default: If RESCALE is input, the histogram method is assumed.	Initiates reading of the module STAT file that must immediately follow this card; if rescaling is performed, it initiates the statistical method.
MODULE	FILE Default: If RESCALE is input, the histogram method is assumed.	Initiates reading of the SAVTAP file; if rescaling is performed, it initiates the statistical method.

TABLE 13-1.- Continued.

Keyword	Parameter and default values	Function
STATFILE	UNIT=n,FILE=m Default: n=20,m=1	n is the number of the Fortran logical unit to which the SAVTAP file has been assigned; m-1 is the number of files to be skipped on the unit. (NOTE: If a module STAT file is input, m is the number of the file on which to store the training statistics. If $m \neq 1$, this control card must precede the module STAT deck.)
SUBCLASSES	$S_1, S_2, \cdots, S_k$ $k \leq$ number of subclasses on SAVTAP ≤ 60 Default: Statistics for all subclasses defined are used in calculating the scaling factors.	Integers which define a subset of subclasses $S_1, S_2, \cdots, S_k$ from the input statistics file (SAVTAP) to be used in calculating the scaling factors and approximating R_i.
LAM	N Default: N=2	An integer multiplied by the standard deviations of the input subclass statistics to derive an approximate range for rescaling the transformed data.

TABLE 13-1.- Continued.

Keyword	Parameter and default values	Function
OPTION	ORIG Default: No statistics printout	Initiates the printout of the original (untransformed) statistics for the subclasses input for the statistical rescaling method.
OPTION	TRANSF Default: No statistics printout	Initiates the printout of the transformed statistics.
OPTION	SCAFAC=(CON_1,MIN_1),(CON_2,MIN_2),···,(CON_i,MIN_i) Default: Histogram method of rescaling	CON and MIN are floating-point values separated by a comma. Blanks between the two values are ignored. The scaling parameters should be ordered according to the transformed data vector components.
OPTION	PUNCH Default: No cards punched	Directs the program to punch the scaling parameters (CON_i,MIN_i) on cards.
TROUT	UNIT=n,FILE=m Default: None	n is the number of the Fortran logical unit assigned to TRFORM; m-1 is the number of files to skip before writing TRFORM.
FORMAT	OUTPUT=UNIVERSAL Default: LARSYS II/III	The transformed data will be output in Universal format.

TABLE 13-1.- Concluded.

Keyword	Parameter and default values	Function
FORMAT	OUTPUT=LARSYS Default: LARSYS II/III	The transformed data will be output in LARSYS II/III format.

Ancillary cards

HED1, HED2, DATE, COMMENT (see table 3-1)

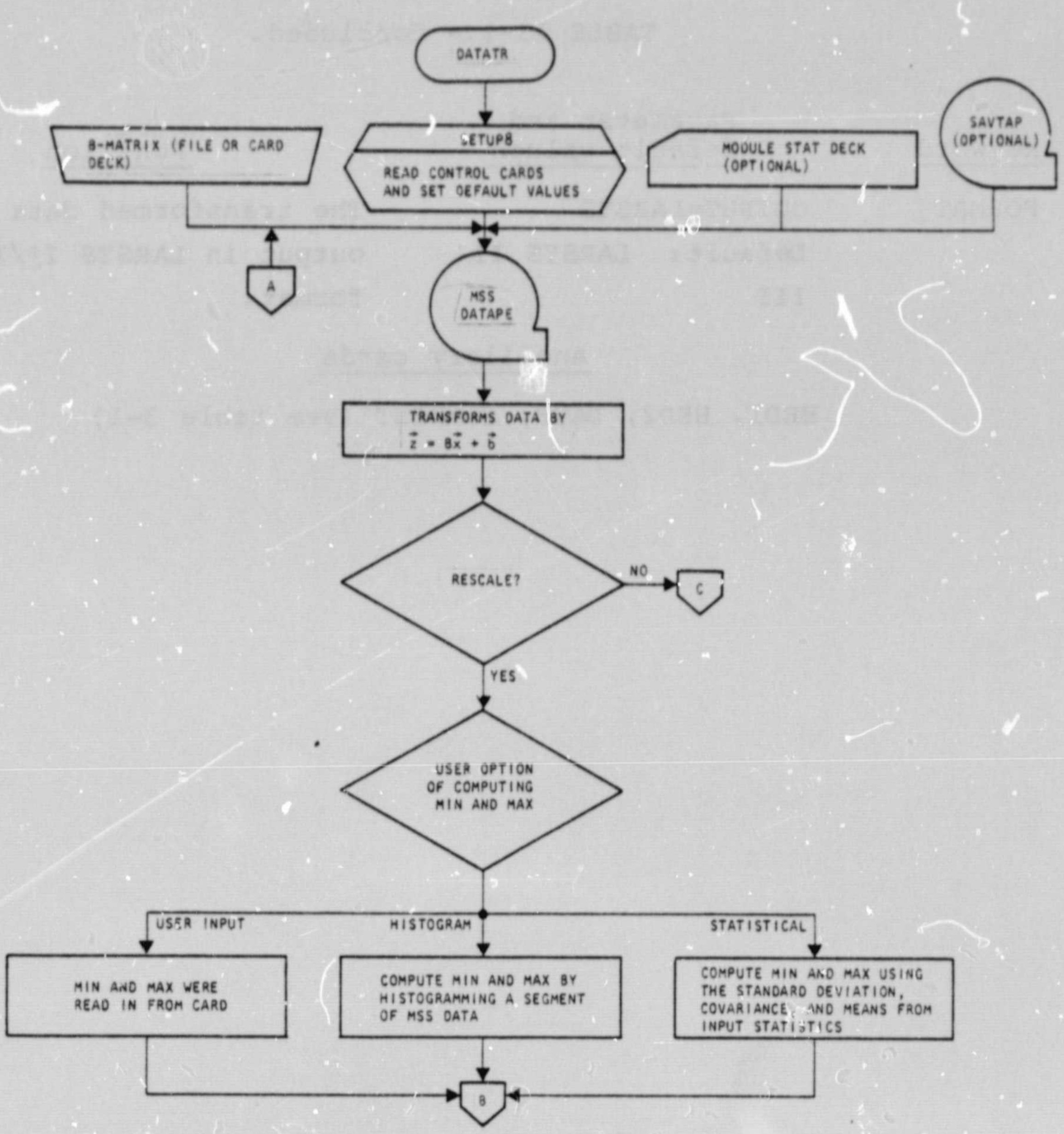


Figure 13-1.– Functional flow chart for the DATATR processor.

180

C-3

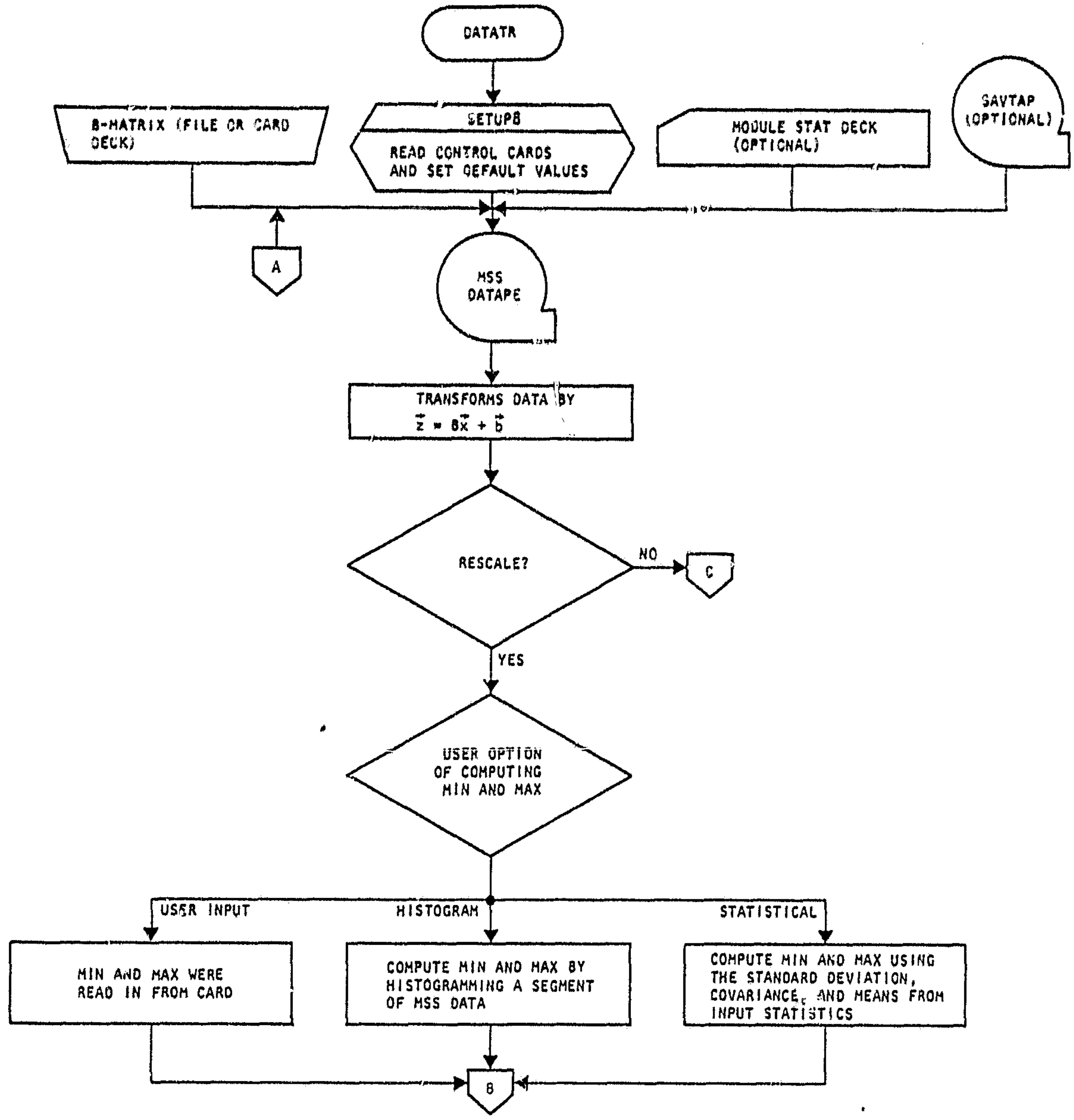


Figure 13-1.— Functional flow chart for the DATATR processor.

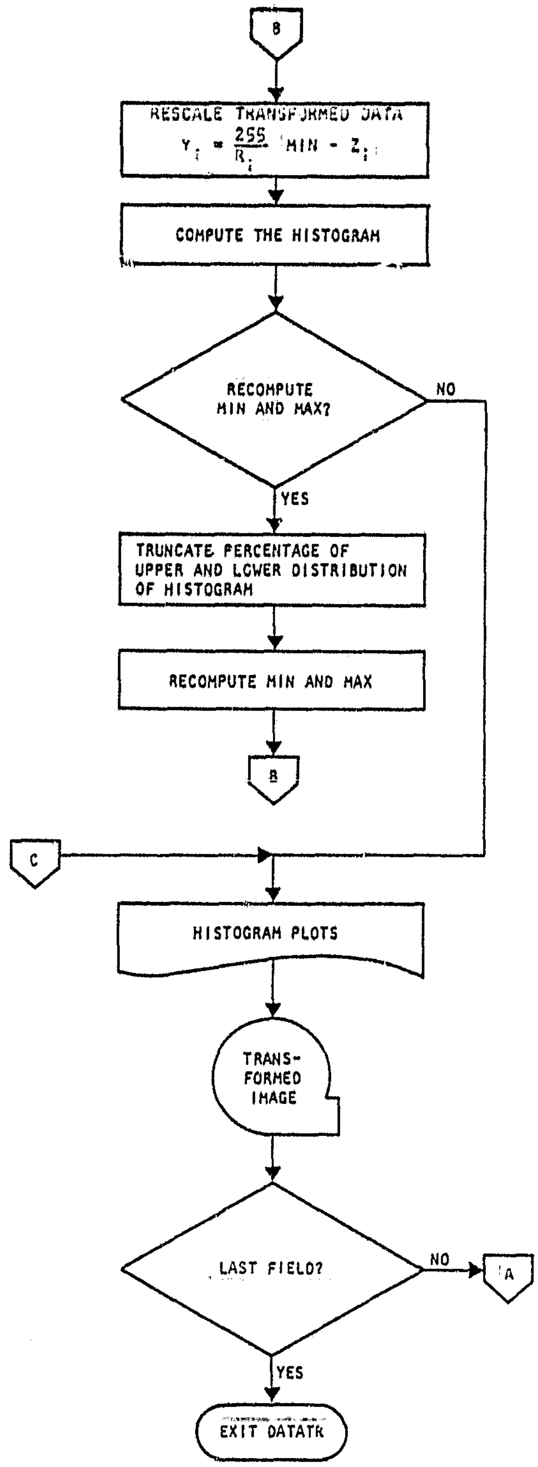


Figure 13-1.— Concluded.

181

14. STATISTICS TRANSFORMATION PROCESSOR — TRSTAT

The TRSTAT processor will read a SAVTAP file or card deck generated by STAT or ISOCLS, perform a linear transformation on the means and covariances, and output the transformed statistics on a new file (fig. 14-1). The equation for the linear transformation of the means is as follows:

$$\mu' = A\mu + b \tag{14-1}$$

where

μ' = a k-by-1 transformed mean vector

$k \leq 16$

A = a k-by-n matrix (see section 14.4.2)

$n \leq 30$

μ = an n-by-1 mean vector

b = a k-by-1 bias vector (see card type 4, section 14.4.2)

The equation for the linear transformation of the covariances is as follows:

$$K' = AKA^T \tag{14-2}$$

where

K' = a k-by-k transformed covariance matrix

K = an n-by-n covariance vector

A^T = an n-by-k transpose of A

14.1 INPUT FILES

A set of statistics must be input either from the SAVTAP file or by cards. (See STATFILE or MODULE control card, table 14-1.)

14.2 OUTPUT FILES

The transformed statistics are output on a file in the SAVTAP format. (See section 4.1 and STATFILE control card.)

14.3 SCRATCH FILES

The TRSTAT processor uses no scratch files.

14.4 CARD INPUT

All system formats referred to in this section are defined in sections 3 and 14.4.2.

14.4.1 PROCESSOR CARD

The processor keyword is left justified starting in column 1, thus,

$TRSTAT

This card directs the system monitor routine to execute the TRSTAT processor and initiates loading of routines used by TRSTAT.

14.4.2 A-MATRIX FILE

The A-matrix file is composed of a transformation matrix and an additive bias vector. Its format is shown below. For additional information on the transformation matrix, see section 3.2.4.2.

Card type	Columns	Format	Definition
1	1-8	A8	Keyword A-MATRIX
2	6-7	I2	Number of linear combinations
	13-14	I2	Number of channels
	17-80	I2	Actual channels used combination*

*The channels in the B-matrix described in section 3.2.4.2 begin in column 18.

183

Card type	Columns	Format	Definition
3	6-20	E15.8	Element 1 of A-matrix (column 1, row 1)
	21-35	E15.8	Element 2 of A-matrix (column 1, row 2)
	⋮		⋮
	60-80	E15.8	Element 5 of A-matrix (Five values are entered on each card until the full matrix has been entered.)
4	6-20	E15.8	Element 1 of b-vector*
	21-35	E15.8	Element 2 of b-vector
	⋮		⋮
	66-80	E15.8	Element 5 of b-vector (Five values are entered on each card until the computer vector of N linear combinations has been entered.)

14.4.3 CONTROL CARDS

Table 14-1 lists the control cards and available options for the TRSTAT processor.

14.4.4 FIELD DEFINITIONS

No field definition cards are input to TRSTAT.

14.5 CARD OUTPUT

The transformed statistics deck will be output in the same format as the module STAT card file.

*Unlike the B-matrix described in section 3.2.4.2, the A-matrix file contains the additive vector.

14.6 RESTRICTIONS

The system-related restrictions in section 24 apply to this processor.

The maximum dimension of the A-matrix is 16 by 30, and the maximum number of elements in the additive b-vector is 30.

14.7 DIAGNOSTIC MESSAGES

Diagnostic messages for the TRSTAT processor are presented in appendix I.

186

TABLE 14-1.- CONTROL CARDS FOR TRSTAT

Keyword (a)	Parameter and default values (b)	Function
	Required cards	
CHANNELS	$N_1, N_2, \cdots, N_k$ k = number of matrix channels $\leq$ 30	N's are integer channel numbers referring to the SAVTAP file. The number of channels requested from SAVTAP must be equal to the number of channels on the A-matrix file.
A-MATRIX	Blank	Initiates input of the A-matrix and b-vector. The A-matrix card images immediately follow this card.
*END	Blank	Signals the end of the control cards.
$END	Blank	Signals the end of all card input for this processor.

[a]The keyword must be left justified in card columns 1 through 10.

[b]The parameter values are in card columns 11 through 72 (beginning in any column past 10).

TABLE 14-1.- Continued.

Keyword	Parameter and default values	Function
	Optional cards	
STATFILE	INPUT/UNIT=n,FILE=m OUTPUT/UNIT=ℓ,FILES=s Default: n=20,m=1, ℓ=20,s=1	n is the number of the Fortran logical unit to which the file containing the statistics to be transformed has been assigned; m-1 is the number of files to be skipped on the unit; ℓ is the number of the Fortran logical unit to which the transformed statistics are to be output; and s-1 is the number of files to skip on the unit before writing the SAVTAP file.
MODULE	Blank	Initiates input of the module STAT deck, which immediately follows this card.
SUBCLASSES	$S_1,S_2,\cdots,S_k$ $k \leq$ number of subclasses on SAVTAP ≤ 60 Default: Statistics for all subclasses defined	Transforms statistics for only subclasses $S_1,S_2,\cdots,S_k$.

187

TABLE 14-1.- Concluded.

Keyword	Parameter and default values	Function
OPTION	P,O,T	P punches the transformed statistics; O prints the original statistics; and T prints the transformed statistics.

Ancillary cards

HED1, HED2, DATE, COMMENT (see table 3-1)

188

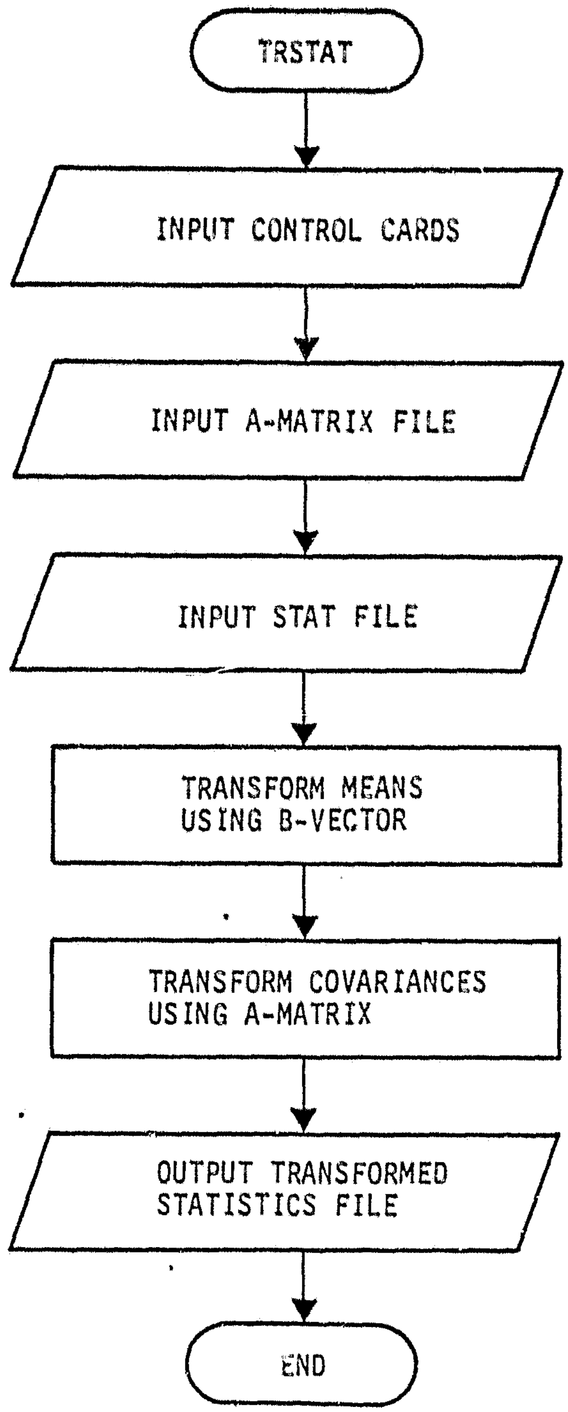


Figure 14-1.— Functional flow chart for the TRSTAT processor.

15. N-DIMENSIONAL HISTOGRAM PROCESSOR — NDHIST

The NDHIST processor computes an n-dimensional histogram of areas of the MSS data file for which the user has requested scatter plots. The user specifies pixel dimensions by the plotting channels. The histogrammed pixels are output on the NHSTUN file, which is written as an interface to the SCTPRL processor.

15.1 PROCEDURES

The number of channels (dimensions) used in histogramming is specified by means of the CHANNELS control card defined in table 15-1. The plotting channels are the primary input channels. The color channels are for further delineation of the frequency in histogramming.

When n is greater than 2, the SCTRPL processor must be directed to reduce the dimensionality to 2 by means of a linear transformation.

The order of the pixels on the output interface file is determined by the plotting channels. The frequency of each such pixel is determined as a function of the color channels (if input) and the plotting channels.

The color assignment for each scatter plot point may be set by the NDHIST or SCTRPL processor. If applicable, the color codes are output on the NHSTUN file. The color codes may be set using the following information.

a. The original radiance values of the pixel (see CHANNELS control card, table 15-1).

b. The mean vector of the cluster or subclass to which the pixel was assigned during clustering or classification. In exercising this option, the user must input a classification or cluster map (see MAPFIL control card, table 15-1) to this processor. To execute the SCTRPL processor, a SAVTAP file related to the MAPUNT must be input (see CHANNELS and STATFILE control cards, section 16, table 16-1). The subclass or cluster numbers assigned to the pixel during classification or clustering are stored on the NHSTUN file, passed to the SCTRPL processor, and used for retrieving the means from the SAVTAP file.

c. The mean vector of the test or training field from which the pixel was extracted (see OPTION MEANS card, table 15-1).

d. User-defined colors (see COLOR control card, table 16-1).

e. From any pass on the MSS data file when using multitemporal Landsat data (see CHANNELS control card, table 15-1).

The areas selected for histogramming are defined by test and/or training fields. The manner in which the fields are collected or grouped for histogramming is user controlled by input parameters. The data vectors may be histogrammed collectively at the class, subclass, or per-field level. The maximum number of fields input at any level is 200, and the maximum number of unique data vectors accumulated at any level is 12 000 divided by one-fourth the number of plotting channels.

A functional flow diagram of the NDHIST processor is given in figure 15-1.

15.2 INPUT FILES

An MSS data file must be input to the NDHIST processor. The assignment defaults to logical unit 11; however, by input of the DATA control card, the user may assign any available logical unit. (See table 4-1 for file assignments and section 3.1, Image Tapes, for further information.) Optionally, a classification or cluster MAPUNT file may be input (see MAPFIL control card).

191

15.3 OUTPUT FILES

An NHSTUN file is always output. It is an interface to the SCTRPL processor and must be assigned to tape or disk. No file-skipping capability is available; the first file created is always file 1. (See HISFIL control card, table 15-1.)

15.4 SCRATCH FILES

The NDHIST processor dynamically assigns random access disk storage for the histogram counters, color codes, identification information, and (optionally) the pixel assignment from the classified or clustered image file (MAPUNT).

15.5 CARD INPUT

15.5.1 PROCESSOR CARD

The processor keyword is left justified starting in column 1, thus,

$NDHIST

This card directs the system monitor routine to select the NDHIST processor and initiates loading of all the NDHIST routines into the system.

15.5.2 SYSTEM CARD FILES

No special system card decks are required for the NDHIST processor.

15.5.3 CONTROL CARDS

Table 15-1 lists the control cards and available options for the NDHIST processor.

15.5.4 FIELD DEFINITIONS

The field cards, which immediately follow the *END control card, define the areas to be histogrammed, and the OPTION control card determines the level of histogramming. The fields may be ordered in one of four ways:

a. As input to STAT (section 8.4.4)

b. As input to ISOCLS (section 9.5.4)

c. As input to CLASSIFY (section 11.5.4)

d. Individually, as input by the user (section 3.2.3)

For example:

```
*END
CLASS               WHT
SUBCLASS            WHT1
(Field card 1)
(Field card 2)
SUBCLASS            WHT2
(Field card 3)
CLASS               NWHT
SUBCLASS            NWH1
(Field card 4)
SUBCLASS            NWH2
(Field card 5)
SUBCLASS            NWH3
(Field card 6)
(Field card 7)
$END
```

If the histogram is accumulated on a class basis, fields 1, 2, and 3 are histogrammed collectively and output as data file 1; and fields 4, 5, 6, and 7 are histogrammed collectively and output as data file 2.

If the histogram is accumulated on a subclass basis, fields 1 and 2 are histogrammed collectively and output as data file 1; field 3 is histogrammed and output as data file 2; field 4 is histogrammed and output as data file 3; field 5 is histogrammed and output as data file 4; and fields 6 and 7 are histogrammed collectively and output as data file 5.

If the histogram is performed on a per-field basis, each field is histogrammed separately and output to a file, making a total of seven data files created.

On a cumulative histogram, a maximum of 200 fields may be input.

See section 3.2.3 for format of the field definition card.

15.6 CARD OUTPUT

The NDHIST processor does not provide punched card output.

15.7 RESTRICTIONS

The system-related restrictions in section 24 apply to this processor. Other restrictions are as follows.

a. A maximum of 16 channels may be histogrammed.

b. A maximum of 4 channels may be used for color codes.

c. The maximum of unique vectors to be histogrammed is

$$n \leq \frac{12\ 000}{1/4(\text{number of channels})} \qquad (15\text{-}1)$$

d. A maximum of 4000 words of storage is allowed for storing the MSS data. The equation for computing the maximum number of pixels is

$$n \leq \frac{4000}{[\text{number of channels}(\text{number of samples per scan line})]} \qquad (15\text{-}2)$$

15.8 DIAGNOSTIC MESSAGES

Diagnostic messages for the NDHIST processor are listed with explanations in appendix I.

TABLE 15-1.- CONTROL CARDS FOR NDHIST

Keyword (a)	Parameter and default values (b)	Function
	Required cards	
CHANNELS	PLOT= $N_1, N_2, \cdots, N_I$, $I \leq 16$ COLOR= $M_1, M_2, \cdots, M_J$ $J \leq 4$	The N's are the channels for determining the position (PLOT) of the pixels to be output on NHSTUN. If I=2, N_1 is the sample location and N_2 the line location on the scatter plot. If I>2, the pixels must be transformed to two components in the SCTRPL processor; component 1 will define the sample location and component 2 the line location. The M's are the channels for the color codes. If the COLOR channels are input, the histogram is a function of both the PLOT and COLOR channels; if the COLOR channels are omitted, the histogram is a function of only the PLOT channels. (See section 16 for further information.)
*END	Blank	Signals the end of the control cards.

[a]The keyword must be left justified in card columns 1 through 10.

[b]The parameter values are in card columns 11 through 72 (beginning in any column past 10).

TABLE 15-1.- Continued.

Keyword	Parameter and default values	Function
$END	Blank	Signals the end of all card input for this processor.
	Optional cards	
DATA	UNIT=n,FILE=n Default: n=11,m=1	n is the number of the logical unit assigned to the MSS data file; m-1 is the number of files to be skipped on the unit.
MAPFIL	UNIT=n,FILE=m Default: None	n is the number of the logical unit assigned to the MAPUNT file; m-1 is the number of files to be skipped on the unit. (The order of the fields to be histogrammed must correspond to the order of the clustered or classified fields on the input MAPUNT file.)
HISFIL	UNIT=N Default: N=4	N is the number of the logical unit assigned to the NHSTUN file.
OPTION	CLASS Default: Field basis	Fields will be histogrammed on the basis of classes.
OPTION	SUBCLS Default: Field basis	Fields will be histogrammed on the basis of subclasses.
OPTION	FIELD	Fields will be histogrammed on a per-field basis.

1971

TABLE 15-1.- Concluded.

Keyword	Parameter and default values	Function
OPTION	MEANS	The means of each field will be computed for the COLOR channels on the CHANNELS card and output on the NHSTUN file.

Ancillary cards

HED1, HED2, DATE, COMMENT (see table 3-1)

198

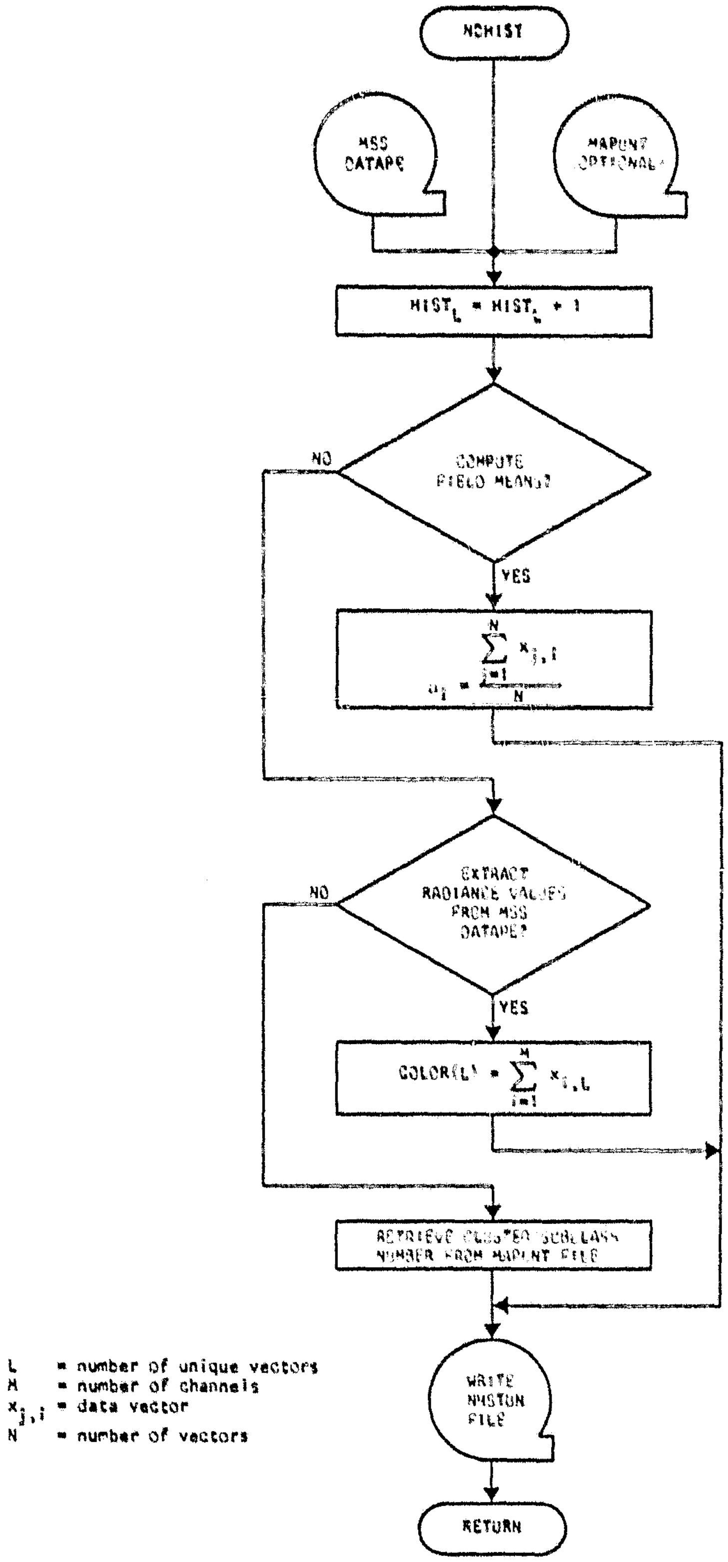


Figure 15-1.– Functional flow chart for the NDHIST processor.

199

16. SCATTER PLOT PROCESSOR — SCTRPL

The SCTRPL processor reads the NHSTUN file written by the NDHIST processor, determines the location of each unique data vector on the scatter plot, and outputs a spectral plot in Universal format. A scatter plot is created and output for each file stored on the NHSTUN.

The location (line and sample intersection) of each pixel on the two-axis scatter plot will be computed using either the radiance values or two linear combinations of radiance values. (This option is controlled in the NDHIST processor by the CHANNELS control card, table 15-1.)

If the data vector is to be transformed (see B-MATRIX control card, table 16-1), the following equation and conditions will be applied:

$$\vec{y} = B\vec{x} + \vec{c} \tag{16-1}$$

where

$\vec{y}$ = a 2-by-1 vector

B = a 2-by-n matrix

$n \leq 16$

$\vec{x}$ = an n-by-1 vector

$\vec{c}$ = a 2-by-1 vector

If the transformed data are to be rescaled (see SCALE control card, table 16-1), the following equation will be applied:

$$Y_i = \left(\frac{HI_i - LO_i}{R_i}\right) \times |MIN_i - Z_i| \tag{16-2}$$

200

where

Y_i = rescaled transformed data value for channel i

HI_i = an input parameter for the upper rescale limit for channel i

LO_i = an input parameter for the lower rescale limit for channel i

R_i = range for channel i [MAX_i - MIN_i]

MAX_i = maximum value for channel i

MIN_i = minimum value for channel i

Z_i = transformed data point for channel i

The scatter plot is created and output line by line. All the pixels belonging to a line, as determined by the second coordinate of the pixel, are collected; and, in the sample location determined by the first coordinate, the color assignment and frequency of occurrence of each pixel are output as channels 1 through n, where n is the channel number of the frequency. [See procedures for NDHIST processor (section 15.1) for definition of color assignments.]

The user controls the dimensions of the output file by using the following input control cards, which are defined in greater detail in table 16-1 (the control cards for the SCTRPL processor) and table 15-1 (the control cards for the NDHIST processor).

Keyword	Parameters
SIZE	XSIZ=129
SIZE	YSIZ=65
SIZE	XHIGH=128
SIZE	XLOW=0
SIZE	YHIGH=64
SIZE	YLOW=0
CHANNELS	PLOT=3,4,COLOR=5,6,7,8

In this case, the output file will contain

a. 129 samples per line with a maximum data resolution of 128.

b. 65 lines per file with a maximum data resolution of 64.

c. 5 channels with channels 1 through 4 containing the color pixel (determined by channels 5, 6, 7, and 8 on CHANNELS control card) and channel 5 the frequency.

The position of each point on the output file is determined by the radiance values of the plotting channels, 3 and 4.

If a MAPUNT file containing the subclass or cluster numbers has been input to the NDHIST processor, either a SAVTAP file related to the MAPUNT file must be input (see STATFILE control card) or the user must input the color codes on cards (see COLOR control card).

Optionally, a line-printer pixel-frequency scatter plot will be output (see PIXPLT control card). The frequency of occurrence or log of frequency of occurrence will be represented by a symbol (see SYMBOL control card). The location of the symbol on the plot will be determined by the radiance value of the pixel. If the data have been transformed, then the data must be rescaled to exercise this option.

A functional flow diagram of the SCTRPL processor is given in figure 16-1.

16.1 INPUT FILES

The NHSTUN file created by NDHIST must be input. (See the HISFIL control card and appendix E for format of the NHSTUN tape.)

The SAVTAP file created by the STAT or ISOCLS processor may be input. (See the STATFILE control card and section 4.1 for a description of the file.)

16.2 OUTPUT FILES

A multifile Universal-formatted tape or disk containing the scatter plots (with color keys) will be output. (See the SCTRUN control card and appendix H for file format.)

16.3 SCRATCH FILES

The program dynamically assigns random access disk storage for scratch files.

16.4 CARD INPUT

16.4.1 PROCESSOR CARD

The processor keyword is left justified starting in column 1, thus,

$SCTRPL

This card directs the system monitor routine to select the SCTRPL processor and initiates loading of all the SCTRPL routines into the system.

16.4.2 SYSTEM CARD FILES

The module STAT and B-matrix card files may be input. See section 3 for formats.

16.4.3 CONTROL CARDS

Table 16-1 lists the options and control cards for the SCTRPL processor.

16.4.4 FIELD DEFINITIONS

Field definitions do not apply to this processor.

16.5 CARD OUTPUT

The SCTRPL processor does not provide punched card output.

16.6 RESTRICTIONS

In addition to the system-related restrictions in section 24, the following restrictions apply to this processor.

a. If the color codes for the scatter plot tape SCTRUN are to be principal component (PC) colors, the user must ensure that the values are positive.

b. The maximum dimension of the B-matrix is 2 by 16; the maximum number of elements in additive vector b is 16.

c. The maximum number of channels on the output tape SCTRUN is 5. Color codes are the first n - 1 channels; the frequency is the n^{th} channel.

d. The maximum number of channels selected from the SAVTAP file is 4.

e. The maximum size of the output tape SCTRUN is 200 samples per scan line and 200 lines.

16.7 DIAGNOSTIC MESSAGES

Error messages for the SCTRPL processor are listed by subroutine in appendix I.

TABLE 16-1.- CONTROL CARDS FOR SCTRPL

Keyword (a)	Parameter and default values (b)	Function
	Required cards	
*END	Blank	Signals the end of the control cards.
$END	Blank	Signals the end of all card input for this processor.
	Optional cards	
HISFIL	UNIT=N Default: N=4	N is the number of the logical unit assigned to the NHSTUN file.
CHANNELS	$N_1, N_2, \cdots, N_I$ $I \leq$ number of channels on SAVTAP ≤ 30 Default: First four channels from NHSTUN file	Statistics for these channels will be extracted from the SAVTAP file; they must be a subset of channels on the SAVTAP file.
STATFILE	UNIT=n,FILE=m Default: None	n is the number of the logical unit assigned to the SAVTAP file; m-1 is the number of files to be skipped on the unit.
MODULE	Blank	Initiates the input of the module STAT deck, which immediately follows this card.

[a]The keyword must be left justified in card columns 1 through 10.

[b]The parameter values are in card columns 11 through 72 (beginning in any column past 10).

TABLE 16-1.- Continued.

Keyword	Parameter and default values	Function
B-MATRIX	CARDS Default: None	The B-matrix is being input by cards.
B-MATRIX	FILE Default: None	The B-matrix is being input by file.
BVEC	T_1,T_2 Default: $T_1=T_2=0.0$	Elements of the additive vector to be used in the transformation; the T's are floating-point numbers.
SCALE	XMAX=T Default: XMAX will be computed from the NHSTUN file.[c]	The upper limit for the transformation of the sample values (x-axis). Used when the B-matrix is applied in SCTRPL. T is a floating-point number.
SCALE	XMIN=T Default: XMIN will be computed from the NHSTUN file.[c]	The lower limit for the transformation of the sample values (x-axis). Used when the B-matrix is applied in SCTRPL. T is a floating-point number.
SCALE	YMAX=T Default: YMAX will be computed from the NHSTUN file.[c]	The upper limit for the transformation of the line values (y-axis). Used when the B-matrix is applied in SCTRPL. T is a floating-point number.

[c]If one of the parameters XMAX, XMIN, YMAX, or YMIN is input, all four parameters must be input.

207

TABLE 16-1.- Continued.

Keyword	Parameter and default values	Function
SCALE	YMIN=T Default: YMIN will be computed from the NHSTUN file.[c]	The lower limit for the transformation of the line values (y-axis). Used when the B-matrix is applied in SCTRPL. T is a floating-point number.
SCALE	FILE	The scale factors will be computed from the NHSTUN file.
SCALE	RESCALE Default: No rescaling of the transformed data	The transformed data will be rescaled to the range of XHIGH, XLOW, YHIGH, and YLOW. (See SIZE control cards.)
SIZE	XSIZ=N Default: XSIZ=101	The number of samples per line to place on the scatter plot output unit; N $\leq$ 200.
SIZE	YSIZ=N Default: YSIZ=101	The number of lines to place on the scatter plot output unit; N $\leq$ 200.
SIZE	XHIGH=N Default: XHIGH=100	The upper limit of the radiance values for the sample axis (x-axis) of the scatter plot; N $\leq$ 255.

[c]If one of the parameters XMAX, XMIN, YMAX, or YMIN is input, all four parameters must be input.

TABLE 16-1.- Continued.

Keyword	Parameter and default values	Function
SIZE	XLOW=N Default: XLOW=0	The lower limit of the radiance values for the sample axis (x-axis) of the scatter plot; $0 \leq N \leq$ XHIGH.
SIZE	YHIGH=N Default: YHIGH=100	The upper limit of the radiance values for the line axis (y-axis) of the scatter plot; $N \leq 255$.
SIZE	YLOW=N Default: YLOW=0	The lower limit of the radiance values for the line axis (y-axis) of the scatter plot; $0 \leq N \leq 255$.
PLOTAP	UNIT=N Default: N=12	N is the number of the logical unit assigned to the spectral plot tape.
BCKGND	N Default: N=255	If N=0, background will be black; if N=255, background will be white.
COLOR	$(m_1),(m_2),\cdots,(m_P)$ or $L*(m_1),K*(m_{L+1}),$ $\cdots,(m_P)$ $P \leq 60$ Default: No user input of color codes	$m_1=n_1,n_2,\cdots,n_\ell$ is the color assignment for cluster 1; $m_2=n_1,n_2,\cdots,n_\ell$ is the color assignment for cluster 2; $m_P=n_1,n_2,\cdots,n_\ell$ is the color assignment for cluster P. $0 \leq n_i \leq 255$ and $\ell \leq 4$. L or K is the number of times a given color assignment is repeated.

209

TABLE 16-1.- Concluded.

Keyword	Parameter and default values	Function
PIXPLT	FREQ Default: No printer plot	Line-printer pixel scatter plot of the frequency of occurrence will be printed.
PIXPLT	LOG	Line-printer pixel scatter plot of the log of frequency of occurrence will be printed.
PIXPLT	RESCALE Default: No rescaling. XSIZ=101, YSIZ=101; the range for x-axis is XLOW+XSIZ-1; the range for y-axis is YLOW+YSIZ-1.	The frequency of occurrence of the pixel for the line-printer scatter plot will be rescaled to ranges XHIGH, XLOW, YHIGH, and YLOW. XSIZ will determine the number of bins on the x-axis; YSIZ, the number of bins on the y-axis. (See SIZE control cards.)
SYMBOLS	$S_1,S_2,\cdots,S_k$ $k \leq 32$ Default: .,/,C,O,Q,O,O,O,B,Θ	Character set separated by commas, with a maximum of 32 characters. The number of symbols/2 determines the number of bin levels. The first set of symbols is overprinted by the second set. A blank is not a legitimate character.

Ancillary cards

HED1, HED2, DATE, COMMENT (see table 3-1)

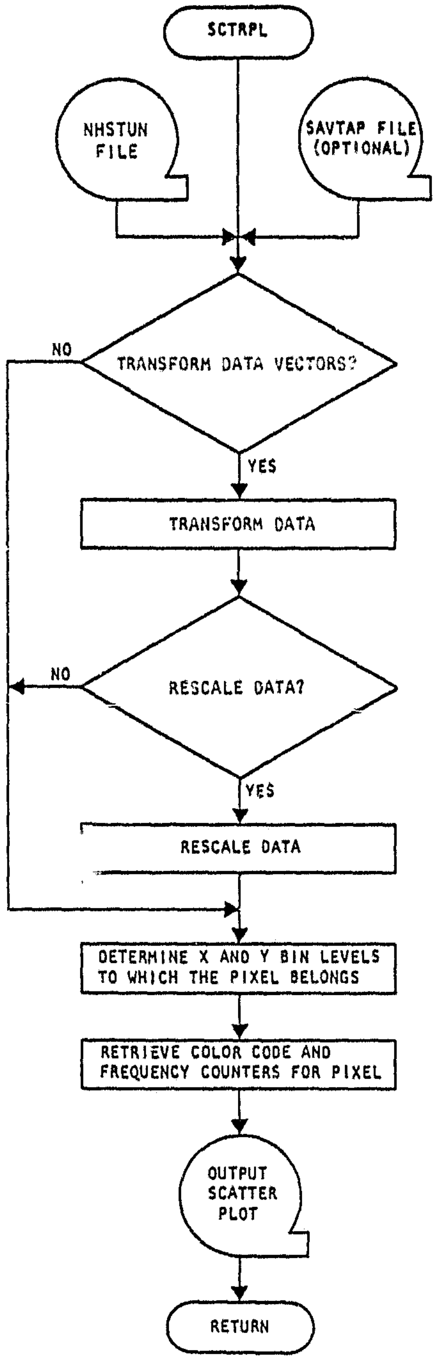


Figure 16-1.— Functional flow chart for the SCTRPL processor.

211

17. DOT DATA PROCESSOR — DOTDATA

17.1 PROCEDURES

In implementing Procedure 1, the DOTDATA processor was added to the system to allow the user to label certain MSS data points corresponding to pixels (known as dots). The main function of this processor is to output a file containing the dots of interest. This file is an interface for three processors, ISOCLS, LABEL, and DISPLAY. (Figure 17-1 shows the functional flow of the DOTDATA processor.)

The dots are defined by field cards. Any subset of the 209 possible grid points may be selected by the user. The dots may be labeled as the file is created or in the LABEL processor. It is not necessary that all the categories of interest be defined in this processor.

This processor actually produces two files. Type 1 dots (starting and labeling dots) are written on one file; type 2 dots (bias correction dots) are written on the other file. As a result, to change the type of a dot, the analyst must execute this processor again.

By an OPTION control card, the user may request that the spatial and spectral information relating to each dot on the file be printed on the line printer.

17.2 INPUT FILES

This processor requires an MSS data file. The assignment defaults to logical unit 11; but by input of the DATA control card, the user may assign any available logical unit. (See section 3.1 for further information on format.)

17.3 OUTPUT FILES

The unformatted DOTUNT files are output as an interface for the processors ISOCLS, LABEL, and DISPLAY. The default for DOTUNT is logical unit 19. (See appendix F for format of the tape.) The logical unit and file number can be controlled by the DOTFIL control card.

17.4 SCRATCH FILES

The DOTDATA processor does not require an additional scratch file.

17.5 CARD INPUT

17.5.1 PROCESSOR CARD

The processor keyword is left justified, beginning in column 1, thus,

$DOTDATA

This card directs the system monitor routine to select the DOTDATA processor and causes all the routines used by it to be loaded into the system.

17.5.2 SYSTEM CARD FILES

The DOTDATA processor does not use any special input files.

17.5.3 CONTROL CARDS

Table 17-1 lists all available options, along with their default values.

17.5.4 FIELD DEFINITIONS

The user defines by field card the grid points to extract from the MSS data file. The order of the field cards determines the order of the dots in the dot data file, DOTUNT. The analyst will need to know the position of the dots to define starting vectors in ISOCLS and to label or relabel the dots in LABEL.

As the fields are defined, the type for each dot is defined by a TYPE card. By option, the analyst may label each dot by a CLASS card. If this card is omitted, the unlabeled dots should be labeled by the control card DOTLABEL or excluded from the set by the control card EXCLUDE in the LABEL processor.

An example of a field data set expected by this processor follows. All names on CLASS cards are read from columns 11 through 15.

```
*END
TYPE      1
CLASS     WHT (optional)
LAB1      (10,10),(10,10),(190,10)     (19 dots)
LAB2      (10,10),(10,20),(190,20)     (19 dots)
CLASS     NWHT (optional)
LAB3      (10,10),(10,50),(100,50)     (10 dots)
TYPE      2
CLASS     WHT (optional)
BIA1      (10,10),(10,40),(190,40)     (19 dots)
CLASS     NWHT (optional)
BIA2      (10,10),(10,70),(190,70)     (19 dots)
$END
```

Two files are written. File 1 contains 38 WHT dots, followed by 10 NWHT dots; all of which are type 1 dots. File 2 contains 19 WHT dots, followed by 19 NWHT dots; all of which are type 2 dots.

If the CLASS cards were omitted, file 1 would contain 48 unlabeled type 1 dots. File 2 would contain 38 unlabeled type 2 dots.

In both cases, the reference numbers for the dots in file 1 defined by the LAB1 field card are 1 through 19, the LAB2 field card reference numbers are 20 through 38, and the LAB3 reference numbers are 38 through 48. The reference numbers for the dots in file 2 defined by the BIA1 field card are 1 through 19; and by the BIA2 field card, 20 through 38.

If the LACIE dot input option is selected, there are no TYPE or CLASS card images. Each input card image has the form

$$\text{DOT}\begin{Bmatrix}1\\2\end{Bmatrix}\begin{Bmatrix}A\\B\\\vdots\\Z\end{Bmatrix}n_1,n_2,\cdots,n_N$$

DOT begins in column 1. The dot type (1 or 2) appears in column 5. A category name appears in columns 8 and 9 (if one character, it should appear in column 8). In columns 11 through 80, integer dot grid numbers specify individual dots. The correspondence is as follows:

Dot grid number	Sample	Line
1	10	10
2	20	10
⋮		
19	190	10
20	10	20
⋮		
39	10	30
⋮		
209	190	110

The dot cards can appear as a separate file in the format described above. This choice is specified by the OPTION U n card, where n is the Fortran unit number of the file. Further information on the LACIE-formatted dot file format can be found in appendix K.

17.6 CARD OUTPUT

DOTDATA does not produce any card decks.

17.7 RESTRICTIONS

System restrictions, presented in section 24, apply to DOTDATA.

17.8 DIAGNOSTIC MESSAGES

Diagnostic messages for the DOTDATA processor are presented with explanations in appendix I.

TABLE 17-1.- CONTROL CARDS FOR DOTDATA

Keyword (a)	Parameter and default values (b)	Function
		Required cards
CHANNELS	DATA=$C_1, C_2, \cdots, C_k$ $k \leq 30$	Integer numbers, separated by commas, referring to the channels on the MSS data file.
*END		Signals the end of the control cards.
$END		Signals the end of all card input for this processor.
		Optional cards
DATA	UNIT=n,FILE=m Default: n=20,m=1	n is the number of the Fortran logical unit assigned to the MSS data file; m-1 is the number of files to be skipped on the unit.
DOTFIL	OUTPUT/UNIT=n,FILE=m Default: n=19,m=1	n is the number of the Fortran logical unit assigned to the DOTUNT file output by this processor; m-1 is the number of files to skip on the unit before writing the DOTUNT file.

[a]The keyword must be left justified in card columns 1 through 10.

[b]The parameter values are in card columns 11 through 72 (beginning in any column past 10).

311

TABLE 17-1.- Concluded.

Keyword	Parameter and default values	Function
OPTION	U n (normally n=29) Default: None	n is the number of the Fortran logical unit assigned to the LACIE-formatted dot file. When this option is taken, the LACIE-formatted dot file is used in lieu of in-line dot cards.
OPTION	LACIE	Dot input will be in LACIE format; card images follow the *END card.
OPTION	PRINT Default: No line-printer output	Prints the DOTUNT file information on the line printer.

Ancillary cards

HED1, HED2, DATE, COMMENT (see table 3-1)

218

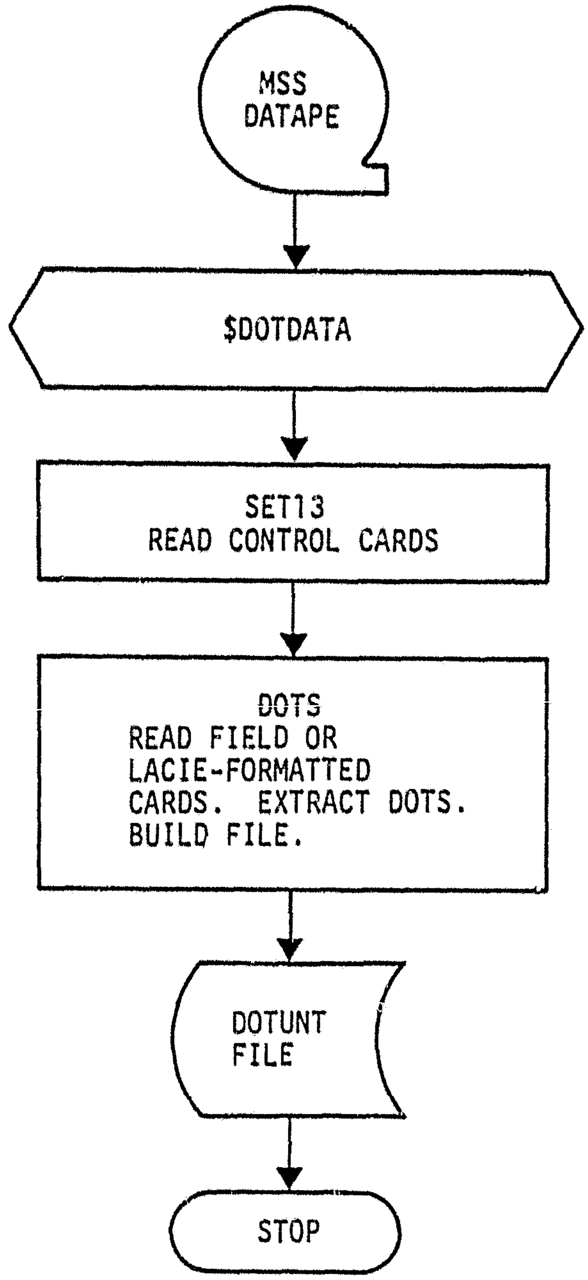


Figure 17-1.— Functional flow chart for the DOTDATA processor.

319

18. CLUSTER LABELING PROCESSOR — LABEL

To aid the analyst in supervising the labeling of the statistics obtained from the clustering processor ISOCLS, a new technique, the processor LABEL, was implemented.

Two procedures for labeling the statistics are provided — the k-nearest-neighbor procedure and the all-of-a-kind procedure (fig. 18-1).

The labels in the dot data file, DOTUNT, or in the previously labeled statistics file, SAVTAP, may be changed by control card input (table 18-1) and the updated file output again. Optionally (1) a conditional or mixed cluster map may be output to tape; (2) unconditional cluster map may be output in the format acceptable to the DISPLAY processor; (3) the spatial and spectral information about the relabeled DOTUNT may be output to the line printer; and (4) the statistics of the SAVTAP file may be output to the line printer.

18.1 PROCEDURES

A distance table containing the L_1 or L_2 distance between each type 1 dot and each cluster mean is computed.

$$L_1 = \sum_{i=1}^{n} |x_i - u_i| \tag{18-1}$$

$$L_2 = \sqrt{\sum_{i=1}^{n} (x_i - u_i)^2} \tag{18-2}$$

where

n = number of channels

x_i = i^{th} element of the dot vector

u_i = i^{th} element of the mean vector

Using the table as input to the k-nearest-neighbor procedure or the all-of-a-kind procedure, the processor labels the statistics generated during clustering.

For the k-nearest-neighbor procedure, the labels of the k labeling dots nearest to a given cluster are polled. The label of the majority of the dots will be the label of the cluster. If a tie occurs, then k - 1 dots are considered.

For the all-of-kind procedure, all of the labeling dots within a cluster are polled. If all the dots are of one category, the cluster is given the label of that category. If the cluster contains labeling dots of more than one category, the label of the majority of the dots labels the cluster. If there are no labeling dots within a cluster, the labeling procedure defaults to k-nearest-neighbor.

Optionally, a conditional cluster map may be output. A cluster is tagged as conditional if the distance between the nearest identically labeled labeling dot and the mean of the cluster is greater than the analyst-input threshold value, t.

Optionally, a mixed cluster map may be output. A cluster is tagged as mixed if the labeling dots within a cluster are of more than one category.

Optionally, a labeled cluster map may be output in the format acceptable to the DISPLAY processor. Information used in the thresholding procedure in DISPLAY is dummied. If thresholding of the clustered data is desired, it can be performed by exercising the conditional map option in this processor.

18.2 INPUT FILES

Either the statistics file (SAVTAP) from ISOCLS or STAT or the dot data file (DOTUNT) from DOTDATA must be input. If either labeling procedure is to be used, both of these files must be input. The cluster map file (MAPUNT) usually from ISOCLS must be input if all-of-a-kind labeling is selected.

For complete descriptions of these files, see section 4.1, appendix F, and section 5.1.

18.3 OUTPUT FILES

Any of the following files may be output.

a. Statistics file (SAVTAP) labeled by using one of the two labeling procedures

b. Statistics file (SAVTAP) relabeled by control card input

c. Relabeled dot data file (DOTUNT)

d. A conditional cluster map

e. A mixed cluster map

f. A labeled cluster map (MAPTAP) in the format acceptable to the DISPLAY processor (see appendix D for this format)

The following items can be output on the line printer:

a. Summary of selected options

b. Table of L_1 or L_2 distances

c. Summary of the labeling dots within a cluster for the all-of-a-kind procedure

d. Summary of the k labeling dots nearest to a cluster for the k-nearest-neighbor procedure

e. Spatial and spectral information about the relabeled DOTUNT file

f. Means and covariances of labeled or relabeled statistics

18.4 SCRATCH FILES

LABEL does not require an additional scratch file.

18.5 CARD INPUT

Formats for all system card input are defined in section 3.2.

18.5.1 PROCESSOR CARD

The processor keyword is left justified starting in column 1, thus,

$LABEL

This card directs the system monitor routine to select the LABEL processor and initiates loading of all the routines needed to execute this processor.

18.5.2 SYSTEM CARD FILES

The LABEL processor does not use any special input decks.

18.5.3 CONTROL CARDS

Table 18-1 lists all available options, along with their default values.

18.5.4 FIELD DEFINITIONS

A field definition card and a MAPUNT file must be input if the all-of-a-kind procedure is selected, a conditional or mixed cluster map is output, or a DISPLAY interface tape is output. This field card defining the area of the unconditional cluster map (input MAPUNT) must be identical to the field card input to ISOCLS and used to create the unconditional cluster map.

223

18.6 CARD OUTPUT

This processor does not produce cards.

18.7 RESTRICTIONS

General system restrictions apply to LABEL.

18.8 DIAGNOSTIC MESSAGES

Diagnostic messages for LABEL are listed by subroutine in appendix I.

TABLE 18-1.- CONTROL CARDS FOR LABEL

Keyword (a)	Parameter and default values (b)	Function
		Required cards
*END		Signals the end of the control cards.
$END		Signals the end of all card input for this processor.
		Optional cards
CHANNELS	STAT=$A_1, A_2, \cdots, A_I$, DATA=$C_1, C_2, \cdots, C_K$ Default: I = all channels on SAVTAP file K = all channels on DOTUNT file	A's and C's are integer numbers, separated by commas, referring to the channels on the SAVTAP file and the DOTUNT file, respectively.
DOTFIL	INPUT/UNIT=n,FILE=m	n is the number of the Fortran unit assigned to the input DOTUNT file; m-1 is the number of files to be skipped on the unit.
EXCLUDE	$n_1, n_2, \cdots, n_i$ $i \leq 250$ Default: No dots are excluded.	$\{n_i\}$ are integer numbers referring to the dots on the DOTUNT file that are to be excluded in all calculations (e.g., dots within a DO/DU area).

[a]The keyword must be left justified in card columns 1 through 10.

[b]The parameter values are in card columns 11 through 72 (beginning in any column past 10).

325

TABLE 18-1.- Continued.

Keyword	Parameter and default values	Function
DOTFIL	OUTPUT/UNIT=n,FILE=m Default: None	n is the number of the Fortran unit assigned to the output DOTUNT file; m-1 is the number of files to skip before writing the relabeled DOTUNT file.
STATFILE	INPUT/UNIT=n,FILE=m Default: None	n is the number of the Fortran unit assigned to the input SAVTAP file; m-1 is the number of files to be skipped on the unit.
MODULE	Blank	Initiates the input of the module STAT file. The file must immediately follow this card.
STATFILE	OUTPUT/UNIT=n,FILE=m Default: None	n is the number of the Fortran unit assigned to the output SAVTAP file; m-1 is the number of files to skip before writing the newly labeled SAVTAP file.
MAPFIL	INPUT/UNIT=n,FILE=m Default: n=16,m=1	n is the number of the Fortran unit assigned to the input MAPUNT file. (If executing back to back with ISOCLS, n must be 16.) m-1 is the number of files to be skipped on the unit.

TABLE 18-1.- Continued.

Keyword	Parameter and default values	Function
MAPFIL	OUTPUT/UNIT=n,FILE=m Default: n=16,m=1	n is the number of the Fortran unit assigned to the output MAPUNT file; m-1 is the number of files to skip before writing MAPUNT. (If both types of maps are output, the conditional map is output on file m; the mixed map, on file m+1.)
MAPTAP	OUTPUT/UNIT=n,FILE=m Default: No DISPLAY interface file will be output.	n is the number of the Fortran unit assigned to the output MAPTAP file. (If executing back to back with DISPLAY, n must be 2.) m-1 is the number of files to skip before writing MAPTAP.
DISTANCE	L1	The L_1 distance between the labeling dots and the cluster means is used.
DISTANCE	L2 Default: L_1 distance	The L_2 distance between the labeling dots and the cluster means is used.
SUNANG	$m_1, m_2, \cdots, m_i$ Default: No Sun angle correction is applied.	$\{m_i\}$ are integer Sun angle numbers used in computing the L_1 or L_2 distance. A Sun angle must be input for each acquisition of interest. An acquisition is assumed to be a four-channel pass.

227

TABLE 18-1.- Continued.

Keyword	Parameter and default values	Function
		Example: If the distance is computed using 16 channels, four Sun angles (m_1, m_2, m_3, and m_4) must be input.
SUNANG	FILE Default: No Sun angle correction is applied.	Sun angles will be extracted from the DOTUNT file.
PROCED	NAME Default: N=K-NEAREST	NAME is an alphabetic word. NAME = K-NEAREST (Use the k-nearest-neighbor procedure.) NAME = ALL (Use the all-of-a-kind procedure.) NAME = MANUAL (Use the manual procedure of relabeling the DOTUNT or SAVTAP file.)
NEAREST	K Default: K=1	K is the number of dots to be used in the k-nearest-neighbor procedure. K is an integer number $\leq$ 11.
DOTLABEL	Category name, $n_1, n_2, \cdots, n_j$ $j \leq 250$ Default: None	The DOTUNT file is labeled by this card. Category name is the label the analyst is assigning to the dots $\{n_j\}$.

TABLE 18-1.- Continued.

Keyword	Parameter and default values	Function
		The category name may be composed of a maximum of four characters. $\{n_j\}$ are integer numbers, separated by commas, referring to the position of the dot on the DOTUNT file.
OPTION	EXIT Default: The run continues.	This card allows the user to exit the EOD-LARSYS run if any label input to the processor is not used to label at least one cluster.
STALABEL	Class name, $n_1,n_2,\cdots,n_j$ $j \leq 250$ Default: None	The SAVTAP file may be manually relabeled by this card. $\{n_j\}$ are the numbers of the subclasses on the SAVTAP that are to be regrouped into another class. Class name is the name of the class to which subclasses $\{n_j\}$ are to be reassigned. The class name must match a name on the SAVTAP file.
OPTION	DOTS Default: Relabeled DOTUNT is not printed.	Spatial and spectral information about the relabeled DOTUNT file is printed.
OPTION	STATS Default: Statistics not printed	Means and covariances for labeled or relabeled statistics on the SAVTAP file are printed.

TABLE 18-1.- Concluded.

Keyword	Parameter and default values	Function
OPTION	COND Default: None	A conditional cluster map will be output.
THRESHOLD	T Default: T=25.0	T is the threshold parameter used in creating the conditional cluster map. T is a floating-point number.
OPTION	MIXED Default: None	A mixed cluster map will be output.
	Ancillary cards	
	HED1, HED2, DATE, COMMENT (see table 3-1)	

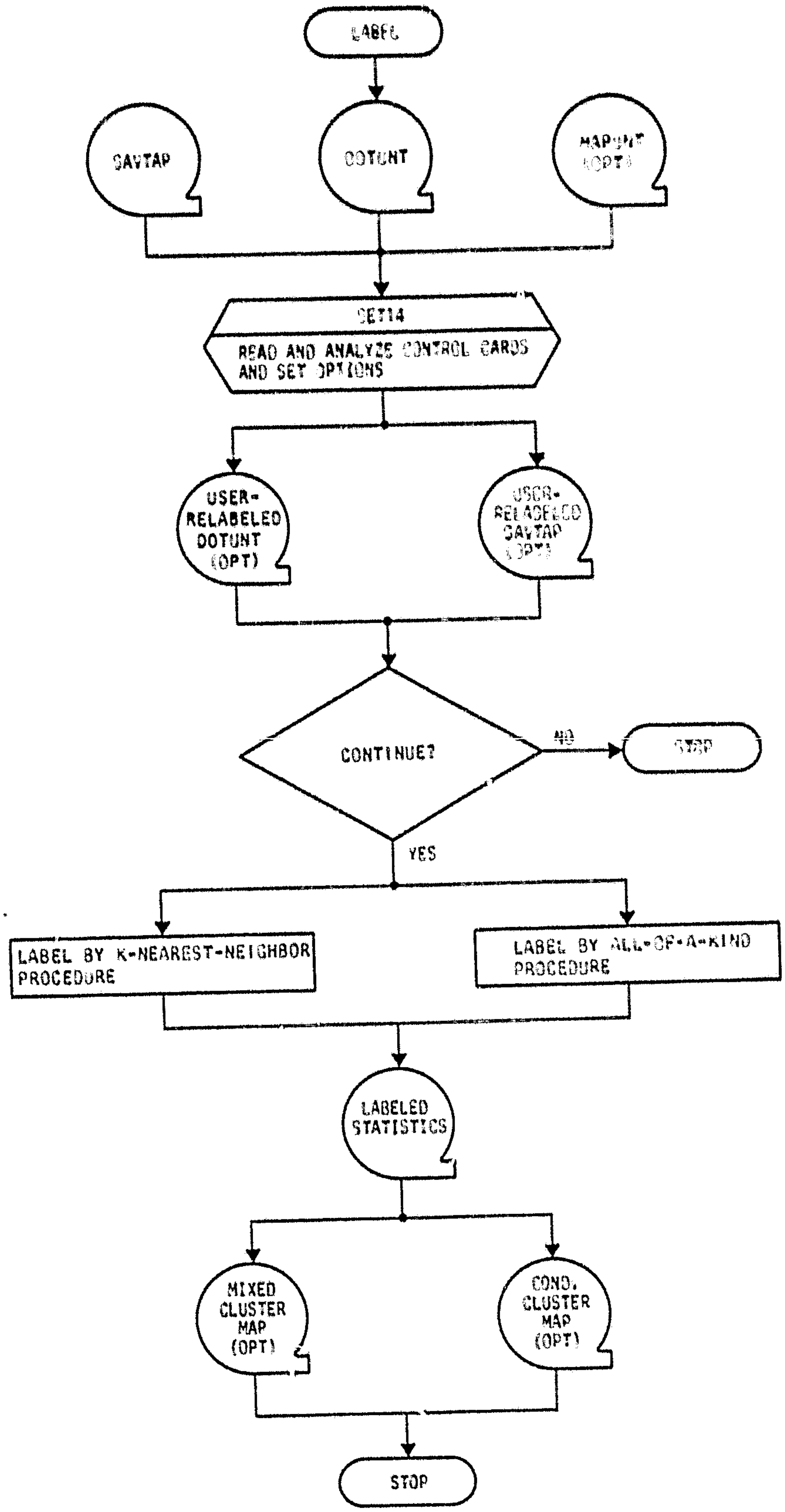


Figure 18-1.– Functional flow chart for the LABEL processor.

231

19. INCORPORATION OF RECENTLY DEVELOPED AND FUTURE PROCESSORS

The processors described in the following sections have been developed recently. The need for multitemporal MSS data and ground truth in applications of pattern recognition such as Procedure 1 prompted their development.

Because of the wide applicability of these processors, they can be regarded as utilities for the rest of the EOD-LARSYS.

Provisions have been made in the monitor routine to incorporate future processors as they become available. Candidates for incorporation include:

AMOEBA (Texas A&M clustering/classification program)

CLASSY (NASA/JSC/Lockheed clustering/classification algorithm)

EQUPRB (Equi-Probable Blocks classifier and distribution function estimator)

MULBAY (Multitemporal Bayes classifier)

PCG (Principal Component Greenness transformation program)

232

20. DATA MERGE PROCESSOR — DAMRG

20.1 DESCRIPTION

DAMRG is a versatile processor which performs the following three types of merge operations:

- Channel merge — Specified channels from selected MSS files are concatenated to provide a field image with more channels.
- Spatial merge — Specified fields are added to the side or bottom of a given field to make a larger image.
- Line merge — Specified scan lines from up to six images are stacked to make an artificial image.

The channel merge is useful for preparing multiple-acquisition images, which can then be used for temporal analysis. It can also be used to aggregate data of several types, provided only that all are registered to the same base, such as a map or reference image.

The spatial merge is useful for combining images of adjacent areas to make larger mosaic images. The processor requires that each input field have the same number of channels, but the analyst may specify the channels from each field.

Line merging is useful for preparing certain composite images for special purposes. Lines must all have the same length; that is to say, they must all contain the same number of samples. The number of channels from each MSS file must be the same, but selection of the channels is up to the user. Certain very restricted types of line merging could also be performed with the spatial merge option, but line merging will handle the general case.

20.2 INPUT/OUTPUT

This processor can take data from up to six files and merge them into a single file. However, only four tape drives are available; hence, to merge six files, some of the files must be on the same tape. Since the output data are written to disk, there is a practical limit to the number and size of files that can be merged.

DAMRG is different from other processors in that it allows no defaults for DATA input cards. These cards, as well as SUNANG, LINES, and CHANNELS cards, if needed, and field definition cards must be present and in correct order. The number and order of files to be merged are deduced from these cards, and the order is preserved as other control cards are read. However, omission of the DATA output card will cause default to unit 11, file 1, so that the merged image can be used normally in subsequent processors. Control cards used by the DAMRG processor are listed in table 20-1. Diagnostic messages are presented in appendix I.

Figure 20-1 is a functional flow chart for the DAMRG processor.

TABLE 20-1.- CONTROL CARDS FOR DAMRG

Keyword (a)	Parameter and default values (b)	Function
	Required cards	
DATA	INPUT/UNIT=n_1,FILE=m_1	Fortran unit numbers and file numbers, one card for each file, in order.
CHANNELS	$n_1,\cdots,n_{m_1}$	Channel numbers, one card for each file, in order.
*END	Blank	Signals the end of the control cards.
$END	Blank	Signals the end of all card input for this processor.
	Optional cards	
DATA	OUTPUT/UNIT=n,FILE=m Default: n=11,m=1	Unit and file numbers for output.
FORMAT	UNIVERSAL or LARSYS II/III Default: UNIVERSAL	Format of output data file.
OPTION	CHANNEL or SPATIAL or PSEUDO Default: CHANNEL	Merging option.
NCPASS	N Default: N=4	Number of channels per acquisition.

[a]The keyword must be left justified in card columns 1 through 10.
[b]The parameter values are in card columns 11 through 72 (beginning in any column past 10).

TABLE 20-1.- Concluded.

Keyword	Parameter and default values	Function
NACROS	N Default: N=1	In SPATIAL option, number of fields to be abutted horizontally (number across).
NLIN	$n_1, \cdots, n_{NUMFIL}$	In PSEUDO option, number of lines from each file.
LINES	$n_1, \cdots, n_J$	In PSEUDO option, scan line numbers for extraction from files.[c]
OPTION	ANGCOR Default: No sun angle correction	Sun angle correction applied to output pixels.
SUNANG	$n_1, \cdots, n_{m_1}$	Sun angles extracted from input cards.
SUNANG	TAPE	Sun angles extracted from tape headers.

Ancillary cards

HED1, HED2, DATE, COMMENT (see table 3-1)

[c]If the scan line numbers will not all fit on one card, continue on subsequent cards, each card having LINE in columns 1 through 4.

236

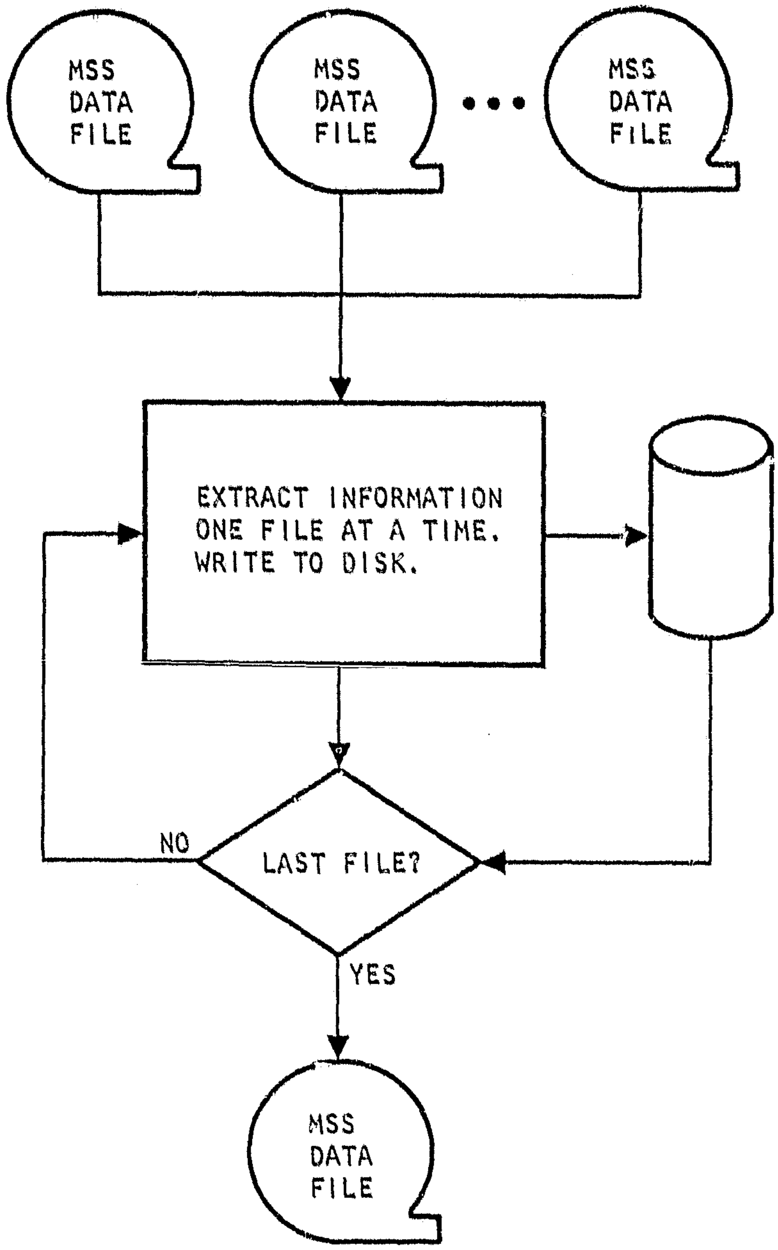


Figure 20-1.— Functional flow chart for the DAMRG processor.

21. GROUND TRUTH DOT LABELING PROCESSOR — GTDDM

21.1 DESCRIPTION

The GTDDM processor labels the 209 LACIE dots on the basis of the converted ground truth file produced by the GTTCN processor (see section 22). It is a very specialized processor, tied to the specific dimensions of LACIE segments and the formats of LACIE dot files.

GTDDM outputs a LACIE-formatted dot file (see appendix K) produced from the converted ground truth file. The user may indicate the source of the dot types (Phase III, Transition Year, or input). He or she may also provide a crop-code-to-category-name transformation.

21.2 INPUT/OUTPUT

Input files are in Universal image format; pixels are represented by code values instead of radiance values.

Transition Year crop codes are used as defaults, as shown in table 21-1 (default for TRANS card). Other code labels can be used in connection with the TRANS card.

The masking operation divides dots into type 1 (labeling dots) and type 2 (bias correction dots). These types are integral to LACIE's Procedure 1. The user may apply either the LACIE Phase III or the LACIE Transition procedure for assigning types (1 or 2) to dots or furnish a different set.

This processor is controlled by the control cards shown in table 21-1. Diagnostic messages are given in appendix I.

Figure 21-1 is a functional flow chart for the GTDDM processor.

TABLE 21-1.- CONTROL CARDS FOR GTDDM

Keyword (a)	Parameter and default values (b)	Function
	Required cards	
*END	Blank	Signals the end of the control cards.
$END	Blank	Signals the end of all card input for this processor.
	Optional cards	
READ	UNIT=n,FILE=m Default: n=12,m=1	n is the number of the Fortran logical unit to which the converted ground truth file has been assigned; m-1 is the number of files to be skipped on the unit.

[a]The keyword must be left justified in card columns 1 through 10.

[b]The parameter values are in card columns 11 through 72 (beginning in any column past 10).

231

TABLE 21-1.- Continued.

Keyword	Parameter and default values	Function
TRANS	Category=n_1, n_2 Default: N=1,256 W=99,99 W=124,124 S=100,100 S=125,125 B=101,101 B=126,126 R=102,102 R=127,127 F=103,103 F=128,128 O=104,104 O=129,129 W=1,15 S=16,30	Crop-code-to-category-name transformation. The category is denoted by a letter of the alphabet. Crop codes n_1 and n_2 are assigned to this category. Subsequent assignments override previous ones.
MASK	PHASE THREE	Use the Phase III dot types.
MASK	TRANSITION YEAR	Use the Transition Year dot types.
MASK	INPUT=$n_1, n_2, \cdots, n_{19}$ (17 cards) Default: TRANSITION YEAR	Input a type matrix (size 17 by 19) where n_i = 1 or 2.
DUMP	FILES=N Default: N=1	Number of dot files to be written.

TABLE 21-1.- Concluded.

Keyword	Parameter and default values	Function
WRITE	UNIT=n,FILE=m Default: n=23,m=1	n is the number of the Fortran logical unit to which the dot files will be written; m-1 is the number of files to skip before writing the dot file.

Ancillary cards

HED1, HED2, DATE, COMMENT (see table 3-1)

241

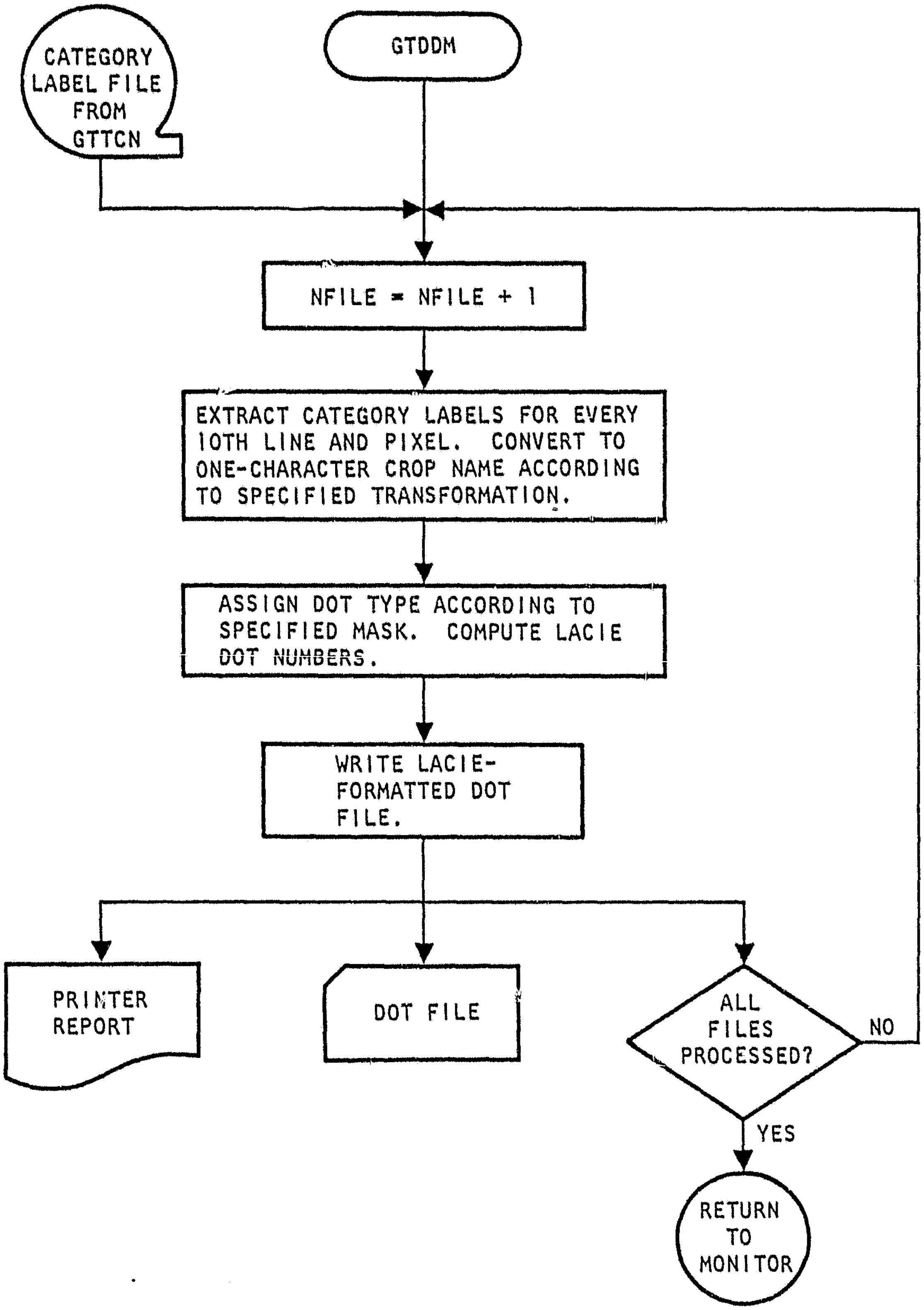


Figure 21-1.— Functional flow chart for GTDDM.

22. GROUND TRUTH DATA CONVERSION PROCESSOR — GTTCN

22.1 DESCRIPTION

GTTCN is a specialized processor constructed to convert artificial LACIE Accuracy Assessment images (392 samples by 351 lines) to standard LACIE segment images (196 samples by 117 lines). After conversion, these one-channel image files contain a ground truth code number for each pixel. The assigned numerical codes are those used in LACIE to identify crops or types of ground cover.

The input image contains six labeled pixels (two samples on each of three lines) for every one LACIE segment pixel. The user chooses a subset of the six to be used in the conversion, and the processor determines the label by majority rule. For example, in the following case,

Ground truth pixel (subpixel)		Corresponding label	
M1	M2	30	20
M3	M4	40	30
M5	M6	40	40

if all six pixels are used (the default case), the majority label would be 40. If the user specified only pixels 1, 3, and 4 by submitting the following card

LABEL 1,0,1,1,0,0

then the label chosen would be 30 (the majority of 30, 40, and 30).

Originally developed for use in LACIE, this processor is also useful for any application requiring ground truth. Such ground-observed data are generally accessible at LARS.

22.2 INPUT/OUTPUT

Both the incoming ground truth file and the converted file are written in Universal format.

The cards shown in table 22-1 control this processor in the standard way. Diagnostics are shown in appendix I. A functional flow chart is shown in figure 22-1. The processor will normally be used in conjunction with the ground truth dot labeling processor (GTDDM). The GTDDM processor is described in section 21.

Codes are not documented formally, but the LACIE Transition Year codes are given in table 21-1.

TABLE 22-1.- CONTROL CARDS FOR GTTCN

Keywords (a)	Parameter and default values (b)	Function
	Required cards	
*END	Blank	Signals the end of the control cards.
$END	Blank	Signals the end of all card input for this processor.
	Optional cards	
CONVERT	FILES=N Default: N=1	Number of files to be converted.
READ	UNIT=n,FILE=m Default: n=11,m=1	n is the number of the Fortran logical unit to which an input tape has been assigned; m-1 is the number of files to be skipped on the unit.
LABEL	VECTOR= N_1,N_2,N_3,N_4,N_5,N_6 Default: N_i=1 for each i	If N_i=1, that subpixel is used in labeling a pixel. If N_i=0, that subpixel is not used.
OPTION	PRINT Default: No print	This card requests a listing of the crop codes for the 209 dots.

[a]The keyword must be left justified in card columns 1 through 10.

[b]The parameter values are in card columns 11 through 72 (beginning in any column past 10).

TABLE 22-1.- Concluded.

Keyword	Parameter and default values	Function
WRITE	UNIT=n,FILE=m Default: n=12,m=1	n is the number of the Fortran logical unit to which the converted ground truth file will be written; m-1 is the number of files to skip before writing the converted file.

Ancillary cards

HED1, HED2, DATE, COMMENT (see table 3-1)

246

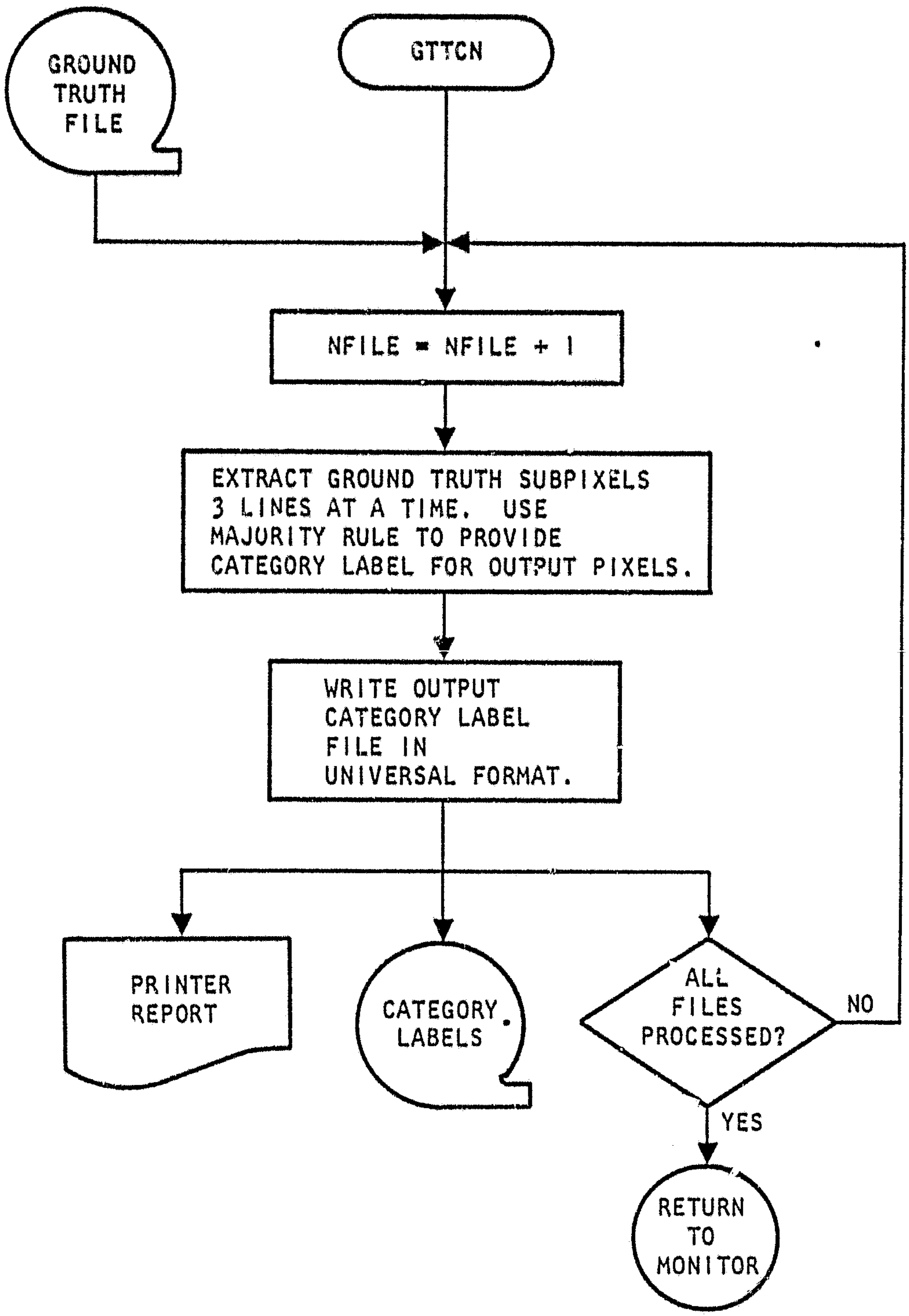


Figure 22-1.— Functional flow chart for GTTCN.

23. EFFICIENT CLUSTERING PROCESSOR — TESTSP

The TESTSP processor is identical to the ISOCLS processor except that in this case the data are stored in packed form on the direct access file, thereby reducing disk storage requirements for this file by a factor of 4. The ISOCLS processor is described in section 9.

24. SYSTEM RESTRICTIONS

EOD-LARSYS is limited in every processor to processing no more than 30 channels of data. The MSS data file (DATAPE) may have more then 30 channels, but for processing purposes a subset of those channels must be selected via the CHANNELS control card.

A maximum of 60 categories, classes, or subclasses may be processed. However, it may not be possible to process the maximum number of channels and subclasses in the same run. The arrays within the system are dimensioned variably according to user requests. The amount of storage available will not accommodate the arrays that are dimensioned (number of subclasses) by (number of channels) if both maximums are used. Restrictions under the STAT, SELECT, and CLASSIFY processors allow the user to compute approximately whether or not the numbers of channels and subclasses selected are acceptable. When core storage requirements are exceeded, a diagnostic message is printed and the user must reduce his or her requirements to get a successful execution. Restrictions specific to individual processors are noted in the description of each processor.

As a result of the virtual storage characteristics of the system, arrays can be enlarged at user request by the system maintenance group, Exploratory Investigations Section, Lockheed.

APPENDIX A

REFERENCES

APPENDIX A

REFERENCES

1. Baker, K.: Algorithm Documentation for CLASSIFY. NASA/JSC informal memorandum, June 1974.

2. Minter, R. T.: Computer Program Documentation, ISOCLS, Iterative Self-Organizing Clustering Program. Program C094, Lockheed Electronics Co. (Houston), Oct. 1972.

3. Baker, K.: User Guide to University of Houston Linear Feature Selection Program. NASA/JSC informal memorandum, Oct. 1973.

4. Quirein, J. A.: Requirements Document to Incorporate Linear Feature Selection Techniques Into the Earth Resources Interactive Processing System. LEC-1403, Lockheed Electronics Co. (Houston), Nov. 1973.

5. Thomas, V. L.: Generation and Physical Characteristics of Landsat 1 and 2 MSS Computer-Compatible Tapes. NASA/Goddard Space Flight Center document X-563-75-223 (Greenbelt, Md.), Nov. 1975.

6. Hollsenbrink, Patrick F.: Manual on Characteristics of Landsat Computer-Compatible Tapes Produced by the EROS Data Center Digital Image Processing System. GPO stock no. 024-001-03116-7, U.S. Geological Survey, 1978.

7. Fletcher, R.; and Powell, M. J. D.: A Rapidly Converging Descent Method for Minimization. Computer J., vol. 6, no. 2, July 1963, pp. 163-168.

8. Decell, H. P.; and Quirein, J. A.: An Iterative Approach to the Feature Selection Problem. Proc. Conf. on Machine Processing of Remotely Sensed Data, IEEE (A74-17543 05-08), 1973, pp. 3B-1 through 3B-12.

9. Hallum, C. R.; and Minter, T. C.: Maximum Likelihood Classification by Thresholding. NASA/JSC Technical Report 640-TR-114, Lockheed Electronics Co. (Houston), June 1972.

10. Feiveson, A. H.: Classification by Thresholding. JSC-08447, Sept. 1973.

11. Eppler, W. G.: Empirical Distribution of Quadratic Form Used for Thresholding. FSD-0001, Lockheed Electronics Co. (Houston), Nov. 1972.

12. Van Rooy, D. L.; and Lynn, M. S.: Use of Spatial Information in Classification of Remotely Sensed Data. Rice Univ. (Houston), May 1973.

APPENDIX B

LARSYS III FORMAT FOR AN MSS DATA TAPE

APPENDIX B
LARSYS III FORMAT FOR AN MSS DATA TAPE

This is the third version of the MSS data storage format used in Purdue's LARSYS. The only difference between the second and third versions of the format is one word in the header record. That difference is transparent to EOD-LARSYS.

There are four types of physical records on the MSS data tapes. They are

1. ID record — 200 4-byte words
2. Data record — long enough to hold one scan line of data
3. End-of-tape record — 200 4-byte words
4. End-of-file record — IBM Standard

An MSS data tape contains one or more data acquisitions consisting of an ID record, several data records, and an end-of-file record. After the last data record on the tape, an end-of-tape record and two end-of-file records are written on the tape.

As used in this document, a "word" is defined to be 32 bits and a "byte" to be 8 bits. Further details regarding the physical records follow.

B.1 ID RECORD (200 4-BYTE WORDS)

Word	Format	Description
ID(1)	I	LARS tape number (e.g., 1, 17, 102)
ID(2)	I	File number on this tape

Word	Format	Description
ID(3)	I	Run number (8 digits aabbbbcc) aa — last 2 digits of the year data were acquired bbbb — running serial number for the year data were taken cc — uniqueness digits for runs that would otherwise have the same run number
ID(4)	I	Continuation code ID(4) = 0 means the first line of data follows this ID record ID(4) = X means that the data following this ID record are a continuation of an acquisition started on tape X
ID(5)	I	Number of data channels (spectral bands) on tape (30 maximum)
ID(6)	I	Number of data samples per channel per scan line
ID(7-10)	A(4A4)	Mission identification (16 characters)
ID(11)	I	Month data were taken
ID(12)	I	Day data were taken
ID(13)	I	Year data were taken
ID(14)	A(1A4)	Time data were taken
ID(15)	I	Altitude of aircraft
ID(16)	I	Ground heading of aircraft
ID(17-19)	A(3A4)	Date data run was generated on this tape (12 characters)
ID(20-50)	I	All zeros (may be changed later)

255

Word	Format	Description
ID(51)	R	Lower limit in micrometers of first spectral band on tape
ID(52)	R	Upper limit in micrometers of first spectral band on tape
ID(53)	R	The suggested value of "C0" calibration pulse
ID(54)	R	The suggested value of "C1" calibration pulse
ID(55)	R	The suggested value of "C2" calibration pulse
ID(56-200)	R	Repeat of ID(51-55) for ID(5) channels in order of appearance in data records
ID(51-200)	R	0.0 if data channels do not exist

B.2 DATA RECORD

Each data record will contain one scan line of data from ID(5) (see ID record) channels. The first half word (2 bytes) will be the record number. The second half word (2 bytes) will be the roll parameter, which is a number indicating relative roll of the aircraft for this scan line of data. If the roll parameter is -32 767, the data for the given line does not exist. If the roll parameter has not been calculated, it will be set to 32 767. The fifth byte will be the first data sample from the first channel. The data samples are ordered $channel_1$, $sample_1$ — $sample_N$; $channel_2$, $sample_2$ — $sample_N$; and so on through ID(5) channels and ID(6) data samples per channel. A data record (scan line) will be ID(5) × ID(6) + 4 bytes long.

256

All but the last 6 bytes in each channel will be scanner data. The last 6 are calibration data, in order of appearance:

1. C_0 "0" or dark level
2. VC_0 Variance of C_0
3. C_1 Calibration source C_1
4. VC_1 Variance of C_1
5. C_2 Calibration source C_2
6. VC_2 Variance of C_2

where

C_i = calibration value i

VC_i = calculated variance of calibration value i

On good data records, all 8-bit data and calibration values will be integers in the range of 0 to 255 with no sign included in the 8 bits. A sample data value of 0 to 255 is the result of the 8-bit analog-to-digital conversion which produces the MSS data tape.

B.3 END-OF-TAPE RECORD (200 4-BYTE WORDS)

Word	Format	Description
ID(1)	I	LARS tape number
ID(2)	I	File number on this tape
ID(3)	I	Set equal to zero
ID(4)	I	Continuation code ID(4) = 0 means end of data ID(4) = X means data in previous file is continued on tape X
ID(5-50)	I	All zeros (may be changed later)
ID(51-200)	R	0.0 (may be changed later)

APPENDIX C

UNIVERSAL FORMAT FOR AN MSS DATA TAPE

APPENDIX C

UNIVERSAL FORMAT FOR AN MSS DATA TAPE

This is an adaptation of the Universal data tape format as defined in the Earth Resources Data Format Control Book.*

C.1 GROUND RULES

The ground rules for the Universal format as accepted by all the processors within EOD-LARSYS are as follows:

a. 8 bits = 1 byte.

b. The header record is the first record on a tape.

c. The header record is 3060 bytes long.

d. Data following the header will be arranged by data sets, where a data set is defined as the ancillary data and all of the MSS data for one scan line for all active channels.

e. Data sets will be recorded in variable length physical records, not to exceed 3000 bytes of information per record. Note, since 3000 bytes is not compatible with the word lengths of all computers, the computer generating the tape will add a sufficient number of fill zeros to the end of the data to make the record length divisible by 32, 36, 48, and 60 bits (180 bytes). Therefore, it is possible to have a physical record length of 3060 bytes, but under no condition will the actual data exceed 3000 bytes.

f. Data sets will be packed into consecutive physical records of equal length. Under no condition will a data set begin in the middle of a physical record unless the data set can be completed in that record. If two or more records are needed

*Vol. 1, Rev. A. NASA/JSC Technical Report PHO-TR543, Philco-Ford Corp. (Houston), Mar. 1975, sec. 7.

for the data set, the data set will be divided, but under no condition will the data for an MSS channel begin in the middle of a physical record unless the data for that MSS channel can be completed in that record. Consequently, lengthy data sets will be divided so that the ancillary block and MSS data from an integral number of channels will be in one record and the remaining MSS data will follow in succeeding records with an integral number of channels per record. Fill zeros will be supplied at the ends of the records as required to satisfy the equal length constraint noted in e.

g. All data in the header record and ancillary blocks will be in binary.

h. The tape format will be as follows:

```
Header record

     IRG*

Ancillary block }  Data set
MSS block       }

     IRG†

Ancillary block }  Data set
MSS block       }

     IRG
      .
      .
      .
     EOF
```

---

*IRG = inter-record gap. This always follows the header record.

†An IRG may appear between the ancillary block and the MSS block so that the recording of a data set requires more than one physical record; or a physical record may contain two or more data sets, not separated by any IRG. See ground rules above and data set description following for criteria determining the placement of IRG's.

266

C.2 HEADER RECORD

Although the header record is 3060 bytes long, only a portion of the information is pertinent to the system at this time. A general description of the data that are unpacked by the TAPHDR routine is as follows:

Byte	Description
89	Processing flag: 0 = raw data 1 = processed data from computing system
90	Number of channels in this job
91	Number of bits per radiance value (currently 8)
92-93	Address (within scan) of start of MSS data
96-97	Number of MSS elements per scan within a single channel
100-101	Physical record size in bytes (must be a multiple of 180 bytes)
102	Number of channels per physical record (This field refers to the second and subsequent records within the recording of a data set. Bytes 1785 and 1786 give the number of channels of data in the first record of a data set. If the number of elements per channel is greater than 3000, this field will equal 0.)
103	Number of physical records per scan per channel (This field is used only when the number of elements per channel is greater than 3000. Otherwise it is equal to 0.)
104	Number of records to make a complete data set
105-106	Length of ancillary block in bytes

261

Byte	Description
107	Data order indicator: 0 = MSS data ordered by channel 1 = MSS data ordered by pixel
108-109	Start pixel number (Number of the first pixel per scan on this tape referenced to original image. The first pixel in the original image is pixel 1.)
110-111	Stop pixel number (Number of the last pixel per scan on this tape referenced to original image.)
1778	Number of data sets per physical record
1785-1786	Number of channels in the first physical record of the data set
1787-1788	Total number of bytes per scan per channel
2201-2202	Sun angle for pass 1
2203-2204	Sun angle for pass 2
2205-2206	Sun angle for pass 3
2207-2208	Sun angle for pass 4
2254-2261	Soil line for pass 1
2262-2269	Soil line for pass 2
2270-2277	Soil line for pass 3
2278-2285	Soil line for pass 4

C.3 DATA SETS

C.3.1 ANCILLARY BLOCK

The first block of a data set is the ancillary block. The length of the ancillary block is variable. The number of bytes is given in the header record.

262

The first word (2 bytes) of every record is a counter giving the number of this physical record within the MSS data set. This is primarily intended for use in data sets that are longer than 3000 bytes and therefore require more than one physical record. This word will always be "1" for the first record of a data set.

Bytes 3 through 6 will contain the current GMT at the start of this data set recorded in tenths of a millisecond.

Bytes 7 through 70 will indicate channel status for this scan, 1 byte per channel, where LSB = 0 indicates the channel is synchronized, and LSB = 1 indicates the channel is not in sync.

Bytes 71 and 72 contain the scan line number. This will be an arbitrary but sequential count for each scan line that appears in the data run.

Bytes 73 through N will be dependent on whether this job contains raw or processed data. (See byte 89 in the header record.) The value of N will be given in bytes 105 and 106 in the header record and will always be equal to or greater than 70. If this job contains raw data, bytes 73 through N will contain the house-keeping data channel from the sensor, if one is available. A job containing processed data will, in addition to the 70 bytes of ancillary data already described, contain, at a minimum, the following pieces of information:

a. Latitude of the aircraft or of the center of the image from EREP* or satellite, in binary

b. Longitude of the aircraft or of the center of the image from EREP or satellite, in binary

*Earth Resources Experiment Package.

c. Altitude in meters, recorded in binary

d. Heading, in tenths of a degree

e. Ground speed, in meters per second

f. Roll — format specific to sensor

g. Pitch — format specific to sensor

h. Yaw — format specific to sensor

i. Sun angle

Other parameters may be added, if required. The total length of the ancillary block is given in the header.

C.3.2 MSS BLOCK

Following the ancillary block in each data set will be an MSS block consisting of the MSS data from all of the active channels for one scan. MSS blocks within a data run will always contain the same number of MSS channels. Each MSS block will be the same number of bytes in length. If MSS data are not available to fill a block, fill zeros will be added to make it the same length as preceding MSS blocks.

MSS data having fewer than 8 bits per pixel will be right justified in an 8-bit byte, with zeros added to the left. MSS data having more than 8 bits per pixel will be right justified in as many 8-bit bytes as necessary to hold the pixel, with zeros added to the left.

If this tape contains raw data, the pulse-code modulated (PCM) sync words associated with the MSS data, if any, will be included with the MSS data on this tape. If this tape contains processed data, no sync words will be present.

If this tape contains raw imagery data, calibration data for each channel used in each scan will be included, in the same sequence in which the raw data appear in the data stream of the mission tape. If this tape contains processed imagery data, the appearance of the calibration data will depend on the specific sensor requirements.

The combined length of the ancillary block and the MSS block will determine the relationship between data sets and physical records. Some data runs may contain data sets which are so small that more than one can be packed into one physical record. Others may contain data sets which will require a whole physical record for each. Still others may contain data sets which are so long that each data set will require two or more physical records.

The way in which a data set is packed into a physical record depends on the length of the data set. The ancillary block will always appear in the first physical record per data set. Following the ancillary block, as many complete channels in this data set will be recorded as will fit in 3000 bytes. If the data set is too long to be recorded in one physical record, the second and subsequent records will begin with the next active channel in the data set.

If an MSS block is divided between records, the number of data channels in the first record may differ from the number of channels in the second and subsequent records; however, the number of channels in all records after the first one in a data set will be the same. The number of channels in the first record and the number in the subsequent records will be given in the header record. In records after the first, if insufficient MSS data are available to allow all records to contain the same number of channels, fill zeros will be added to make all of these records the same length. Finally, fill zeros will be added to either the first record or all of the subsequent records, depending on which is shorter, so as to make all of the records the same length.

Within the context of EOD-LARSYS, pixels within a scan of data will be arranged by channel. The Universal format* will be as follows:

- Line 1
  - Channel 1
    - Pixel 1
    - Pixel 2
    - ⋮
    - Pixel N
  - Channel 2
    - Pixel 1
    - Pixel 2
    - ⋮
    - Pixel N
  - ⋮
  - Channel M
    - Pixel 1
    - Pixel 2
    - ⋮
    - Pixel N
- Line 2
  - Channel 1
    - Pixel 1
    - ⋮
    - Pixel N
  - ⋮
  - Channel M
    - Pixel 1
    - ⋮
    - Pixel N
- ⋮

*If this tape contains raw imagery data, the PCM sync words, if any, associated with the data on the mission tape will be included with the data.

APPENDIX D

MAPTAP FILE FORMAT

APPENDIX D

MAPTAP FILE FORMAT

The file MAPTAP is output by the processor CLASSIFY. It contains the statistics used in classification; the training field, category, class, and subclass information; and the classified data.

Each file consists of the following types of records:

	4 run header records
Repeated for each classified field {	1 field header record
	N data records
	1 end-of-field record
	1 end-of-run record
	1 end-of-file record

All records are written using unformatted Fortran WRITE statements.

D.1 RUN HEADER RECORD 1

```
WRITE(MAPTAP)(DATE(I),I=1,2),BMFLG,BMCOMB,BMFEAT,NOCLS2,
             NOFLD2,NOSUB2,NOFET2,TOTVT2,NOCAT,VARSZ2,
             (FETVC2(I),I=1,NOFET2)
```

Fortran name and dimension	Description
DATE(2)	Date the classification was performed
BMFLG	Flag indicating B-matrix was used in classification
BMCOMB	Number of linear combinations in B-matrix
BMFEAT	Number of channels used in computing the B-matrix
NOCLS2	Number of classes

Fortran name and dimension	Description
NOFLD2	Number of training fields
NOSUB2	Number of subclasses
NOFET2	Number of channels used in classification
TOTVT2	Number of vertices in training fields
NOCAT	Number of categories
VARSZ2	Size of covariance for each subclass
FETVC2(NOFET2)	Actual channels used in classification

D.2 RUN HEADER RECORD 2

```
WRITE(MAPTAP)(CATNAM(I),I=1,NOCAT1),(CLSMTX(I),I=1,NOCLS2),
             (SUBNO(I),I=1,NOCLS2),(SUBDES(I),I=1,NOSUB2),
             ((FLDMTX(I,J),I=1,4),J=1,NOFLD2),
             ((VERTEX(I,J),I=1,2),J=1,TOTVT2),
             (SUBCAT(I),I=1,NOSUB2),(CLSVC2(I),I=1,NOSUB2),
             (KATNO(I),I=1,NOCLS2),(KEPPTS(I),I=1,NOSUB2)
```

Fortran name and dimension	Description
CATNAM(NOCAT1)	Category names (if available) NOCAT1 = number of categories if CATEGORY classifier was applied NOCAT1 = number of classes if STANDARD classifier was applied
CLSMTX(NOCLS2)	Class names
SUBNO(NOCLS2)	Number of subclasses in each class
SUBDES(NOSUB2)	Subclass names

Fortran name and dimension	Description
FLDMTX(4,NOFLD2)	Training field information: 1 — field name 2 — number of class to which field belongs 3 — number of subclass to which field belongs 4 — number of vertices in this field
VERTEX(2,TOTVT2)	Vertices for all the fields; ordered $(sample,line)_1,(sample,line)_2,\cdots,(sample,line)_{TOTV2}$
SUBCAT(NOSUB2)	Contains the number of the category to which each subclass belongs
CLSVC2(NOSUB2)	Contains the number of the class to which each subclass belongs
KATNO(NOCLS2)	Contains the number of the category to which each class belongs
KEPPTS(NOSUB2)	Contains the total number of training field pixels in each subclass

D.3 RUN HEADER RECORD 3

```
WRITE(MAPTAP)((COVMTX(I,J),I=1,VARSZ2),J=1,NOSUB2)
             ((AVEMTX(I,J),I=1,NOFET2),J=1,NOSUB2)
```

Fortran name and dimension	Description
COVMTX(VARSZ2,NOSUB2)	Original or B-matrix-transformed covariance matrix for all subclasses
AVEMTX(NOFET2,NOSUB2)	Mean vector for each subclass

D.4 RUN HEADER RECORD 4

```
WRITE(MAPTAP)((COVMTX(I,J),I=1,VARSZ2),J=1,NOSUB2),
             (CON(I),I=1,NOSUB2),(DET(I),I=1,NOSUB2)
```

Fortran name and dimension	Description
COVMTX(VARSZ2,NOSUB2)	Modified Cholesky factorization of the covariance matrix for all subclasses
CON(NOSUB2)	Natural logorithm of determinant divided by a priori value squared
DET(NOSUB2)	Determinant of covariance matrix for each subclass

D.5 FIELD HEADER RECORD

```
WRITE(MAPTAP)(FLDINF(I),I=1,6),PTS,LINES,FLDESC,NC,
             (VERTCS(I),I=1,NC),(VERTCS(I+NC),NC=1,NC)
```

Fortran name and dimension	Description
FLDINF(6)	Rectangular coordinates surrounding the field classified: 1 — line start 2 — line stop 3 — line increment 4 — sample start 5 — sample stop 6 — sample increment
PTS	Number of points in the rectangular field defined in FLDINF
LINES	Number of lines in the rectangular field defined in FLDINF
FLDESC	Name of the classified field
NC	Number of vertices in the classified field
VERTCS(2,NC)	Vertices for the classified field; ordered $(\text{sample,line})_1,(\text{sample,line})_2,\cdots,(\text{sample,line})_{NC}$

D.6 DATA RECORD

```
WRITE(MAPTAP)ILINE,(IR(I),I=1,PTS),(VR(I),I=1,PTS)
```

Fortran name and dimension	Description
ILINE	Line number in reference to the MSS data tape
IR(PTS)	Subclass number to which each classified data point belongs
VR(PTS)	Likelihood that the point belongs to that subclass

D.7 END-OF-FIELD RECORD

An end-of-field record has the same format as a data record with ILINE = 0.

D.8 END-OF-RUN RECORD

An end-of-run record has the same format as the field header record with PTS = 0.

APPENDIX E

NHSTUN FILE FORMAT

APPENDIX E

NHSTUN FILE FORMAT

The interface file written to the NHSTUN is output by the NDHIST processor and read by the SCTRPL processor.

All records are written using unformatted Fortran WRITE statements. The header record is always the first file on NHSTUN.

The format of NHSTUN is as follows:

File 1	Header record
	EOF
Data file 1	Record 1
	Record 2
	Record 3 (optional)
	Record 4
	Record 5
	Record 6
	Record 7 (optional)
	EOF
⋮	⋮
Data file N	EOF

The contents of each record are as follows.

E.1 HEADER RECORD

Fortran name and dimension	Description
TOTMNS	Total number of means computed
SIZE	NOFET2 divided by 4

Fortran name and dimension	Description
NOFET2	Number of channels to histogram
(FETVC2(I),I=1,NOFET2)	Actual channels to histogram
NCLRCH	Number of color code channels
(CLRVEC(I),I=1,NCLRCH)	Actual color code channels

E.2 RECORD 1

Fortran name and dimension	Description
NOFLD2	Number of fields histogrammed
NOSUB2	Number of subclasses histogrammed
TOTVT2	Number of vertices
NOVEC	Number of unique vectors histogrammed

E.2 RECORD 2

Fortran name and dimension	Description
CLSVC2	Class name
(SUBVC2(I),I=1,NOSUB2)	Subclass names
((FIELDS(I,J),I=1,4),J=1,NOFLD2)	Field information
((VERTEX(I,J),I=1,2),J=1,TOTVT2)	Field vertices

E.4 RECORD 3 (OPTIONAL)

Fortran name and dimension	Description
(MEANS(I),I=1,TOTMNS)	Mean statistics for input fields

PRECEDING PAGE BLANK NOT FILMED

E.5 RECORD 4

Fortran name and dimension	Description
((PLOT(I,J),I=1,SIZE),J=1,NOVEC)	Data vectors

E.6 RECORD 5

Fortran name and dimension	Description
(ID(I),I=1,NOVEC)	Class, subclass, and field to which the data vectors belong

E.7 RECORD 6

Fortran name and dimension	Description
(COUNTR(I),I=1,NOVEC)	Number of occurrences of the data vectors

E.8 RECORD 7 (OPTIONAL)

Fortran name and dimension	Description
(COLOR(I),I=1,NOVEC)	Color extracted from MSS data file

PRECEDING PAGE BLANK NOT FILMED

APPENDIX F

DOTUNT FILE FORMAT

C-41

APPENDIX F

DOTUNT FILE FORMAT

The file written on the DOTUNT is output by the DOTDATA processor. The records are written using unformatted Fortran WRITE statements.

A file is output for each type of field. The file consists of the following records:

Repeat for each type	Record 1 / Record 2	Field information
	Record 3 / EOF	Data record

In the context of Procedure 1, type 1 dots are used for cluster seeding and labeling; type 2 dots are used for bias correction of proportion estimates.

F.1 RECORD 1

```
WRITE(DOTUNT)NOCAT,NOFEAT,NOFLD,TOTVRT,TOTDOT,NOSUN,
             (CATNAM(I),I=1,NOCAT),SIZE
```

Parameter	Dimension (words)	Definition
NOCAT	1	Number of category names
NOFEAT	1	Number of channels
NOFLD	1	Number of fields
TOTVRT	1	Number of vertices
TOTDOT	1	Number of dots
NOSUN	1	Number of Sun angles

Parameter	Dimension (words)	Definition
CATNAM	NOCAT	Array containing the category names
SIZE	1	4 + NOFEAT

F.2 RECORD 2

```
WRITE(DOTUNT)(FETVEC(I),I=1,NOFEAT),
             ((FLDSAV(I,J),I=1,4),J=1,NOFLD),
             ((VERTEX(I,J),I=1,2),J=1,TOTVRT),
             (ANGLE(I),I=1,NOSUN)
```

Parameter	Dimension (words)	Definition
FETVEC	NOFEAT	Array containing the channel numbers
FLDSAV	(4,NOFLD)	Array containing the field description
VERTEX	(2,NOFLD)	Array containing the field vertices
ANGLE	NOSUN	Array containing the Sun angles

F.3 RECORD 3

```
WRITE(DOTUNT)((DOTS(I,J),I=1,SIZE),J=1,TOTDOT)
```

Parameter	Dimension (words)	Definition
DOTS	(TOTDOT,SIZE)	Array containing the dot information DOTS(1,I) = sample number for dot i DOTS(2,I) = line number for dot i DOTS(3,I) = type number for dot i

Parameter	Dimension (words)	Definition
		DOTS(4,I) = category number for dot i (optional) DOTS(5,I) ⋮ = dot vector i DOTS(4+NOFEAT,I)

APPENDIX G

DESCRIPTION OF CLUSTER IMAGE DISPLAY
WITH COLOR KEYS

APPENDIX G

DESCRIPTION OF CLUSTER IMAGE DISPLAY WITH COLOR KEYS

The cluster image data tape output by the ISOCLS processor contains the mean vector to which each corresponding pixel was assigned during clustering and a color key. The color key consists of n square images, each 10 samples by 10 lines in dimension. Each color code square represents the mean vector for a given cluster. The color codes are ordered according to cluster number or greenness. The greenness ordering (G) is a function of the four Landsat channels:

$$\left.\begin{aligned} G_{i,N} &= -0.29\mu_{1,N} - 0.56\mu_{2,N} + 0.60\mu_{3,N} + 0.49\mu_{4,N} \\ G_N &= \sum_{i=1}^{M} G_{i,N} \end{aligned}\right\} \qquad \text{(G-1)}$$

where

i = number of pass

N = cluster number

μ_1 = first channel of pass i

μ_2 = second channel of pass i

μ_3 = third channel of pass i

μ_4 = fourth channel of pass i

M = number of passes for multitemporal Landsat data

The number of color codes per scan line is computed by

$$K = \frac{\text{number of samples per scan line}}{11} \qquad \text{(G-2)}$$

304

The number of lines required to display the color codes is computed by

$$L = \left[\frac{(\text{number of clusters} - 1)}{K} + 1 \right] \times 11 \qquad \text{(G-3)}$$

The cluster field and color key are separated by a scan line of zeros. Each color code square is separated from the next by a vertical line of zeros.

The data are output in LARSYS II/III or Universal format (see appendixes B and C, respectively).

The structure of the file is as follows.

<u>HEADER RECORD</u>

Inter-record gap (IRG)

N records — Mean vector for each corresponding pixel

IRG

Record (N + 1) — Scan line of zeros

IRG

```
                                        0           0           0
                                        0           0           0
Record (N + 2) — 10 lines  { color      0   color   0   . . .   0   color
                           { code 1     :   code 2  :           :   code K
                                        0           0           0
                             \________/   \________/              \________/
                             10 samples
```

IRG

⋮

Record (N + L) — Color code (last cluster)

EOF

APPENDIX H

SCTRUN FILE FORMAT

APPENDIX H

SCTRUN FILE FORMAT

The scatter plot image, written to the SCTRUN, contains two-axis color-coded spectral plot(s) and is output in the Universal format (appendix C) by the SCTRPL processor. Each file on SCTRUN contains (1) a single scatter plot image, of which N - 1 channels are color assignments and the N^{th} channel is the frequency channel, and (2) a color key, unless the color assignment is made up of the radiance values of each output pixel.

The color key consists of n square images dimensioned 10 samples by 10 lines. A color code square is composed of the colors assigned to a given cluster. The color codes are ordered according to their cluster association; i.e., the color code associated with cluster 1 is output first, followed by the color code associated with cluster 2, etc. The number of color codes per scan line is computed by

$$K = \frac{\text{number of samples per scan line}}{11} \qquad \text{(H-1)}$$

The number of lines required to display the color codes is computed by

$$L = \left[\frac{(\text{number of clusters} - 1)}{K} + 1\right] \times 11 \qquad \text{(H-2)}$$

The scatter plot image and color key are separated by a scan line of zeros. Each color code square is separated from the next by a vertical line of zeros.

The dimensions of the output tape are user controlled via the control cards SIZE and CHANNELS (see section 16, table 16-1, and section 15, table 15-1, respectively), where the number of samples per scan line = XSIZ; the number of channels = dimensions of color pixel plus the frequency channel; and the number of scan lines = YSIZ + L.

The structure of the file is as follows.

HEADER RECORD

IRG

YSIZ records — Scatter plot image

IRG

or

EOF (if color key is omitted)

Record (YSIZ + 1) — Scan line of zeros

IRG

```
                                         0            0          0
                                         0            0          0
Record (YSIZ + 2) — 10 lines { color     0   color    0   ...    0   color
                             { code 1    :   code 2   :          :   code K
                             {           0            0          0
                              \_________/  \_______/              \_______/
                               10 samples
```

IRG

⋮

Record (YSIZ + L) — Color code (last cluster)

EOF

APPENDIX I

DIAGNOSTIC MESSAGES

APPENDIX I

DIAGNOSTIC MESSAGES

I.1 HIST - SUBROUTINE SETUPS

	Message	Explanation
a.	CHANNEL ___ IS NOT ONE OF THE CHANNELS GIVEN ON CHANNELS CARD.	A channel on the DISPLAY card is not a member of the set of channels on the CHANNELS card.
b.	TOO MANY CHANNELS ARE BEING HISTOGRAMMED AND PLOTTED -- NO. OF CHANNELS WAS RESET TO ________.	User requested too many histograms to be plotted. The number of histograms plotted varies according to the number of channels histogrammed.
c.	XHIGH - XLOW WAS LESS THAN 100. XHIGH WAS RESET TO XXX, OR XLOW WAS RESET TO XXX.	Range of pixel radiances required to be ≥ 100.
d.	ERROR ON DATA CARD.	Check unit assignment and file number.
e.	BAD SUPERVISOR CONTROL CARD.	Check spelling of keywords.
f.	INVALID CARD -- IGNORED.	Inappropriate or defective card read. Make sure cards are punched correctly.

I.2 GRAYMAP

I.2.1 SUBROUTINE PICT

	Message	Explanation
a.	THE NO. OF CHANNELS FOR THIS FIELD HAS BEEN REDUCED TO XXX SO ALL THE INFORMATION WILL FIT ON DISK. MAKE ANOTHER RUN TO GRAYMAP OTHER CHANNELS.	Self-explanatory.
b.	FIELD TOO LARGE, TERMINATING.	Data exceed allocated storage.
c.	YOU HAVE ASKED FOR TOO MANY SAMPLES. THE LAST SAMPLE IS ____________.	The last sample is reset to the last sample on the data tape.

I.2.2 SUBROUTINE SETUP6

	Message	Explanation
a.	THIS CHANNEL IS OUT OF NUMERICAL RANGE AND WAS IGNORED.	All channels requested must be in the range 1 to 30.
b.	THIS CHANNEL IS NOT HISTOGRAMMED.	Check CHANNELS control card and make sure all channels requested have been histogrammed.
c.	ONLY 16 BINLEVELS PERMITTED.	Reduce the number of bin levels to 16.
d.	ERROR ON DATA CARD.	Check for format error and unit assignment.
e.	BAD SUPERVISOR CONTROL CARD.	Check spelling of keywords.

I.3 STAT

I.3.1 SUBROUTINE SETUP1

	Message	Explanation
a.	///// FROM SUBR. SETUP1 -- BAD CONTROL CARD ENCOUNTERED -- INPUT CARD IS ________, 'CCCC ··· CCC'	The input card read has none of the legitimate keywords to identify it as a recognizable control card. The faulty card is printed out as part of the message. Although the processor will continue to read more control cards, this is an indication of an error in the deck setup. The deck should be checked for proper control cards and proper sequence of cards.
b.	*** STAT/SETUP1 -- ERROR IN OPTION(S) REQUESTED -- SCAN OF OPTION(S) DISCONTINUED AT CARD COLUMN XX ***	An OPTION control card is not acceptable to the processor. The scan of the options will be discontinued by the processor, and any options specified beyond the erroneous one will not be activated for the run. The processor continues by reading the next control card. (See section 3.2.2 and table 8-1 for correct OPTION control card usage.)
c.	ERROR ON DATA CARD	Check format and unit assignment.
d.	ERROR ON STATFILE CARD	Check format and unit assignment.

Message	Explanation
e. *** MAXSUB=XX -- MAX. NO. OF SUBCLASSES CAN-NOT BE GREATER THAN YY MAXSUB SET=YY PROCEEDING TO NEXT OPTIONS(S) ***	The maximum subclass number input on the OPTION MAXSUB control card exceeds the maximum number of subclasses that can be handled by EOD-LARSYS. The processor will set the maximum number of subclasses, which will apply to subclasses read in from the input subclass field definition deck.
f. /////FROM SUBR. SETUP1 -- DECREASE OPTIONS ***** TERMINATING PROGRAM EXECUTION FROM SUBR. SETUP1 *****	The STAT processor has run out of internal storage to handle the combination of the quantities of input training fields, subclasses, and channels. Internal storage is fixed at 10 600 locations. Each subclass required roughly 1/2(number of channels)^2 locations for the subclass statistics. If histograms or spectral plots of subclasses and/or fields are requested, additional internal storage is required. The options specified in the run deck (i.e., histograms and spectral plots) and possibly the quantities of subclasses, channels, and training fields must be decreased or eliminated in order to get a successful run within the core storage limitation.

243

Message	Explanation
g. CHECK CHANNELS OR CLASS NOS. REQUESTED -- CANNOT BE LESS THAN OR EQUAL 0, OR GREATER THAN 30 ***** TERMINATING PROGRAM EXECUTION FROM SUBR. SETUP1 *****	If the channel numbers specified on a HISTO or CHANNELS control card are not integers within the range 1 through 30, this message results. The processor halts after printing this message. Check the format of the applicable processor control cards (see section 3.2.2 and table 8-1).

I.3.2 SUBROUTINE LEARN

Message	Explanation
a. ***** STAT/LEARN -- MAX. OF XX SUBCLASSES EXCEEDED -- FIRST XX SUBCLASSES USED -- REMAINDER IGNORED	The processor has read the maximum allowable number of subclass names and associated training fields. The first MAXSUB subclasses and associated training fields input are computed and the remainder are ignored by the processor.
b. ***** STAT/LEARN -- MAX. OF XX FIELDS EXCEEDED -- XX FIELDS RETAINED FOR YY SUBCLASSES -- REMAINDER OF INPUT TRAINING FIELDS NOT USED	The STAT processor has read the maximum allowable number of training fields. The available internal storage has been filled, and no further training fields can be accepted. Training statistics will be computed for the subclasses and fields which have been read to this point, and the remainder are ignored by the processor.

I.4 ISOCLS AND TESTSP

I.4.1 SUBROUTINES ISOCLS AND TESTSP

Message	Explanation
a. NO. CHANNELS FOR STARTING NOT EQUAL THAT FOR CLUSTER.	The number of channels of starting vectors from the STAT file must equal the number of requested data channels.
b. DIMENSION LIMITS EXCEEDED IN ISOCLS BY ______. REDUCE CHANNELS OR MAX. CLUSTERS.	The user has exceeded storage. The number of channels or maximum clusters per class should be reduced.
c. DIMENSION LIMIT OF ______ FOR COVARIANCES EXCEEDED.	Same.

I.4.2 SUBROUTINES PSPLIT AND PSPPAT

Message	Explanation
ERROR READING DISK -- ISTAT=XXXX.	Operating system returns non-standard status from disk input output.

I.4.3 SUBROUTINES RDDATA AND RDDPAT

Message	Explanation
a. TOO MANY DO OR DU FIELDS. THESE IGNORED.	There can be up to 10 DO fields and 10 DU fields.
b. TOO MUCH DATA REQUESTED -- PIXELS * (CHANNELS + 1) CANNOT EXCEED ______.	Disk file will not hold all of the data for one class. Reduce channels or size of fields.

Message	Explanation
c. STORAGE REQUIRED FOR FIELD DEFINITION INFORMATION EXCEEDS THE DIMENSION LIMIT OF _____.	All vertices, names, and rectangular coordinates are saved for each field. The user has exceeded storage. Reduce the number of fields.
d. END OF TAPE REACHED BEFORE END OF FIELD.	A field has been defined beyond the limits of the MSS DATAPE.
e. INPUT ERROR -- A CLASS CARD MUST BE INPUT BEFORE A GROUP OF FIELDS.	See section 9.5.4 on defining classes and fields.
f. NO. OF PIXELS TO BE UNPACKED PER SCAN EXCEEDS THE DIMENSION LIMIT OF _____.	Decrease the number of channels or pixels per scan line in the field.

I.5 SELECT

I.5.1 SUBROUTINE AVEDIV

Message	Explanation
a. REDUCED COVARIANCE MATRIX FOR CLASS XXX IS NOT POSITIVE DEFINITE.	Check subclass/cluster statistics for singularity.
b. MORE STORAGE NEEDED IN SUBR. AVEDIV FOR WORK ARRAY -- IWRKSZ=XXXXXXX.	Storage inadequate; adjust parameters.

I.5.2 SUBROUTINE BHTCHR

Message	Explanation
a. COVAR. FOR CLASS XXX IS NOT POSITIVE DEFINITE.	Check subclass/cluster statistics for singularity.

Message	Explanation
b. COVAR. FOR SUM OF CLASSES XXXX,XXXX IS NOT POSITIVE DEF.	Same as above.
c. NOT ENOUGH WORK AREA AVAILABLE IN BHTCHR -- IWRKSZ=XXXXX.	Storage inadequate; adjust parameters.

I.5.3 SUBROUTINE BSTCHK

Message	Explanation
a. "BEST" XXX IS GREATER THAN OR EQUAL TO NO. OF FEATURES IN GIVEN DATA -- IGNORED.	The channels included in "best" must be a subset of the input channels.
b. INVALID EVALUATE REQUEST.	The channels whose separabilities are to be evaluated must be a subset of total input channels.

I.5.4 SUBROUTINE DAVDN1

Message	Explanation
ERROR ON DISK FILE -- SUBR. DAVDN1 -- ISTAT=XXX.	Tape hardware read error.

I.5.5 SUBROUTINE DAVDN2

Message	Explanation
MINIMUM IS AT ORIGIN -- PROGRAM CANNOT CONTINUE.	Davidon-Fletcher-Powell iteration has reached an unusable minimum.

I.5.6 SUBROUTINE DAVDN3

Message	Explanation
DAVDN3 -- EITHER SIGYI OR YHY TOO CLOSE TO ZERO TO UPDATE H MATRIX -- SIGYI=XXXXXXX.XXXXXXX. YHY=XXXXXXX.XXXXXXX.	Davidon-Fletcher-Powell iteration cannot continue.

I.5.7 SUBROUTINE DAVIDN

Message	Explanation
a. NOT ENOUGH WORK AREA AVAILABLE IN DAVIDN -- IWRKSZ=XXXXXX.	Storage inadequate; adjust parameters.
b. ERROR ON DISK FILE -- SUBR. DAVIDN == LSTAT=XXX.	Read error on disk file.

I.5.8 SUBROUTINE DIVERG

Message	Explanation
a. COVAR. FOR CLASS XXXX IS NOT POSITIVE DEFINITE.	Check subclass/cluster statistics for singularity.
b. NOT ENOUGH WORK AREA AVAILABLE IN DIVERG -- IWRKSZ=XXXXX.	Storage inadequate; adjust parameters.

I.5.9 SUBROUTINE GTSTAT

Message	Explanation
NOT ENOUGH WORK AREA IN GTSTAT -- IWRKSZ=XXXXX.	Storage inadequate; adjust parameters.

I.5.10 SUBROUTINE SELECT

	Message	Explanation
a.	ERROR IN INPUT CHANNELS.	User should review choice of input channels with respect to input parameters related to channels.
b.	CORE OVERFLOW IN SUBRAY -- NN STORAGE LOCATIONS NEEDED FOR THIS PROBLEM.	User might reduce the number of subclasses or channels or try another procedure. The SUBRAY array is used for temporary storage in SELECT only. (See restrictions, section 10.6.)
c.	CORE OVERFLOW IN ARRAY -- NN*2 STORAGE LOCATIONS NEEDED FOR THIS PROBLEM.	See suggestions for previous diagnostic message. The ARRAY is used throughout the system for variably dimensioned storage.

I.5.11 SUBROUTINE SETUP4

	Message	Explanation
a.	ERROR ON STATFILE CARD.	Check format.
b.	TOO MANY EVALUATE REQUESTS -- REMAINDER IGNORED.	The buffer to hold EVALUATE requests is dimensioned 100. The number of channels and channels to be evaluated for each EVALUATE request are stored in this array.
c.	GROUP CARD IN ERROR -- IGNORED.	Check format of GROUP option.
d.	PROGRAM CANNOT PROCESS LESS THAN 2 CHANNELS.	At least two channels must be input.

319

Message	Explanation
e. PROGRAM CANNOT PROCESS LESS THAN 2 CLASSES.	At least two classes must be input.
f. INVALID CONTROL CARD -- IGNORED.	Check spelling of keyword.
g. CORE NEEDED IN ARRAY FOR THIS PROBLEM IS XXXXXX WORDS.	Storage is inadequate; adjust parameters.

I.5.12 SUBROUTINE TRNDIV

Message	Explanation
a. REDUCED COVARIANCE MATRIX FOR CLASS N IS NOT POSITIVE DEFINITE.	The indicated covariance matrix cannot be inverted.
b. NOT ENOUGH WORK AREA IN TRNDIV -- IWRKSZ=XXXXX.	Storage is inadequate; adjust parameters.

I.5.13 SUBROUTINE WGTCHK

Message	Explanation
SUBCLASS IS NOT AMONG INPUT SUBCLASSES -- WEIGHT INPUT IGNORED.	Weights for every pair involving this subclass are ignored.

I.5.14 SUBROUTINE WGTSCN

Message	Explanation
a. SYNTAX ERROR ON WEIGHT CARD -- REMAINDER OF CARD IGNORED.	Self-explanatory.
b. WEIGHT BUFFER IS FILLED -- ONLY XXXXX CLASS NAME PAIRS ALLOWED.	Buffer storage is inadequate for all class name pairs.

I.5.15 SUBROUTINE WHRPLC

Message	Explanation
THE INCLUDE REQUEST FOR CHANNEL N IS NOT A LEGITIMATE REQUEST -- IGNORED.	The indicated channel to be included is not among the input channels.

I.6 CLASSIFY

I.6.1 SUBROUTINE CLSFY1

Message	Explanation
***** CLSFY/CLSFY1 -- THE COVARIANCE MATRIX FOR SUBCLASS XX IS EITHER SINGULAR OR NOT POSITIVE DEFINITE -- THE DETERMINANT = XXXX.XXXX ***** TERMINATING PROGRAM EXECUTION *****	The determinant of each subclass covariance matrix is checked by CLASSIFY to see that it is a positive value. A zero value indicates a singular matrix, and a negative value indicates a nonpositive definite matrix. If either condition occurs for any subclass covariance matrix to be used in classification, the processor will stop. (NOTE: A probable source of an invalid covariance matrix is a module STAT deck which has been incorrectly formatted and thus is not producing good training class statistics. Another possible source is a SAVTAP file which does not contain valid statistical data.)

I.6.2 SUBROUTINE CLSFY2

Message	Explanation
a. WIDTH OF RECTANGULAR FIELD SURROUNDING CLASSIFICATION FIELD CANNOT EXCEED 1000 POINTS.	The difference between the largest sample number of the classification field and the smallest sample number of the classification field cannot exceed 1000. Reduce count of samples per scan line.
b. AS THE COMPUTER CANNOT EXPONENTIATE A NUMBER SMALLER THAN EXP (-88), XXXXXX PTS. WERE NOT CLASSIFIED IN THIS FIELD.	Self-explanatory.
c. TOO MUCH DATA REQUESTED.	When too much data has been requested, (1) for the standard classifier, reduce parameters so that $\left(\text{number of subclasses} - 1\right)\left(\frac{\text{number of subclasses} - 2}{2}\right) + \left(\text{number of subclasses}\right) + \left(\text{points per scan line}\right) \times \left(\text{number of channels}\right) \leq 12\,500$; or (2) for category classifier, reduce data so that the number of points per scan line × number of channels $\leq$ 12 500.

I.6.3 SUBROUTINE REDIF2

	Message	Explanation
a.	ERROR ON CHANNELS CARD.	Check job setup and unit assignments.
b.	**** CLSFY/REDIF2 -- BAD CARD INPUT DETECTED ON ATTEMPT TO READ B-MATRIX INFORMATION AS DIRECTED BY CONTROL CARD... **** TERMINATING PROGRAM EXECUTION FROM REDIF2 ****	The input B-MATRIX control card is printed out as part of the error message. One of the data cards following it is incorrectly formatted. Check deck setup and B-matrix card file.
c.	**CLSFY/REDIF2 -- B-MATRIX INPUT FROM BMFILE -- BAD VALUES DETECTED: NO. COMBINATIONS (BMCOMB) = ______, NO. CHANNELS (BMFEAT) = ______, CHANNEL VECTOR (BMVEC) = ______. ***** TERMINATING PROGRAM EXECUTION FROM REDIF2 *****	Invalid data from the BMFILE have been deleted.
d.	***CLSFY/REDIF2 -- BAD CARD INPUT ON APRIORI CARD -- DEFAULT A PRIORI PROBABILITY VALUES WILL BE USED.	Check format of APRIORI card.
e.	AT LEAST TWO (2) CATEGORIES MUST BE ASSIGNED. EXITING FROM REDIF2.	In exercising the category option, two or more categories must be used.
f.	**** CLSFY/REDIF2 -- BAD PROCESSOR CONTROL CARD **** TERMINATING PROGRAM EXECUTION FROM REDIF2 ****	Check spelling of keyword.

I.6.4 SUBROUTINE SETUP2

Message	Explanation
a. AN ERROR HAS OCCURRED IN GROUPING CLASSES INTO CATEGORIES. CHECK THE FOLLOWING: NOT ALL OF THE CLASSES HAVE BEEN ASSIGNED TO A CATEGORY. A CLASS NAME ON THE CATEGORY CARD HAS BEEN MISSPELLED. CLASS NAMES FROM SAVTAP FILE ARE: ______. CLASS NAMES FROM CATEGORY CARDS ARE: ______.	When an error occurs in grouping classes into categories, either one or more class names (1) have not been assigned or (2) have been misspelled. The program lists the class names as submitted from the SAVTAP file or the module STAT deck. Check these for errors. If neither (1) nor (2) is applicable, check the module STAT deck to assure that class names are left justified in the field.
b. USER-INPUT A PRIORI VALUES DO NOT SUM TO 1.0. INPUT VALUES WERE NORMALIZED.**	Self-explanatory.
c. ** ERROR IN A PRIORI CONTROL CARD. USER-INPUT VALUES IGNORED.**	Check format of APRIORI card.
d. NO. OF CHANNELS REQUESTED FROM DATA TAPE AND NO. OF CHANNELS ON STAT FILE MUST BE EQUAL.	Self-explanatory.

I.7 DISPLAY

I.7.1 SUBROUTINE DISTCV

Message	Explanation
OVERFLOW.	Error flag set by CHIN indicates overflow condition.

I.7.2 SUBROUTINE DISPLAY

Message	Explanation
** DISPLAY ** FIELDS MUST BE DEFINED FOR SUBCLASSES FOR EMPIRICAL THRESHOLDS.	Self-explanatory.

I.7.3 SUBROUTINE DSPLY1

Message	Explanation
NOT ENOUGH STORAGE FOR COVARIANCE MATRICES -- DSPLY1.	Adjust parameters.

I.7.4 SUBROUTINE DSPLY2

Message	Explanation
a. DISPLAY WILL ACCEPT ONLY 1000 PTS/SCAN LINE.	Adjust parameters.
b. **** DSPLY2/DOTSUM -- DISCREPANCY IN DOT FILE INFORMATION **** NO. OF DOT CATEGORY LABELS MATCHING MAPTAP CATEGORY NAMES = XXXXXX. NO. OF DOT CATEGORIES IS GIVEN AS = XXXXXX.	Check if correct dot file is used.
c. **** DSPLAY/DSPLY2 **** NO. OF DOTS = XXXXXX -- EXCEEDS THE MAX. ALLOWABLE (250) **** DOT PERFORMANCE SUMMARIES WILL NOT BE PRODUCED.	Maximum number of dots exceeded.

I.7.5 SUBROUTINE EMTHRS

Message	Explanation
ERROR BACK SPACING MAPTAP. ISTAT = XXXXX.	Hardware tape-read error.

I.7.6 SUBROUTINE FDIST

Message	Explanation
FDIST -- OVERFLOW CONDITION IN FISHIN ROUTINE FOR SUBCLASS = XXXX. THRESHOLD SET TO 999.999.	The FISHIN system subroutine has returned an overflow condition. The threshold value is set to 999.999 by the program.

I.7.7 SUBROUTINE PRTSUM

Message	Explanation
THE CROP NAME XXXX DOES NOT MATCH A CATEGORY, CLASS, OR SUBCLASS NAME. THE INTENSIVE TEST SITE SUMMARY REPORT CANNOT BE PRINTED.	Check spelling on CROP cards.

I.7.8 SUBROUTINE REDIF3

Message	Explanation
a. **** DSPLY/REDIF3 -- ERROR IN 'OPTION' CARD ··· **** SCAN OF THIS CARD DISCONTINUED -- PROCEEDING TO NEXT CARD ****	Check format and spelling of parameter.

	Message	Explanation
b.	***** FISHER THRESHOLD REQUESTED -- NOT PERFORMED -- NO. SAMPLES FOR SUB-CLASS NAME (=N) IS LESS THAN OR EQUAL TO NUMBER OF CHANNELS (=M).	The program compares the number of samples to the number of channels. If the number of samples $\leq$ number of channels, the threshold request is bypassed.
c.	ERROR IN ACREAGE CARD -- CARD IGNORED.	Check format.
d.	*** A THRESHOLD VALUE IS OUTSIDE THE ALLOWABLE RANGE 0 - 1, THEREFORE NO THRESHOLDING HAS BEEN DONE IN THIS RUN *** XXXXX, XXXXXXXXX.XXXXX	Check format of THRESHOLD card. The first number is the cluster number. The second number is the threshold value for that cluster.
e.	* ERROR ON SUBCLASS NAME CARD XXXX DOES NOT MATCH A SUBCLASS FROM THE MAPTAP FILE.	Self-explanatory.
f.	* ERROR ON CLASS NAME CARD XXXX DOES NOT MATCH A CLASS NAME FROM THE MAPTAP FILE *	Self-explanatory.
g.	INVALID CONTROL CARD -- CHECK SPELLING OF KEYWORD.	Self-explanatory.

I.7.9 SUBROUTINE SETUP3

	Message	Explanation
a.	***** DISPLAY/SETUP3 -- ERROR CONDITION ON ATTEMPT TO POSITION MAPTAP OVER ______ FILES ***** FSBSFL STATUS CODE = ________ -- ABORTING RUN *****	The system routine for positioning files (FSBSFL) has encountered difficulties in positioning MAPTAP to the correct file. The error occurred in the SETUP3 routine for DISPLAY. User should make sure that the correct file number for MAPTAP has been indicated and that MAPTAP does in fact have the correct number of files.
b.	****** DISPLAY/SETUP3 -- CORE OVERFLOW (TOP - TOP2) BY XXXXXX -- EXECUTION TERMINATED ******	Subroutine SETUP3 has computed the storage needed for the specific problem; if more is needed than is available, this diagnostic is printed.
c.	CLASSIFICATION BY CATEGORY (ON MAPTAP) IS REQUIRED IN ORDER TO PROCESS THE DOT DATA *** *** DOT PERFORMANCE SUMMARIES WILL NOT BE OUTPUT ***	Procedure 1 uses category classifier only.

I.8 DATATR

I.8.1 SUBROUTINE LNTRAN

	Message	Explanation
a.	*** THE NUMBER OF COMPONENTS IN Y-VECTOR TIMES THE NUMBER OF SAMPLES EXCEEDS THE SIZE OF STORAGE AREA -- TERMINATING ***	Self-explanatory.
b.	*** NUMBER OF CHANNELS TIMES NUMBER OF SAMPLES EXCEEDS 10 600 ***	Storage exceeded. Adjust parameters.
c.	***** DATATR/LNTRAN ***** ERROR ON INPUT FIELD DEFINITION CARD, FOR FIELD NAME XXXX ***** CONTINUING TO NEXT FIELD DEFINITION CARD.	Check format and parameters.

I.8.2 SUBROUTINE SETREM

Message	Explanation
SETREM ERROR -- THERE WERE XX SCALE FACTORS AND MINIMUM VALUES INPUT THROUGH THE SCAFAC OPTION. YY LINEAR COMBINATIONS WERE REQUESTED. THERE MUST BE A SCALE FACTOR AND A MINIMUM VALUE FOR EACH LINEAR COMBINATION. THE PROGRAM WILL TERMINATE THROUGH CMERR.	This message indicates that the number of input scaling parameter pairs does not correspond to the number of components of the transformed data. Too many or too few pairs were input.

I.8.3 SUBROUTINE SETUP8

	Message	Explanation
a.	*** BAD CONTROL CARD -- DATATR/SETUP8 ***	Check spelling of keyword.
b.	***** DATATR/SETUP8 ***** ERROR ON INPUT DATA CARD -- CONTINUING TO PROCESS INPUT *****	Check control card, correct, and resubmit.
c.	***** DATATR/SETUP8 ***** ERROR ON INPUT STATFILE CARD -- CONTINUING TO PROCESS INPUT *****	Check control card, correct, and resubmit.
d.	*** INVALID CONTROL CARD REJECTED BY DATATR/ SETUP8 ***	Check spelling of parameter.

I.9 TRSTAT

I.9.1 SUBROUTINE SETUP9

	Message	Explanation
a.	ERROR ON STATFILE CARD.	Check spelling of keyword and parameters.
b.	NUMBER OF CHANNELS FROM STAT FILE DOES NOT EQUAL THE NUMBER OF CHANNELS ON A-MATRIX FILE. CHANNELS ON STAT FILE = _______. CHANNELS ON A-MATRIX = _______.	Self-explanatory.
c.	*** BAD SUPERVISOR CONTROL CARD SETUP9 ***	Invalid control card. Check spelling of keyword.

Message	Explanation
d. INVALID CONTROL CARD REJECTED *** SETUP9 ***	The parameter field of the control card is in error.

I.9.2 SUBROUTINE TRAMTX

Message	Explanation
ERROR IN TRYING TO POSITION TRANSFORMED STAT FILE TO BEGINNING OF FILE XXX.	Check file assignment.

I.10 NDHIST

I.10.1 SUBROUTINE ADDRES

Message	Explanation
TOO MUCH DATA REQUESTED. REDUCE NO. OF SAMPLES PER SCAN LINE AND/OR NO. OF CHANNELS.	Self-explanatory.

I.10.2 SUBROUTINE NDHST1

Message	Explanation
a. N VECTORS WERE NOT HISTO-GRAMMED, BUT USED IN COMPUTING FIELD MEANS, IF APPLICABLE.	The histogrammed vector table is full. N unique vectors were not histogrammed.
b. ERROR IN FIELD CARD. ABORTING.	Check format and parameters.

I.10.3 SUBROUTINE NDHST2

Message	Explanation
CORE LIMITS EXCEEDED. MAXIMUM NO. OF VECTORS ACCEPTED IS _____.	Self-explanatory.

I.10.4 SUBROUTINE RESTO

Message	Explanation
ERROR READING DISK.	Self-explanatory.

I.10.5 SUBROUTINE SET10

Message	Explanation
a. ERROR ON CHANNELS CARD.	Check parameter field of CHANNELS control card.
b. ERROR ON DATA FILE CARD.	Check parameter field of DATA control card.
c. ERROR ON DAS FILE CARD.	Check parameter field of MAPFIL control card.
d. ERROR ON N-DIM HISTOGRAM FILE CARD.	Check parameter field of HISFIL control card.
e. ERROR ON OPTION CARD.	Check parameter field of the OPTION control card printed out just above this message.
f. INVALID CONTROL CARD -- IGNORED.	Check spelling of keyword.

I.10.6 SUBROUTINE STODAT

Message	Explanation
NOT ENOUGH DRUM SPACE TO STORE DAS TAPE DATA.	Actually refers to the disk space used to store the MAPUNT file.

I.11 SCTRPL

I.11.1 SUBROUTINE LINPLT

Message	Explanation
A TOTAL OF _____ POINTS WERE NOT DISPLAYED ON THE LINE-PRINTER GRAPH. THE POINTS WERE OUT OF RANGE OF EITHER THE X DIRECTION OR THE Y DIRECTION.	Data may be rescaled to a resolution of 100.

I.11.2 SUBROUTINE SETADR

Message	Explanation
NOT ENOUGH DISK SPACE. TOTAL WORDS OF DISK SPACE = XXXXXXXXXXXX. TOTAL WORDS OF DISK SPACE AVAILABLE = XXXXXXXXXXXX.	Adjust parameters.

I.11.3 SUBROUTINE SET11

Message	Explanation
a. ERROR IN POSITIONING NHSTUN FILE TO FILE _____.	Physical tape error occurred. Resubmit run.
b. ERROR ON N-DIM HISTOGRAM FILE CARD.	Check parameter field of HISFIL card.
c. ERROR ON CHANNELS CARD.	Check parameter field of CHANNELS card.
d. ERROR ON STATFILE CARD.	Check parameter field of STATFILE card.
e. ERROR ON B-MATRIX CARD.	Check spelling and parameter.

Message	Explanation
f. NO. OF PLOTTING CHANNELS AND NO. OF B-MATRIX CHANNELS MUST BE EQUAL. CHANNELS ARE _____ AND _____, RESPECTIVELY.	Number of channels to be transformed must equal the number of channels in transformed matrix.
g. ERROR ON SCALING CARD.	Check parameter field of SCALE cards.
h. ERROR ON TAPE SIZE CARD.	Check parameter field of SIZE card.
i. ERROR ON SCATTER PLOT TAPE CARD.	Check parameter field of PLOTAP card.
j. ERROR ON OPTION CARD.	Check parameter field of PIXPLT cards.
k. DATA MUST BE RESCALED BEFORE PIXEL FREQUENCY PLOT OPTION MAY BE SELECTED.	Transformed data must be rescaled for line-printer plot.
l. INVALID CONTROL CARD -- IGNORED.	Check spelling of keyword.

I.11.4 SUBROUTINE SORTVC

Message	Explanation
ERROR IN SORTING VECTORS.	Self-explanatory.

I.11.5 SUBROUTINE VECSCN

Message	Explanation
ERROR OCCURRED SCANNING VECTOR CARD.	Check keypunching error on cards.

I.12 DOTDATA

I.12.1 SUBROUTINE DOTS

Message	Explanation
**** NOTE -- TOTVEC WAS GREATER THAN 250, THEREFORE TOTVEC WAS SET TO 250 ****	Total number of dots allowable is 250 of type 1 and 250 of type 2.

I.12.2 SUBROUTINE SET13

Message	Explanation
a. INVALID CONTROL CARD -- IGNORED.	Check spelling of keyword.
b. ERROR ON DATA CARD.	Check parameter field.
c. ERROR ON DOTFIL CARD.	Check parameters.
d. ERROR ON OPTION CARD.	Check format and parameters.

I.13 LABEL

I.13.1 SUBROUTINE ALLKIN

Message	Explanation
a. LABELING BY ALL-OF-A-KIND PROCEDURE.	Supervisory messages.
b. ** DEFAULTING TO K-NEAREST-NEIGHBOR PROCEDURE **	
c. A TIE OCCURRED. THE FOLLOWING DOTS WERE DISCARDED.	

I.13.2 SUBROUTINE FILERD

Message	Explanation
a. NOT ENOUGH CORE TO STORE DOT FILE.	Revise data par[illegible]rs.

Message	Explanation
b. NOT ENOUGH DISK SPACE FOR CLUSTER MAP INFO.	Revise data parameters.

I.13.3 SUBROUTINE KNEAR

Message	Explanation
a. LABELING BY XXX NEAREST NEIGHBOR PROCEDURE.	Supervisory messages.
b. A TIE OCCURRED. THE FOLLOWING DOTS WERE DISCARDED.	

I.13.4 SUBROUTINE LABLR

Message	Explanation
CATEGORIES HAVE NOT BEEN DEFINED.	Check spelling and keywords on CATEGORY control card.

I.13.5 SUBROUTINE MANORD

Message	Explanation
a. ERROR IN INPUT OF CLASS NAMES. NAMES ON STAT FILE ARE:	Self-explanatory.
b. NAMES INPUT ARE:	Self-explanatory.

I.13.6 SUBROUTINE SET14

Message	Explanation
a. ERROR ON CHANNELS CARD.	Check format and parameters.
b. NO. OF STAT CHANNELS AND DOT DATA CHANNELS MUST BE EQUAL.	Self-explanatory.

316

	Message	Explanation
c.	ERROR ON STATFILE CARD.	Check format and parameters.
d.	ERROR ON MAPFIL CARD.	Check format and parameters on both the MAPFIL and the DOTFIL card.
e.	USER HAS NOT INPUT ONE OF THE REQUIRED FILES: SAVTAP, MAPUNT, OR DOTUNT.	Control cards missing.
f.	ERROR ON MAPTAP CARD.	Check format and parameters.
g.	ERROR ON PROCEDURE CARD.	Check spelling and parameter.
h.	A LABELING PROCEDURE MAY NOT BE CHOSEN WHEN UPDATING THE DOTUNT OR SAVTAP FILE.	You may want to go through $LABEL again after files have been updated.
i.	ERROR ON OPTION CARD.	Check parameter.
j.	INVALID CONTROL CARD -- IGNORED.	Check spelling of keyword.

I.13.7 SUBROUTINE STOMAP

Message	Explanation
NOT ENOUGH DISK SPACE TO STORE DAS TAPE DATA.	A MAPTAP output unit should be assigned.

I.14 DAMRG - SET18

	Message	Explanation
a.	INVALID CARD -- IGNORED.	Self-explanatory.
b.	ERROR ON ABOVE INPUT CONTROL CARD.	Self-explanatory.

Message	Explanation
c. ERROR IN FIELD -- OR $END CARD MISSING.	After the *END card, at least one field definition card must appear.
d. NUMBER OF CHANNEL CARDS -- DOES NOT MATCH NUMBER OF DATA FILES ______.	Self-explanatory.
e. NUMBER OF SUN ANGLE CARDS -- DOES NOT MATCH NUMBER OF DATA FILES _____.	Self-explanatory.
f. NUMBER OF FEATURES ON -- FILE -- IS NOT EQUAL TO NUMBER OF FIRST FILE ______.	Self-explanatory.
g. FEATURES ADD UP TO A NUMBER GREATER THAN 30 ______. EXITING.	EOD-LARSYS is restricted to 30 channels.

I.15 GTDDM

I.15.1 ALPHA

Message	Explanation
THE SYMBOL ___ CANNOT BE USED.	The characters must be alphabetic.

I.15.2 SET19

Message	Explanation
a. INVALID CONTROL CARD -- IGNORED.	Self-explanatory.
b. ERROR ON TRANSFORMATION CARDS. DEFAULT TRANSFORMATION USED.	Self-explanatory.

Message	Explanation
c. ERROR ON READ TAPE CARD.	Either the word UNIT or the word FILE is missing from the READ control card.
d. ERROR ON WRITE FILE CARD.	Either the word UNIT or the word FILE is missing from the WRITE control card.
e. ERROR ON MASK CARD/// TRANSITION YEAR MASK USED.	Self-explanatory.
f. ERROR ON CONVERT CARD.	DUMP card is not written properly. The word FILE has been omitted.

I.16 GTTCN - SET17

Message	Explanation
a. INVALID CONTROL CARD -- IGNORED.	Self-explanatory.
b. ERROR ON CONVERT CARD.	Self-explanatory.
c. ERROR ON OPTION CARD.	Self-explanatory.
d. ERROR ON WRITE TAPE CARD.	Self-explanatory.

I.17 UTILITY SUBPROGRAMS

I.17.1 SUBROUTINE BUFILL

Message	Explanation
XXXX BYTES EXPECTED. XXX BYTES ON RECORD.	Self-explanatory.

319

I.17.2 SUBROUTINE CLSCHK

	Message	Explanation
a.	** CLSCHK ** REQUESTED SUBCLASS NO. XXX IS NOT AVAILABLE IN INPUT STATISTICS FILE -- REQUEST IGNORED.	Either by the SUBCLASS control card or by default, a subclass number has been requested for use in classification which is greater than the largest subclass number available in the input training subclasses. The CLASSIFY processor ignores the requested subclass number and deletes it from use in classification.
b.	** CLSCHK ** REQUESTED SUBCLASS NO. XXX FOR GROUP NO. XXX IS NOT AVAILABLE IN INPUT STATISTICS FILE -- REQUEST IGNORED.	A subclass number input by either the SUBCLASS or the GROUP control card is greater than the largest training subclass number available. The requested subclass will be ignored in classification.
c.	***** CLSFY/FETCHK -- CHANNEL XX NOT IN TRAINING DATA -- TRAINING DATA CHANNELS ARE $C_1, C_2, C_3, \cdots, C_N$.	A channel requested on the CHANNELS control card or in the B-matrix input for use in classification is not in the set of channels which was used to obtain training subclass statistics. The available set of channels is printed out as part of the diagnostic message. Message d or e is added to this, d when the B-matrix is not involved and e when B-matrix channels are input.

Message	Explanation
d. ***** CHANNEL XX IGNORED (NOT USED) IN CLASSIFICATION.	If the B-matrix is not being input to the CLASSIFY processor (i.e., B-matrix channels are not involved), the requested channel causing the previous message will be deleted from the list of channels and ignored by the CLASSIFY processor. In this case, message d is added to message c.
e. ***** B-MATRIX CHANNELS MUST BE EQUAL TO OR A SUBSET OF AVAILABLE TRAINING DATA CHANNELS -- THE INPUT B-MATRIX CHANNEL SET IS $B_1, B_2, \cdots, B_M$.	If the B-matrix is input to the CLASSIFY processor, the B-matrix channels become the set to be used in classification; and if one of the B-matrix channels is not in the training subclass statistics, the processor cannot continue. In this case, message e will be added to message c.

I.17.3 SUBROUTINE CMERR

Message	Explanation
ERROR HAS OCCURRED.	Self-explanatory.

I.17.4 SUBROUTINE CRDSTA

Message	Explanation
EXCEEDED CORE LIMITS. REDUCE NO. OF TRAINING CLASSES OR FEATURES. EXITING FROM CRDSTA.	The combination of total number of channels, subclasses, and training fields must be reduced to fit in the internal core storage available to the processor. Total storage is 10 600 locations.

I.17.5 SUBROUTINE DSTAPE

Message	Explanation
THE NUMBER OF CHANNELS TIMES THE NUMBER OF SAMPLES HAS EXCEEDED 11500. DECREASE THE NUMBER OF CHANNELS OR THE NUMBER OF SAMPLES. TERMINATING RUN FROM DSTAPE.	Storage available has been exceeded.

I.17.6 SUBROUTINE FLDINT

Message	Explanation
a. FEATURE NUMBERS XXXXX AND ABOVE ARE NOT ON DATA TAPE.	User has requested a channel not on MSS DATAPE.
b. FIRST SCAN ON THIS TAPE IS NUMBERED XXXXXX. FIELD DEFINITION IN ERROR.	Self-explanatory.
c. NUMBER OF SAMPLES PER SCAN LINE ON THIS TAPE IS XXXXXX. FIELD DEFINITION IN ERROR.	Self-explanatory.

Message	Explanation
d. THIS TAPE CONTAINS ONLY XXXXXX CHANNELS.	Self-explanatory.

I.17.7 SUBROUTINE FSBSFL

Message	Explanation
FSBSFL ONLY SKIPS FORWARD.	Self-explanatory.

I.17.8 SUBROUTINE FSFMFL

Message	Explanation
FSFMFL ONLY SKIPS FORWARD.	Self-explanatory.

I.17.9 SUBROUTINE GETST

Message	Explanation
a. ERROR IN POSITIONING UNIT XXX TO FILE XXX.	Self-explanatory.
b. REQUESTED SUBCLASS IS NOT ON STAT FILE. STAT FILE CONTAINS XXX SUBCLASSES.	Self-explanatory.
c. CHANNEL NO. XX IS NOT ON TRAINING STAT FILE. CHANNELS ARE $C_1, C_2, C_3, \cdots, C_N$.	Self-explanatory.

323

I.17.10 SUBROUTINE GRPSCN

Message	Explanation
///// FROM SUBR. GRPSCN -- CLASS XXXXX INCORRECT -- CLASS XXXXX IGNORED. CARD BEING SCANNED IS XXXX···XXXX.	One of the class numbers listed on the GROUP control card (1) is not in ascending order, (2) is greater than the largest class number allowable (30), or (3) has already been used in another group. The erroneous GROUP control card is printed as part of the message. The processor will delete the erroneous class number from the list and proceed to group all other listed classes.

I.17.11 SUBROUTINE HISTGM

Message	Explanation
TOO MUCH DATA REQUESTED -- SAMPLE END WAS RESET TO XXXXX.	The data for all channels for one scan line are unpacked into an array dimensioned 12 000. If the number of channels times [(sample end - sample begin)/sample increment] exceeds 12 000, this diagnostic is printed. Sample end is reset to fit the dimensions and execution continues.

I.17.12 SUBROUTINE HISTIC

Message	Explanation
ONLY THE FIRST 50 FIELD DESCRIPTIONS WERE PRINTED, BUT ALL THE FIELDS WERE INCLUDED IN THE TOTAL HISTOGRAMMED STATS.	The user has input more than 50 fields, and only the first 50 field descriptions will be printed in the "Data Blocks Histogrammed" portion of the total report; however, all the input fields were included in the calculations of the "Total Histogrammed Statistics."

I.17.13 SUBROUTINE I4A1BN

Message	Explanation
EBCDIC TO BINARY INTEGER CONVERSION ERROR AT CHARACTER XXXXX OF XXXXX CHARACTER FIELD: XXXX···XXXX.	EBCDIC is the Extended Binary Coded Decimal Interchange Code.

I.17.14 SUBROUTINE LABMAN

Message	Explanation
ERROR IN POSITIONING SIG. EXTENSION TAPE TO FILE XXX. OUTPUT FILE NOT WRITTEN.	Self-explanatory.

I.17.15 SUBROUTINE LAREAD

Message	Explanation
a. ERROR IN FIELD CARD. TERMINATING RUN.	A field description card has an incorrect format. All vertices must be separated by commas and enclosed in parentheses, and sample and line numbers must be integers. The card which caused the error is printed out with this message.
b. INCORRECT FIELD CARD. TERMINATING RUN.	Same.

I.17.16 SUBROUTINE LINERD

Message	Explanation
a. FIELD BOUNDARY FOR THIS FIELD DEFINED BEYOND SCOPE OF DATA. THIS FLIGHT LINE CONTAINS XXXXXX SCAN LINES.	User has requested scan line not on MSS DATAPE.
b. FLDINT MUST BE CALLED TO INITIALIZE PARAMETERS FOR A NEW FIELD.	For every field input there must be a call to FLDINT to reset parameters for positioning the MSS DATAPE.

I.17.17 SUBROUTINE RANK

Message	Explanation
THE NUMBER OF CHANNELS ARE NOT A MULTIPLE OF 4. THE COLOR KEYS WILL BE ORDERED BY CLUSTER NUMBER.	Currently used greenness/ brightness transformations require four channels per pass.

I.17.18 SUBROUTINE RDDOTS

Message	Explanation
a. CHANNEL XX IS NOT ON DOT FILE. CHANNELS ARE $C_1, C_2, \cdots, C_N$.	Self-explanatory.
b. DOT NO. XXX IS NOT ON DOT FILE. FILE CONTAINS XXX DOTS.	Self-explanatory.

I.17.19 SUBROUTINE RDMEAN

Message	Explanation
MEANS FOR CHANNEL XXXX ARE NOT ON FILE -- DUMMY VALUES WILL NOT BE USED.	Self-explanatory.

I.17.20 SUBROUTINE RDMODK

Message	Explanation
ERROR IN TRYING TO POSITION STAT FILE TO FILE XXX IN CRDSTA.	An error occurred in positioning the SAVTAP file, and no statistics were written. Resubmit the run.

I.17.21 SUBROUTINE REDDAT

Message	Explanation
CHANNEL NO. XX IS NOT A TRAINING CHANNEL. XX TRAINING CHANNELS ARE $C_1, C_2, \cdots, C_N$.	Self-explanatory.

I.17.22 SUBROUTINE REDSAV

Message	Explanation
a. STAT FILE WAS NOT CREATED. EXITING FROM **REDSAV**	An error occurred in positioning the SAVTAP file, and no statistics were written. Resubmit run.
b. ERROR IN POSITIONING STAT FILE TO FILE XXX. EXITING FROM REDSAV.	Same.
c. USER HAS REQUESTED XX CHANNELS, XX SUBCLASSES, AND XX CLASSES. THIS COMBINATION OF STATS WILL NOT FIT IN CORE. PLEASE REDUCE REQUEST.	The fixed amount of internal core storage available to the processor for storing class descriptions, number of subclasses in each class, subclass descriptions, field information, vertices, covariances, means, and working area has been exceeded. The total amount of storage available for the above information is 10 600 locations.* Reduce the requested combination.

I.17.23 SUBROUTINE SEARCH

Message	Explanation
a. SEARCHING FOR LINE.	Self-explanatory.
b. RECORDS PER SCAN, XXXXX. SCANS PER RECORD, XXXXX.	Self-explanatory.

*The equation for computing the required storage is STORAGE = 2(number of classes) + (number of subclasses) + 4(number of fields) + 2(total number of vertices for all the fields) + (number of subclasses + 1)[(number of channels)(number of channels + 1)/2] + (number of subclasses)(number of channels).

	Message	Explanation
c.	FOUND IT AFTER XXX TRIES.	Self-explanatory.
d.	FAILED AFTER XXXXX TRIES -- ABORTING.	Self-explanatory.
e.	SCAN XXXXX IS MISSING -- USING PREVIOUS SCAN INSTEAD.	Self-explanatory.

I.17.24 SUBROUTINE SETUP7

	Message	Explanation
a.	ERROR ON CHANNEL CARD.	Check format.
b.	ERROR ON DATA CARD.	Check format.
c.	ERROR ON STATFILE CARD.	Check format.
d.	ERROR ON DOTFIL CARD.	Check format.
e.	INVALID INPUT -- CARD ______ IGNORED.	Check table 9-1 for correct spelling of keywords for card input and make sure the keyword is left justified in the field.
f.	CHANNELS CANNOT BE CHANGED UNTIL THIS EXECUTION OF ISOCLS IS COMPLETED.	The channels to be used should be included in the first set of control cards input after the ISOCLS card. That set of channels will be used for all classes. If the user attempts to input a CHANNELS card into the SETUP routine on a later entry, the card will be ignored.

	Message	Explanation
g.	NO. OF CLASSES CANNOT BE CHANGED UNTIL THIS EXECUTION OF ISOCLS IS COMPLETED.	The number of classes to be clustered must be input only in the first set of control cards input after the ISOCLS card. If the user attempts to change this parameter, the input will be ignored.
h.	**WARNING** NMIN IS LESS THAN NO. OF CHANNELS, COVARIANCES WILL NOT BE INVERTIBLE.	NMIN should be made greater than the total number of channels.

I.17.25 SUBROUTINE TAPHDR

	Message	Explanation
a.	UNRECOVERABLE ERROR READING HEADER RECORD.	Error occurred while trying to read header record.
b.	A LINE NO. IS LESS THAN OR EQUAL TO ZERO.	The first line number on the data tape is less than or equal to zero.
c.	LAST SCAN LINE READ XXXXX. ISTAT = XXXXX.	Self-explanatory.
d.	INTERNAL DIMENSIONS TOO SMALL FOR DATA. NUMBER OF CHANNELS ON DATA TAPE = XXXXXXX. NUMBER OF POINTS/CHANNEL = XXXXXXX.	The maximum record size of the data record exceeded 6800 words.
e.	DATA TAPE IS NOT IN UNIVERSAL OR LARSYS FORMAT.	The MSS DATAPE must be in Universal or LARSYS II/III format.

Message	Explanation
f. ONLY ONE RECORD OR LESS PER CHANNEL ACCEPTABLE AT THIS TIME.	All of the samples of one channel must be contained within one record.
g. NO. OF RECORDS PER DATA SET = XXXXX. MUST BE LESS THAN OR EQUAL TO 15.	One data set cannot contain more than 15 records.
h. NO. OF BITS/PIXEL = XXXXX. ONLY 8 BITS ACCEPTABLE AT THIS TIME.	According to the header record, the samples on the MSS DATAPE do not equal 8 bits. It is assumed that the header record is in error, and execution continues.
i. DATA ORDER INDICATOR = XXXXX. DATA MUST BE ORDERED BY PIXEL.	Information from header record indicates data are not ordered properly. All radiance values in channel 1 must be given first, followed by the radiance values in channel 2, and so on for each scan line.

I.17.26 SUBROUTINE WRTHED

Message	Explanation
NUMBER OF SAMPLES WAS RESET TO 2998.	Number of samples per line cannot be more than 2998.

331

APPENDIX J

SAMPLE PROCEDURE 1 EXECUTION
OF EOD-LARSYS

** VM/370 VERSION 05 LEVEL 00 PLC 0012 ************** CLASS A ********** PRINTED 11/14/79 AT 11.45.41 AT HOUSTON **************

USER NAME	SUSANNE
USERID	JSC780
SPOOL FILE NAME TYPE	
CREATION DATE	11/13/79 16:01:16
SPOOL FILE ID	7135
RECORD COUNT	3588

LARS

JSC780

SUSANNE

```
ATA TAPE BEING USED IS 3915 , DENSITY IS 1600 BPI

                                                      LYNDON B. JOHNSON SPACE CENTER
                                                             HOUSTON, TEXAS

FORM
$DOTDATA

          INPUT SUMMARY

DATA      UNIT=11,FILE=2
DOTF      OUTPUT/UNIT=19,FILE=1
CHAN      DATA=1,2,3,4
COMM      PROCEDURE 1 RUN
OPTI      U    28
*END

USER HAS REQUESTED THE FOLLOWING OPTIONS :

SELECTED CHANNELS ARE  1  2  3  4
SKIPPING FILE        1

SUN ANGLES :     34     34     34     34     20     20     20   . 20

INPUT IMAGE DATA TAPE INFORMATION

    FORMAT                     UNIVERSAL
    NO. OF CHANNELS              16
    NO. OF PIXELS/LINE          196
    FIRST SCAN LINE NO.           1
    FIRST PIXEL REFERENCE PT      1

SITE =  1851 DAY=    1 MONTH=    4 YEAR=   77 MASK = TY  TYPE = GT
```

LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

PROCEDURE 1 RUN

INPUT FIELDS

FIELD		CLASS		SUBCLASS	VERTICES (SAMPLE,LINE)
1	SS	1	SS		(120, 10)
2	SS	1	SS		(20, 30)
3	SS	1	SS		(180, 30)
4	SS	1	SS		(110, 40)
5	SS	1	SS		(130, 40)
6	SS	1	SS		(80, 50)
7	SS	1	SS		(100, 50)
8	SS	1	SS		(160, 50)
9	SS	1	SS		(70, 60)
10	SS	1	SS		(190, 60)
11	SS	1	SS		(20, 70)
12	SS	1	SS		(60, 70)
13	SS	1	SS		(80, 70)
14	SS	1	SS		(10, 80)
15	SS	1	SS		(30, 80)
16	SS	1	SS		(130, 80)
17	SS	1	SS		(40, 90)
18	SS	1	SS		(60, 90)
19	SS	1	SS		(80, 90)
20	SS	1	SS		(100, 90)
21	SS	1	SS		(120, 90)
22	SS	1	SS		(10, 100)
23	SS	1	SS		(130, 100)
24	N	2	N		(20, 10)
25	N	2	N		(40, 10)
26	N	2	N		(60, 10)
27	N	2	N		(80, 10)
28	N	2	N		(100, 10)
29	N	2	N		(140, 10)
30	N	2	N		(160, 10)
31	N	2	N		(180, 10)
32	N	2	N		(10, 20)
33	N	2	N		(30, 20)
34	N	2	N		(50, 20)
35	N	2	N		(70, 20)
36	N	2	N		(90, 20)
37	N	2	N		(110, 20)
38	N	2	N		(130, 20)
39	N	2	N		(170, 20)
40	N	2	N		(40, 30)
41	N	2	N		(60, 30)
42	N	2	N		(80, 30)
43	N	2	N		(100, 30)
44	N	2	N		(120, 30)
45	N	2	N		(140, 30)
46	N	2	N		(160, 30)
47	N	2	N		(10, 40)
48	N	2	N		(30, 40)
49	N	2	N		(50, 40)
50	N	2	N		(70, 40)

```
51  N        2  N                                     (  90,  40)
52  N        2  N                                     ( 150,  40)
53  N        2  N                                     ( 170,  40)
54  N        2  N                                     ( 190,  40)
55  N        2  N                                     (  20,  50)
56  N        2  N                                     (  40,  50)
57  N        2  N                                     (  60,  50)
58  N        2  N                                     ( 120,  50)
59  N        2  N                                     ( 140,  50)
60  N        2  N                                     ( 180,  50)
61  N        2  N                                     (  10,  60)
62  N        2  N                                     (  30,  60)
63  N        2  N                                     (  50,  60)
64  N        2  N                                     (  90,  60)
65  N        2  N                                     ( 110,  60)
66  N        2  N                                     ( 150,  60)
67  N        2  N                                     ( 170,  60)
68  N        2  N                                     (  40,  70)
69  N        2  N                                     ( 100,  70)
70  N        2  N                                     ( 120,  70)
71  N        2  N                                     ( 160,  70)
72  N        2  N                                     ( 180,  70)
73  N        2  N                                     (  50,  80)
74  N        2  N                                     (  70,  80)
75  N        2  N                                     ( 110,  80)
76  N        2  N                                     ( 150,  80)
77  N        2  N                                     ( 170,  80)
78  N        2  N                                     ( 190,  80)
79  N        2  N                                     (  20,  90)
80  N        2  N                                     ( 140,  90)
81  N        2  N                                     ( 160,  90)
82  N        2  N                                     ( 180,  90)
83  N        2  N                                     (  30, 100)
84  N        2  N                                     (  50, 100)
85  N        2  N                                     (  70, 100)
86  N        2  N                                     (  90, 100)
87  N        2  N                                     ( 110, 100)
88  N        2  N                                     ( 150, 100)
89  N        2  N                                     ( 170, 100)
90  N        2  N                                     ( 190, 100)
91  N        2  N                                     (  20, 110)
92  N        2  N                                     (  40, 110)
93  N        2  N                                     (  60, 110)
94  N        2  N                                     (  80, 110)
95  N        2  N                                     ( 140, 110)
96  N        2  N                                     ( 160, 110)
97  M        3  M                                     (  90,  40)
SKIPPING FILE      1

SUN ANGLES :     34    34    34    34    20    20    20    20


INPUT IMAGE DATA TAPE INFORMATION

     FORMAT                  UNIVERSAL
     NO. OF CHANNELS            16
     NO. OF PIXELS/LINE        196
     FIRST SCAN LINE NO.         1
     FIRST PIXEL REFERENCE PT    1
```


336

LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

PROCEDURE 1 RUN

INPUT FIELDS

FIELD	CLASS	SUBCLASS	VERTICES (SAMPLE,LINE)
1 SS	1 SS		(60, 40)
2 SS	1 SS		(120, 40)
3 SS	1 SS		(70, 50)
4 SS	1 SS		(170, 50)
5 SS	1 SS		(60, 60)
6 SS	1 SS		(140, 60)
7 SS	1 SS		(30, 70)
8 SS	1 SS		(70, 70)
9 SS	1 SS		(150, 70)
10 SS	1 SS		(120, 80)
11 SS	1 SS		(90, 90)
12 SS	1 SS		(40, 100)
13 SS	1 SS		(120, 100)
14 SS	1 SS		(180, 100)
15 SS	1 SS		(90, 110)
16 N	2 N		(10, 10)
17 N	2 N		(30, 10)
18 N	2 N		(50, 10)
19 N	2 N		(70, 10)
20 N	2 N		(90, 10)
21 N	2 N		(110, 10)
22 N	2 N		(130, 10)
23 N	2 N		(150, 10)
24 N	2 N		(170, 10)
25 N	2 N		(190, 10)
26 N	2 N		(20, 20)
27 N	2 N		(40, 20)
28 N	2 N		(60, 20)
29 N	2 N		(80, 20)
30 N	2 N		(100, 20)
31 N	2 N		(140, 20)
32 N	2 N		(160, 20)
33 N	2 N		(180, 20)
34 N	2 N		(10, 30)
35 N	2 N		(30, 30)
36 N	2 N		(50, 30)
37 N	2 N		(70, 30)
38 N	2 N		(90, 30)
39 N	2 N		(110, 30)
40 N	2 N		(130, 30)
41 N	2 N		(150, 30)
42 N	2 N		(170, 30)
43 N	2 N		(190, 30)
44 N	2 N		(40, 40)
45 N	2 N		(80, 40)
46 N	2 N		(100, 40)
47 N	2 N		(160, 40)
48 N	2 N		(180, 40)
49 N	2 N		(10, 50)
50 N	2 N		(30, 50)

51	N	2	N	(50, 50)
52	N	2	N	(90, 50)
53	N	2	N	(110, 50)
54	N	2	N	(150, 50)
55	N	2	N	(190, 50)
56	N	2	N	(20, 60)
57	N	2	N	(40, 60)
58	N	2	N	(80, 60)
59	N	2	N	(100, 60)
60	N	2	N	(120, 60)
61	N	2	N	(160, 60)
62	N	2	N	(180, 60)
63	N	2	N	(10, 70)
64	N	2	N	(50, 70)
65	N	2	N	(90, 70)
66	N	2	N	(110, 70)
67	N	2	N	(130, 70)
68	N	2	N	(170, 70)
69	N	2	N	(20, 80)
70	N	2	N	(40, 80)
71	N	2	N	(60, 80)
72	N	2	N	(80, 80)
73	N	2	N	(100, 80)
74	N	2	N	(140, 80)
75	N	2	N	(160, 80)
76	N	2	N	(180, 80)
77	N	2	N	(10, 90)
78	N	2	N	(30, 90)
79	N	2	N	(50, 90)
80	N	2	N	(70, 90)
81	N	2	N	(110, 90)
82	N	2	N	(130, 90)
83	N	2	N	(150, 90)
84	N	2	N	(170, 90)
85	N	2	N	(190, 90)
86	N	2	N	(20, 100)
87	N	2	N	(60, 100)
88	N	2	N	(80, 100)
89	N	2	N	(100, 100)
90	N	2	N	(140, 100)
91	N	2	N	(160, 100)
92	N	2	N	(10, 110)
93	N	2	N	(30, 110)
94	N	2	N	(50, 110)
95	N	2	N	(70, 110)
96	N	2	N	(110, 110)
97	N	2	N	(130, 110)
98	N	2	N	(150, 110)
99	N	2	N	(190, 110)
100	M	3	M	(120, 20)
101	M	3	M	(130, 20)
102	M	3	M	(190, 20)
103	M	3	M	(140, 40)
104	M	3	M	(130, 50)
105	M	3	M	(130, 60)
106	M	3	M	(140, 70)
107	M	3	M	(190, 70)
108	M	3	M	(120, 110)
109	M	3	M	(170, 110)
110	M	3	M	(180, 110)
111	M	3	M	(100, 110)
112	M	3	M	(20, 40)

NO.	SAMPLE	LINE	TYPE	CATEGORY	DATA			
					CH(1)	CH(2)	CH(3)	CH(4)
1.	60	40	2	1	39	56	59	25
2.	120	40	2	1	34	43	49	20
3.	70	50	2	1	45	58	72	31
4.	170	50	2	1	39	50	49	22
5.	60	60	2	1	39	50	54	23
6.	140	60	2	1	32	37	48	20
7.	30	70	2	1	42	51	60	23
8.	70	70	2	1	44	56	59	23
9.	150	70	2	1	39	42	54	23
10.	120	80	2	1	44	45	64	30
11.	90	90	2	1	44	53	51	23
12.	40	100	2	1	42	51	65	27
13.	120	100	2	1	47	61	59	23
14.	180	100	2	1	39	43	49	22
15.	90	110	2	1	39	45	46	19
16.	10	10	2	2	22	19	78	39
17.	30	10	2	2	50	61	73	28
18.	50	10	2	2	39	48	59	25
19.	70	10	2	2	31	29	49	25
20.	90	10	2	2	32	29	56	25
21.	110	10	2	2	32	33	54	23
22.	130	10	2	2	32	33	54	22
23.	150	10	2	2	32	32	46	20
24.	170	10	2	2	37	32	64	25
25.	190	10	2	2	31	32	54	25
26.	20	20	2	2	34	33	49	22
27.	40	20	2	2	34	37	49	20
28.	60	20	2	2	34	30	76	36
29.	80	20	2	2	31	22	68	33
30.	100	20	2	2	39	37	46	19
31.	140	20	2	2	34	37	57	23
32.	160	20	2	2	34	33	57	22

33.	180	20	2	2	34	32	60	27
34.	10	30	2	2	31	32	54	22
35.	30	30	2	2	39	50	54	23
36.	50	30	2	2	34	30	70	35
37.	70	30	2	2	34	30	59	27
38.	90	30	2	2	34	37	59	25
39.	110	30	2	2	31	30	32	15
40.	130	30	2	2	34	29	51	25
41.	150	30	2	2	36	38	54	23
42.	170	30	2	2	36	30	64	25
43.	190	30	2	2	34	37	49	20
44.	40	40	2	2	34	42	62	28
45.	80	40	2	2	39	43	54	22
46.	100	40	2	2	32	37	43	19
47.	160	40	2	2	32	42	43	19
48.	180	40	2	2	31	38	54	22
49.	10	50	2	2	42	43	68	30
50.	130	50	2	2	32	32	57	26
51.	50	50	2	2	39	43	64	30
52.	90	50	2	2	39	40	54	23
53.	110	50	2	2	29	27	35	17
54.	150	50	2	2	29	30	40	19
55.	190	50	2	2	32	30	36	17
56.	20	60	2	2	37	35	59	27
57.	40	60	2	2	36	37	59	25
58.	80	60	2	2	32	27	68	31
59.	100	60	2	2	32	33	54	27
60.	120	60	2	2	39	38	51	22
61.	160	60	2	2	36	38	51	22
62.	180	60	2	2	32	38	51	22
63.	10	70	2	2	37	38	59	25
64.	50	70	2	2	54	71	73	30

340

65.	90	70	2	2	34	33	56	25
66.	110	70	2	2	34	48	56	23
67.	130	70	2	2	31	32	33	15
68.	170	70	2	2	34	42	56	25
69.	20	80	2	2	29	25	36	15
70.	40	80	2	2	29	29	44	20
71.	60	80	2	2	57	67	76	31
72.	80	80	2	2	29	30	43	17
73.	100	80	2	2	67	75	83	33
74.	140	80	2	2	27	27	36	17
75.	160	80	2	2	34	43	49	23
76.	180	80	2	2	34	37	64	27
77.	10	90	2	2	31	32	54	27
78.	30	90	2	2	26	29	40	15
79.	50	90	2	2	31	35	44	19
80.	70	90	2	2	31	30	48	20
81.	110	90	2	2	31	27	54	30
82.	130	90	2	2	36	45	59	27
83.	150	90	2	2	34	38	54	27
84.	170	90	2	2	39	45	59	23
85.	190	90	2	2	31	33	54	23
86.	20	100	2	2	37	42	73	30
87.	60	100	2	2	32	35	46	20
88.	80	100	2	2	31	32	36	17
89.	100	100	2	2	31	32	54	28
90.	140	100	2	2	32	43	43	17
91.	160	100	2	2	34	42	62	28
92.	10	110	2	2	34	40	52	22
93.	30	110	2	2	37	40	52	22
94.	50	110	2	2	42	50	57	23
95.	70	110	2	2	37	40	57	27

96.	110	110	2	2	34	33	60	30
97.	130	110	2	2	31	33	54	23
98.	150	110	2	2	31	30	54	23
99.	190	110	2	2	27	27	43	20
100.	120	20	2	3	45	53	60	23
101.	130	20	2	3	34	37	54	22
102.	190	20	2	3	32	32	46	22
103.	140	40	2	3	32	33	56	27
104.	130	50	2	3	34	37	49	20
105.	130	60	2	3	44	53	59	22
106.	140	70	2	3	34	37	54	23
107.	190	70	2	3	31	33	36	15
108.	120	110	2	3	42	50	54	23
109.	170	110	2	3	52	64	68	27
110.	180	110	2	3	50	58	64	27
111.	100	110	2	3	34	40	51	20
112.	20	40	2	3	34	35	51	23

ON 11/13/79 AT 16: 5:30 TIME FOR $DOT WAS - TOTAL 0.558 MINUTES VIRTUAL 0.424 MINUTES.

```
                                                    LYNDON B. JOHNSON SPACE CENTER
                                                          HOUSTON, TEXAS
          PROCEDURE 1 RUN

    $ISOCLS


    INPUT SUMMARY

         DATA     UNIT=11,FILE=2
         DOTF     INPUT/UNIT=19,FILE=1
         STAT     OUTPUT/UNIT=20,FILE=1
         CHAN     DATA=1,2,3,4
         CHAN     STAT=1,2,3,4
         SFP      1.0
         STDM     15.0
         DLMI     0
         NMIN      0
         PMIN      0
         ISTO     0
         FORM     UNIVERSAL
         OPTI     CLUSTER
         OPTI     STATS
         COMM     PROCEDURE 1 RUN
         DOTS         1,11,21,31,41,51
         *END


    YOU HAVF SELECTFD THE FOLLOWING PARAMETER VALUES AND   OPTIONS

    STOP AFTER     0 ITERATION(S)
    ALLOW A MINIMUM OF      0 PIXELS PER CLUSTER
    PRINT A CLUSTER SUMMARY EVERY   20 ITERATION(S)
    PRINT A CLUSTER MAP EVERY   20 ITERATION(S)
    ALLOW A MAXIMUM OF   60 CLUSTERS PER CLASS
    THE STATISTICS FILE WILL BE WRITTEN AFTER   1 CLASS(ES) HAVE BEEN CLUSTERED
    CHANNELS ARE---  1  2  3  4
    DLMIN =  0.0
    STDMAX= 15.000
    PERCENT =   80
    PMIN =     0
    NSDOTS =     6
    NO. SUN ANGLES FROM CARDS =     0
    SUN ANGLE TAPE SW =     0

    SFP    =  1.000
    WRITE A CLUSTER MAP OUTPUT TAPE IN UNIVERSAL FORMAT

    **WARNING**  NMIN IS LESS THAN NO. OF CHANNELS,COVARIANCES WILL NOT BE INVERTIBLE
    SKIPPING FILE       1


    INPUT IMAGE DATA TAPE INFORMATION

       FORMAT                     UNIVERSAL
       NO. OF CHANNELS               16
       NO. OF PIXELS/LINE           196
       FIRST SCAN LINE NO.            1
       FIRST PIXEL REFERENCE PT       1
```

LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

PROCEDURE 1 RUN

FIELDS TO BE CLUSTERED FOR CLASS WHEA

	FIELD NAME	SAMPLE INC.	LINE INC.	VERTICES (SAMPLE,LINE)
1	SFGH	10	10	(10, 10) (196, 10) (196, 117) (10, 117)

DO/DU CLUSTER POP FOR THIS CLASS 0 0

LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

PROCEDURE 1 RUN

FINAL CLUSTER SUMMARY FOR CLASS WHEA

TOTAL NUMBER OF CLUSTERS = 6

TOTAL NUMBER OF POINTS = 209

CLUSTER	SYMBOL	POINTS IN CLUSTER
1	1	44
2	2	16
3	3	17
4	4	61
5	5	19
6	6	52

MEANS

CLUSTER	CH(1)	CH(2)	CH(3)	CH(4)
1	31.61	32.93	56.59	25.39
2	45.19	55.06	58.31	24.06
3	36.47	44.94	48.65	20.35
4	36.31	38.08	58.59	26.13
5	46.16	54.16	70.37	29.84
6	31.54	32.98	42.48	18.79

STANDARD DEVIATIONS

CLUSTER	CH(1)	CH(2)	CH(3)	CH(4)
1	2.29	5.50	5.91	3.90
2	2.92	4.13	3.42	1.48
3	2.48	2.78	3.12	1.53
4	2.27	5.45	5.51	3.25
5	7.60	10.78	5.31	2.01
6	2.50	4.16	4.76	1.94

DISTANCES BETWEEN CLUSTERS

CLUSTER	1	2	3	4	5	6
1	0.0	7.04	4.61	2.30	5.32	3.58
2	7.04	0.0	5.85	5.06	4.40	8.88
3	4.61	5.85	0.0	3.95	8.10	4.44
4	2.30	5.06	3.95	0.0	4.11	4.86
5	5.32	4.40	8.10	4.11	0.0	9.13
6	3.58	8.88	4.44	4.86	9.13	0.0

ORIGINAL PAGE IS OF POOR QUALITY

LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

PROCEDURE 1 RUN

SEGM

TOTAL NUMBER OF POINTS IN THIS FIELD 209

```
       0000000001111111111
       1234567890123456789
       0000000000000000000

 10    1151461414111664441
 20    1656241166421161446
 30    1444454444564644466
 40    4114121466233163314
 50    5414565442666563346
 60    4414146141642544313
 70    4224522644116146456
 80    5646656625656613644
 90    1663626323134644441
100    5515166661125364631
110    1344214436441416526
```

POINTS PER CLUSTER IN THIS FIELD

CLUSTER	SYMBOL	POINTS
1	1	44
2	2	16
3	3	17
4	4	61
5	5	19
6	6	52

346

```
                                   LYNDON B. JOHNSON SPACE CENTER
                                          HOUSTON, TEXAS
PROCEDURE 1 RUN

                                     FILE NO.     -     1
                                     FIELD NAME   - SEGM
                                     FORMAT       - UNIVERSAL
                                     NO. OF SCAN LINES -     11
                                     NO. OF COLOR KEY SCAN LINES -      0
```

CLASS : WHEA
SUBCLASS: WH01

MEAN:	31.61	32.93	56.59	25.39

COVARIANCE MATRIX:

5.24			
8.77	30.20		
-9.45	-22.41	34.97	
-6.30	-16.86	20.20	15.24

CLASS : WHEA
SUBCLASS: WH02

MEAN:	45.19	55.06	58.31	24.06

COVARIANCE MATRIX:

8.53			
7.55	17.06		
3.88	6.17	11.71	
1.99	1.81	2.73	2.18

CLASS : WHEA
SUBCLASS: WH03

MEAN:	36.47	44.94	48.65	20.35

COVARIANCE MATRIX:

6.13			
4.03	7.70		
3.17	2.75	9.76	
1.54	1.55	2.54	2.35

CLASS : WHEA
SUBCLASS: WH04

MEAN:	36.31	38.08	58.59	26.13

COVARIANCE MATRIX:

5.17			
7.17	29.65		
-3.28	-11.97	30.37	
-2.73	-8.21	14.97	10.57

CLASS : WHEA
SUBCLASS: WH05

MEAN:	46.16	54.16	70.37	29.84

COVARIANCE MATRIX:

57.82			
75.40	116.24		
29.94	40.31	28.23	
3.66	2.08	5.48	4.03

CLASS : WHEA
SUBCLASS: WH06

MEAN:	31.54	32.98	42.48	18.79

COVARIANCE MATRIX:

6.25			
7.65	17.33		
4.38	8.13	22.67	
1.56	2.67	7.14	3.78

ORIGINAL PAGE IS
OF POOR QUALITY

350

```
THE STATISTICS FILE FOR   1 CLASSES AND   6 SUBCLASSES HAS BEEN WRITTEN

THE STATS FOR A PARTICULAR CLASS OR SUBCLASS SHOULD BE REFERRED TO IN LATER RUNS BY
THE FOLLOWING NAMES AND NUMBERS (W HICHEVER APPLICABLE)

    CLASS  1    WHEA    SUBCLASSES (TOTAL=  6)

                         1    WH01
                         2    WH02
                         3    WH03
                         4    WH04
                         5    WH05
                         6    WH06
ON 11/13/79 AT 16: 6: 2 TIME FOR $ISO WAS - TOTAL     0.079 MINUTES VIRTUAL     0.047 MINUTES.
```

```
        PROCEDURE 1 RUN

$LABEL

            INPUT SUMMARY

DOTF        INPUT/UNIT=19,FILE=1
STAT        INPUT/UNIT=20,FILE=1
MAPF        INPUT/UNIT=16,FILE=1
MAPT        OUTPUT/UNIT=2,FILE=1
STAT        OUTPUT/UNIT=20,FILE=2
PROC        K-NEAREST
THRE        25.0
COMM        PROCEDURE 1 RUN
*END

USER HAS REQUESTED THE FOLLOWING OPTIONS :

CLUSTER/CLASSIFICATION TAPE IS BEING INPUT
MAPTAP FILE WILL BE OUTPUT
K-NEAREST PROCEDURE WILL BE USED
L1 DISTANCE WILL BE USED
THRESHOLD DISTANCE =      25.000
  1-NEAREST DOTS WILL BE USED
NO SUN ANGLE CORRECTION WILL BE APPLIED
DOTFIL FILE IS BEING INPUT
SAVTAP FILE IS BEING INPUT

INPUT IMAGE DATA TAPE INFORMATION

    FORMAT                  UNIVERSAL
    NO. OF CHANNELS             1
    NO. OF PIXELS/LINE         19
    FIRST SCAN LINE NO.         1
    FIRST PIXEL REFERENCE PT    1
```

357

LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

PROCEDURE 1 RUN

INPUT FIELDS

FIELD		CLASS		SUBCLASS	VERTICES (SAMPLE,LINE)
1	SS	1	SS		(120, 10)
2	SS	1	SS		(20, 30)
3	SS	1	SS		(180, 30)
4	SS	1	SS		(110, 40)
5	SS	1	SS		(130, 40)
6	SS	1	SS		(80, 50)
7	SS	1	SS		(100, 50)
8	SS	1	SS		(160, 50)
9	SS	1	SS		(70, 60)
10	SS	1	SS		(190, 60)
11	SS	1	SS		(20, 70)
12	SS	1	SS		(60, 70)
13	SS	1	SS		(80, 70)
14	SS	1	SS		(10, 80)
15	SS	1	SS		(30, 80)
16	SS	1	SS		(130, 80)
17	SS	1	SS		(40, 90)
18	SS	1	SS		(60, 90)
19	SS	1	SS		(80, 90)
20	SS	1	SS		(100, 90)
21	SS	1	SS		(120, 90)
22	SS	1	SS		(10, 100)
23	SS	1	SS		(130, 100)
24	N	2	N		(20, 10)
25	N	2	N		(40, 10)
26	N	2	N		(60, 10)
27	N	2	N		(80, 10)
28	N	2	N		(100, 10)
29	N	2	N		(140, 10)
30	N	2	N		(160, 10)
31	N	2	N		(180, 10)
32	N	2	N		(10, 20)
33	N	2	N		(30, 20)
34	N	2	N		(50, 20)
35	N	2	N		(70, 20)
36	N	2	N		(90, 20)
37	N	2	N		(110, 20)
38	N	2	N		(150, 20)
39	N	2	N		(170, 20)
40	N	2	N		(40, 30)
41	N	2	N		(60, 30)
42	N	2	N		(80, 30)
43	N	2	N		(100, 30)
44	N	2	N		(120, 30)
45	N	2	N		(140, 30)
46	N	2	N		(160, 30)
47	N	2	N		(10, 40)
48	N	2	N		(30, 40)
49	N	2	N		(50, 40)
50	N	2	N		(70, 40)

51	N	2	N	(90, 40)
52	N	2	N	(150, 40)
53	N	2	N	(170, 40)
54	N	2	N	(190, 40)
55	N	2	N	(20, 50)
56	N	2	N	(40, 50)
57	N	2	N	(60, 50)
58	N	2	N	(120, 50)
59	N	2	N	(140, 50)
60	N	2	N	(180, 50)
61	N	2	N	(10, 60)
62	N	2	N	(30, 60)
63	N	2	N	(50, 60)
64	N	2	N	(90, 60)
65	N	2	N	(110, 60)
66	N	2	N	(150, 60)
67	N	2	N	(170, 60)
68	N	2	N	(40, 70)
69	N	2	N	(160, 70)
70	N	2	N	(120, 70)
71	N	2	N	(160, 70)
72	N	2	N	(180, 70)
73	N	2	N	(50, 80)
74	N	2	N	(70, 80)
75	N	2	N	(110, 80)
76	N	2	N	(150, 80)
77	N	2	N	(170, 80)
78	N	2	N	(190, 80)
79	N	2	N	(20, 90)
80	N	2	N	(140, 90)
81	N	2	N	(160, 90)
82	N	2	N	(180, 90)
83	N	2	N	(30, 100)
84	N	2	N	(50, 100)
85	N	2	N	(70, 100)
86	N	2	N	(90, 100)
87	N	2	N	(110, 100)
88	N	2	N	(150, 100)
89	N	2	N	(170, 100)
90	N	2	N	(190, 100)
91	N	2	N	(20, 110)
92	N	2	N	(40, 110)
93	N	2	N	(60, 110)
94	N	2	N	(80, 110)
95	N	2	N	(140, 110)
96	N	2	N	(160, 110)
97	M	3	M	(90, 80)

555

351

LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

PROCEDURE 1 RUN

CLUSTER-DOT INTER-DISTANCE TABLE

DOTS	CLST (1)	CLST (2)	CLST (3)	CLST (4)	CLST (5)	CLST (6)
1.	10.43	37.63	19.41	14.11	55.53	19.21
2.	18.34	42.88	38.47	16.05	41.53	39.17
3.	33.43	56.63	24.41	33.11	74.53	12.75
4.	49.25	12.38	44.59	38.15	20.89	69.21
5.	36.43	27.63	9.88	29.33	45.53	29.21
6.	34.34	64.88	59.41	36.67	47.11	55.17
7.	53.25	16.38	48.59	42.15	28.89	73.21
8.	32.43	29.63	3.29	25.33	47.53	27.21
9.	34.43	59.63	27.41	36.11	77.53	13.75
10.	27.43	30.63	3.24	20.95	48.53	26.21
11.	43.25	6.38	38.59	32.15	20.89	63.21
12.	41.25	4.38	36.59	30.15	18.89	61.21
13.	19.43	54.63	22.41	31.11	72.53	2.21
14.	46.48	40.88	46.47	33.89	23.11	67.21
15.	3.25	35.50	27.41	10.11	51.53	23.21
16.	29.43	46.63	14.41	26.95	64.53	10.21
17.	28.43	23.63	9.29	21.33	41.53	33.21
18.	40.43	11.63	20.59	33.33	29.53	45.21
19.	25.43	32.63	5.12	18.95	50.53	24.21
20.	38.43	25.63	8.59	31.33	43.53	31.21
21.	29.43	36.63	4.41	22.95	54.53	20.21
22.	63.48	27.38	59.59	50.89	9.47	84.21
23.	42.48	22.88	38.59	29.89	11.84	63.21
24.	9.43	36.63	20.41	13.11	54.53	20.21
25.	14.43	27.63	15.41	11.95	45.53	29.21
26.	17.30	52.63	20.41	29.11	70.53	6.17
27.	10.48	32.88	27.41	4.67	43.53	31.21
28.	9.25	30.88	25.41	4.93	45.53	29.21
29.	21.30	56.63	24.41	33.11	74.53	2.17
30.	14.66	32.50	19.47	6.23	44.53	30.21
31.	13.43	30.50	17.47	6.49	46.53	28.21
32.	3.43	40.50	26.41	15.11	56.53	18.21
33.	31.48	21.88	27.59	18.89	22.53	52.21
34.	35.48	9.88	31.59	22.89	18.53	56.21
35.	26.57	61.88	56.41	33.67	56.84	47.25
36.	21.66	49.63	17.41	26.11	67.53	8.29
37.	13.48	28.50	25.47	6.23	40.53	34.21
38.	21.30	52.63	20.41	29.11	70.53	10.17
39.	18.25	21.88	27.47	7.31	36.53	38.21
40.	23.43	24.63	11.47	16.33	42.53	32.21
41.	46.48	16.88	42.59	33.89	7.53	67.21
42.	13.48	31.88	26.41	3.67	40.53	34.21
43.	28.48	34.88	29.41	16.51	25.53	49.21
44.	17.48	37.88	32.41	9.67	36.53	38.21
45.	31.43	62.63	30.41	39.11	80.53	10.75
46.	12.66	34.50	20.41	6.85	46.53	28.21
47.	17.48	37.88	32.41	9.67	36.53	38.21
48.	3.66	41.50	31.41	13.85	51.53	23.21
49.	2.30	39.50	29.41	14.11	55.53	21.17
50.	9.43	36.63	20.41	13.11	54.53	20.21

51.	22.43	45.63	13.41	22.11	63.53	11.21
52.	29.52	65.63	33.41	42.11	83.53	9.21
53.	28.43	41.63	9.41	25.95	59.53	15.21
54.	26.49	32.88	28.47	13.89	27.84	47.21
55.	24.43	23.63	8.59	17.33	41.53	33.21
56.	22.48	28.88	28.47	9.89	31.84	43.21
57.	22.52	58.63	26.41	35.11	76.53	8.83
58.	35.52	71.63	39.41	48.11	89.53	15.21
59.	37.48	23.88	33.59	24.89	18.84	58.21
60.	11.48	41.88	36.41	13.67	44.84	32.21
61.	24.48	30.88	34.47	12.05	29.53	45.21
62.	10.34	44.88	39.41	18.93	59.53	30.25
63.	13.43	38.63	16.41	15.11	56.53	18.21
64.	18.43	35.63	11.41	12.11	53.53	21.21
65.	24.43	51.63	19.41	28.11	69.53	5.21
66.	19.43	24.50	10.53	12.95	40.53	34.21
67.	27.43	20.63	12.53	20.95	38.53	3[illegible].21
68.	7.25	32.88	27.41	6.93	47.53	27.21
69.	25.43	16.63	15.59	18.33	34.53	40.21
70.	5.52	43.50	29.41	15.85	55.53	21.17
71.	19.43	54.63	22.41	31.11	72.53	2.21
72.	86.48	50.38	82.59	73.89	32.47	107.21
73.	28.43	51.63	19.41	28.11	69.53	8.17
74.	20.43	43.63	14.71	20.11	61.53	13.21
75.	19.43	42.63	15.71	19.11	60.53	14.21
76.	8.30	43.63	27.41	20.11	61.53	19.17
77.	25.43	42.63	10.41	22.95	60.53	14.21
78.	9.11	38.88	33.41	12.93	51.53	29.17
79.	28.52	64.63	32.41	41.11	82.53	7.79
80.	26.52	62.63	30.41	39.11	80.53	7.25
81.	32.48	48.88	43.41	20.67	31.11	53.21
82.	5.43	38.50	24.41	13.11	54.53	20.21
83.	48.57	83.88	78.41	55.67	66.11	69.25
84.	4.11	37.88	32.41	11.93	52.53	24.17
85.	32.52	68.63	36.41	45.11	86.53	11.79
86.	31.52	67.63	35.41	44.11	85.53	14.25
87.	3.43	40.50	26.41	15.11	56.53	18.21
88.	30.43	43.63	11.41	27.95	61.53	13.21
89.	21.52	57.63	25.41	34.11	75.53	10.25
90.	6.34	40.88	35.41	12.67	49.53	27.17
91.	23.43	28.63	7.47	16.33	46.53	28.21
92.	18.43	25.50	12.47	11.33	41.53	33.21
93.	17.43	28.63	12.41	14.95	46.53	28.21
94.	5.43	38.50	24.41	13.11	54.53	20.21
95.	23.48	35.88	35.47	13.05	30.84	44.21
96.	22.52	58.63	26.41	35.11	76.53	11.25
97.	34.43	11.63	20.59	27.33	29.53	45.21

356

LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

PROCEDURE 1 RUN

LABELING BY 1-NEAREST NEIGHBOR PROCEDURE

CLUSTER LABELING DETAILS

CLUSTER NUMBER	CLUSTER LABFL	DOT LABEL	DOT NUMBER	DOT DISTANCE	DOT CLUSTER
1	N	N	49	2.30	1
2	SS	SS	12	4.38	2
3	SS	SS	10	3.24	3
4	N	N	42	3.67	4
5	N	N	41	7.53	5
6	N	N	29	2.17	6

CLUSTER LABELING SUMMARY

CLUSTER NUMBER	LABEL	TOTAL	NUMBER OF DOTS USED (BY CATEGORY NAME) SS	N	M
1	N	1	0	1	0
2	SS	1	1	0	0
3	SS	1	1	0	0
4	N	1	0	1	0
5	N	1	0	1	0
6	N	1	0	1	0

```
THE STATISTICS FILE FOR    2 CLASSES AND    6 SUBCLASSES HAS BEEN WRITTEN

THE STATS FOR A PARTICULAR CLASS OR SUBCLASS SHOULD BE REFERRED TO IN LATER RUNS BY
THE FOLLOWING NAMES AND NUMBERS (W HICHEVER APPLICABLE)

    CLASS  1    SS        SUBCLASSES (TOTAL=  2)

                           1    SS01
                           2    SS02

    CLASS  2    N         SUBCLASSES (TOTAL=  4)

                           3    N 01
                           4    N 02
                           5    N 03
                           6    N 04
ON 11/13/79 AT 16: 6: 8 TIME FOR $LAB WAS - TOTAL       0.043 MINUTES VIRTUAL      0.029 MINUTES.
```

PROCEDURE 1 RUN

```
$CLASSIF
   MAPT        OUTPUT/UNIT=2,FILE=2
   DATA         UNIT=11,FILE=2
   STAT         UNIT=20,FILE=2
   CATE         FILE
   APRI         FILE
   COMM         PROCEDURE 1 RUN
   *END
```

LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

PROCEDURE 1 RUN

THE FOLLOWING OPTIONS HAVE BEEN SELECTED

```
   CATEGORY CLASSIFIER OPTION HAS BEEN SELECTED.
     ALSO CLASSES FROM STATFILE WILL BE CONSIDERED THE CATEGORIES FOR CLASSIFICATION
   APRIORI VALUES FROM STATFILE  2   APRIORI= NO. PI   XELS IN SUBCLASS/TOTAL NO. PIXELS IN ALL SUBCLASSES ***
```

SUPERVISOR INFORMATION :

```
   FILE NUMBER ....        2
   NO. OF FIELDS.......    1
   NO. OF CLASSES.........  2
   NO. OF SUBCLASSES.....  6
   NO. OF CHANNELS.......  4
```

LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

PROCEDURE 1 RUN

AREA USED TO COMPUTE TRAINING STATISTICS

FIFLD	CLASS	SUBCLASS	VERTICES (SAMPLE,LINE)
1 SFGM	1 SS		(10, 10) (196, 10) (196, 117) (10, 117)

362

LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

PROCEDURE 1 RUN

*** CLASSIFICATION STUDY *** MAPTAP FILE 2

SUBCLASSES CONSIDERED

SYMBOL	SUBCLASS	A PRIOR
1	SS01	0.0766
2	SS02	0.0813
3	N 01	0.2105
4	N 02	0.2919
5	N 03	0.0909
6	N 04	0.2488

CHANNELS CONSIDERED

TRAINING	RECOGNITION
1	1
2	2
3	3
4	4

SKIPPING FILE 1

INPUT IMAGE DATA TAPE INFORMATION

FORMAT	UNIVERSAL
NO. OF CHANNELS	16
NO. OF PIXELS/LINE	196
FIRST SCAN LINE NO.	1
FIRST PIXEL REFERENCE PT	1

LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

PROCEDURE 1 RUN

AREA OF CLASSIFICATION

FIELD NAME	NO. OF VERTICES	SAMPLE INC.	LINE INC.	VERTICES
SFGM	4	10	10	(10, 10) (196, 10) (196, 117) (10, 117) (10, 10)

LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

PROCEDURE 1 RUN

MAP OF CATEGORY CLASSIFIER CLASSIFICATION RESULTS

CATEGORY NO.	CATEGORY NAME	CLASS NO.	CLASS NAME	SUBCLASS NO.	SUBCLASS NAME	SUBCLASS SYMBOL
1	SS	1	SS	1	SS01	1
				2	SS02	2
2	N	2	N	3	N 01	3
				4	N 02	4
				5	N 03	5
				6	N 04	6

```
            0000000001111111111
            1234567890123456789
            0000000000000000000

     10     3353463434333664443
     20     3656143364413363446
     30     3424454445643644466
     40     4334313466122362234
     50     5234465441666562246
     60     4434326443641644232
     70     4114511634436346456
     80     5646656615656634644
     90     3662616212324644433
    100     5535366663315224643
    110     3444134323413436516
```

*** $CLASSIFY - COMPLETED ***

ON 11/13/79 AT 16: 6:36 TIME FOR $CLA WAS - TOTAL 0.068 MINUTES VIRTUAL 0.041 MINUTES.

LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

```
        PROCEDURE 1 RUN

$DISPLAY
     GTUN      UNIT=23,FILE=1
     PPUN      UNIT=27,FILE=1
     AIUN      UNIT=28,FILE=1
     *END


   YOU HAVE SELFCTED THE FOLLOWING OPTIONS:
       PROCESS THE CLASSIFICATION RESULTS FROM MAPTAP (      B ), UNIT    2  ,  FILE    1
       APPLY NO THRESHOLDING

    NUMBERS OF GROUND TRUTH ,  AI, AND          DISCRIMINATOR UNITS AND FILES ARE AS FOLLOWS

     23      1      28      1      27      1
```

LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

PROCEDURE 1 RUN

AREA USED TO COMPUTE TRAINING STATISTICS

FIELD		CLASS		SUBCLASS	VERTICES (SAMPLE,LINE)
1	SS	1	SS		(120, 10)
2	SS	1	SS		(20, 30)
3	SS	1	SS		(180, 30)
4	SS	1	SS		(110, 40)
5	SS	1	SS		(130, 40)
6	SS	1	SS		(80, 50)
7	SS	1	SS		(100, 50)
8	SS	1	SS		(160, 50)
9	SS	1	SS		(70, 60)
10	SS	1	SS		(190, 60)
11	SS	1	SS		(20, 70)
12	SS	1	SS		(60, 70)
13	SS	1	SS		(80, 70)
14	SS	1	SS		(10, 80)
15	SS	1	SS		(30, 80)
16	SS	1	SS		(130, 80)
17	SS	1	SS		(40, 90)
18	SS	1	SS		(60, 90)
19	SS	1	SS		(80, 90)
20	SS	1	SS		(100, 90)
21	SS	1	SS		(120, 90)
22	SS	1	SS		(10, 100)
23	SS	1	SS		(130, 100)
24	N	2	N		(20, 10)
25	N	2	N		(40, 10)
26	N	2	N		(60, 10)
27	N	2	N		(80, 10)
28	N	2	N		(100, 10)
29	N	2	N		(140, 10)
30	N	2	N		(160, 10)
31	N	2	N		(180, 10)
32	N	2	N		(10, 20)
33	N	2	N		(30, 20)
34	N	2	N		(50, 20)
35	N	2	N		(70, 20)
36	N	2	N		(90, 20)
37	N	2	N		(110, 20)
38	N	2	N		(150, 20)
39	N	2	N		(170, 20)
40	N	2	N		(40, 30)
41	N	2	N		(60, 30)
42	N	2	N		(80, 30)
43	N	2	N		(100, 30)
44	N	2	N		(120, 30)
45	N	2	N		(140, 30)
46	N	2	N		(160, 30)
47	N	2	N		(10, 40)
48	N	2	N		(30, 40)
49	N	2	N		(50, 40)
50	N	2	N		(70, 40)

51	N	2	N	(90, 40)
52	N	2	N	(150, 40)
53	N	2	N	(170, 40)
54	N	2	N	(190, 40)
55	N	2	N	(20, 50)
56	N	2	N	(40, 50)
57	N	2	N	(60, 50)
58	N	2	N	(120, 50)
59	N	2	N	(140, 50)
60	N	2	N	(180, 50)
61	N	2	N	(10, 60)
62	N	2	N	(30, 60)
63	N	2	N	(50, 60)
64	N	2	N	(90, 60)
65	N	2	N	(110, 60)
66	N	2	N	(150, 60)
67	N	2	N	(170, 60)
68	N	2	N	(40, 70)
69	N	2	N	(100, 70)
70	N	2	N	(120, 70)
71	N	2	N	(160, 70)
72	N	2	N	(180, 70)
73	N	2	N	(50, 80)
74	N	2	N	(70, 80)
75	N	2	N	(110, 80)
76	N	2	N	(150, 80)
77	N	2	N	(170, 80)
78	N	2	N	(190, 80)
79	N	2	N	(20, 90)
80	N	2	N	(140, 90)
81	N	2	N	(160, 90)
82	N	2	N	(180, 90)
83	N	2	N	(30, 100)
84	N	2	N	(50, 100)
85	N	2	N	(70, 100)
86	N	2	N	(90, 100)
87	N	2	N	(110, 100)
88	N	2	N	(150, 100)
89	N	2	N	(170, 100)
90	N	2	N	(190, 100)
91	N	2	N	(20, 110)
92	N	2	N	(40, 110)
93	N	2	N	(60, 110)
94	N	2	N	(80, 110)
95	N	2	N	(140, 110)
96	N	2	N	(160, 110)
97	M	3		(90, 80)

LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

PROCEDURE 1 RUN

```
DISPLAY OF CLASSIFIED FIELD...........SEGM
           CLASSIFICATION  DATE.......
           CLASSIFICATION CHANNELS... 2 LINEAR COMBINATIONS OF CHANNELS
                                       1  2  3
```

MAP OF CATEGORY CLASSIFIER CLASSIFICATION RESULTS

CATEGORY NO.	CATEGORY NAME	CLASS NO.	CLASS NAME	SUBCLASS NO.	SUBCLASS NAME	SYMBOL
1	SS	1	SS	1	SS01	1
				2	SS02	2
2	N	2	N	3	N 01	3
				4	N 02	4
				5	N 03	5
				6	N 04	6

```
      0000000001111111111
      1234567890123456789
      0000000000000000000

 10   3353463434333664443
 20   3656143366413363446
 30   3444454445644644466
 40   4334313466122362234
 50   5434565441666562246
 60   4434346343641644232
 70   4114511644336346456
 80   5646656615656632644
 90   3662616212324644443
100   5535366663315264623
110   3244134426443436516
```


368

LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

PROCEDURE 1 RUN

CLASSIFICATION SUMMARY FOR FIELD SEGM

TOTAL NUMBER OF SAMPLED POINTS 209

SUBCLASS	PTS. BEFORE THRES.	PCT. OF TOTAL CLSF.FLD.	PTS. AFTER THRES.	PCT.OF TOTAL CLSF.FLD.	PCT.OF SUBCLASS	PTS. THRES.	PCT.OF TOTAL CLSF.FLD.	PCT. OF SUBCLASS THRES.
SS01	16	7.66	16	7.66	100.00	0	0.0	0.0
SS02	17	8.13	17	8.13	100.00	0	0.0	0.0
N 01	44	21.05	44	21.05	100.00	0	0.0	0.0
N 02	61	29.19	61	29.19	100.00	0	0.0	0.0
N 03	19	9.09	19	9.09	100.00	0	0.0	0.0
N 04	52	24.88	52	24.88	100.00	0	0.0	0.0

PTS. THRESHOLDED IN DISPLAY 0
PTS. THRESHOLDED IN CLASSIFY 0
TOTAL 0 PCT. = 0.0

LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

PROCEDURE 1 RUN

CLASSIFICATION SUMMARY FOR FIELD SEGM

TOTAL NUMBER OF SAMPLED POINTS 209

CLASS	PTS. BEFORE THRES.	PCT. OF TOTAL CLSF.FLD.	PTS. AFTER THRES.	PCT.OF TOTAL CLSF.FLD.	PCT.OF CLASS	PTS. THRES.	PCT.OF TOTAL CLSF.FLD.	PCT. OF CLASS THRES.
SS	33	15.79	33	15.79	100.00	0	0.0	0.0
N	176	84.21	176	84.21	100.00	0	0.0	0.0

PTS. THRESHOLDED IN DISPLAY 0
PTS. THRESHOLDED IN CLASSIFY 0
TOTAL 0 PCT. = 0.0

376

LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

PROCEDURE 1 RUN

CLASSIFICATION SUMMARY FOR FIFLD SEGM

TOTAL NUMBER OF SAMPLED POINTS 209

CATEGORY	PTS. BEFORE THRES.	PCT. OF TOTAL CLSF.FLD.	PTS. AFTER THRES.	PCT.OF TOTAL CLSF.FLD.	PCT.OF CATEGORY	PTS. THRES.	PCT.OF TOTAL CLSF.FLD.	PCT. OF CATEGORY THRES.
SS	33	15.79	33	15.79	100.00	0	0.0	0.0
N	176	84.21	176	84.21	1^0.00	0	0.0	0.0

PTS. THRESHOLDED IN DISPLAY 0
PTS. THRESHOLDED IN CLASSIFY 0
TOTAL 0 PCT. = 0.0

SITE = 1851 DAY= 1 MONTH= 4 YEAR= 77 MASK = TY TYPE = GT

PPTC VS CLASSIFIED LABELS

TYPE 1 DOT CLASSIFICATION

	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190
10		N /N N 01		N /N N 01		N /N N 04		N /N N 02		N /N N 02		N /N N 01		N /N N 04		N /N N 02		N /N N 02	
20	N /N N 01		N /N N 03		N /SS SS01		N /N N 01		N /N N 04		N /N N 02				N /N N 04		N /N N 02		
30		N /N N 02		N /N N 02		N /N N 03		N /N N 02		N /N N 03		N /N N 02		N /N N 04		N /N N 02		N /N N 04	
40	N /N N 02		N /N N 01		N /N N 01		N /N N 01		N /N N 04		N /SS SS01		N /SS SS02		N /N N 04		N /SS SS02		N /N N 02
50		N /N N 02		N /N N 02		N /N N 04		N /N N 02		N /SS SS01		N /N N 04		N /N N 03		N /SS SS02		N /N N 02	
60	N /N N 02		N /N N 01		N /N N 01		N /N N 04		N /N N 02		N /N N 04				N /N N 02		N /SS SS02		N /SS SS02
70		N /SS SS01		N /N N 02		N /SS SS01		N /N N 04		N /N N 02		N /N N 01				N /N N 04		N /N N 03	
80	N /N N 03		N /N N 02		N /N N 04		N /N N 04		S /SS SS01		N /N N 04		N /N N 04		N /N N 01		N /N N 04		N /N N 02
90		N /N N 04		N /SS SS02		N /SS SS01		N /SS SS02		N /SS SS02		N /SS SS02		N /N N 04		N /N N 02		N /N N 02	
100	N /N N 03		N /N N 01		N /N N 01		N /N N 04		N /N N 04		N /N N 01		N /N N 03		N /N N 04		N /N N 04		N /N N 01
110		N /SS SS02		N /N N 02		N /N N 01		N /N N 02						N /N N 02		N /N N 04			

PPTC VS CLASSIFIED LABELS

TYPE 2 DOT CLASSIFICATION

	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190
10	N /N N 01		N /N N 03		N /N N 02		N /N N 01		N /N N 01		N /N N 01		N /N N 01		N /N N 04		N /N N 02		N /N N 01
20		N /N N 04		N /N N 04		N /N N 02		N /N N 01		N /N N 04		M /SS SS01	M /N N 01	N /N N 01		N /N N 01		N /N N 02	M /N N 04
30	N /N N 01		N /N N 02		N /N N 02		N /N N 02		N /N N 02		N /N N 04		N /N N 02		N /N N 02		N /N N 02		N /N N 04
40		M /N N 01		N /N N 02		S /SS SS01		N /N N 02		N /N N 04		S /SS SS02		M /N N 01		N /SS SS02		N /N N 01	
50	N /N N 03		N /N N 01		N /N N 03		S /N N 03		N /N N 02		N /N N 04		M /N N 04		N /N N 04		S /SS SS02		N /N N 04
60		N /N N 02		N /N N 02		S /N N 02		N /N N 01		N /N N 01		N /N N 02	M /SS SS01	S /N N 04		N /N N 02		N /N N 01	
70	N /N N 02		S /SS SS01		N /N N 03		S /SS SS01		N /N N 02		N /N N 01		N /N N 04	M /N N 01	S /N N 02		N /N N 02		M /N N 04
80		N /N N 04		N /N N 04		N /N N 03		N /N N 04		N /N N 03		S /N N 03		N /N N 04		N /SS SS02		N /N N 02	
90	N /N N 01		N /N N 04		N /N N 04		N /N N 04		S /SS SS01		N /N N 01		N /N N 02		N /N N 02		N /N N 02		N /N N 01
100		N /N N 03		S /N N 03		N /N N 04		N /N N 04		N /N N 01		S /SS SS01		N /SS SS02		N /N N 02		S /SS SS02	
110	N /N N 01		N /N N 02		N /SS SS01		N /N N 02		S /SS SS02	M /N N 04	N /N N 02	M /N N 02	N /N N 01		N /N N 01		M /N N 03	M /SS SS01	N /N N 04

TYPE II DOT REPORTS

CONFUSION MATRIX

CLASS	S	N	M
PPC LABEL			
S	0	6	0
N	0	80	0
M	0	10	0

ALPHA VALUE MATRIX

CLASS	S	N	M
PPC LABEL			
S	0.0	0.063	0.0
N	0.0	0.833	0.0
M	0.0	0.104	0.0

DOT DATA PERFORMANCE SUMMARY

CATEGORY	CLASSIFIED	BIAS CORRECTED PROPORTION		RANDOM SAMPLE PROPORTION		VARIANCE
NAME	ESTIMATE	ESTIMATE	VARIANCE	ESTIMATE	VARIANCE	RATIO
S	0.0	5.2632	4.3738	13.3929	10.4497	0.4186
N	84.2105	70.1754	10.3676	75.0000	16.8919	0.6138
M	0.0	8.7719	6.9657	11.6071	9.2431	0.7536

SITE = 1851 DAY= 1 MONTH= 4 YEAR= 77 MASK = TY TYPE = GT

GROUND TRUTH VS CLASSIFIED LABELS

TYPE 1 DOT CLASSIFICATION

	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190
10		N /N N 01		N /N N 01		N /N N 04		N /N N 02		N /N N 02		SS/N N 01		N /N N 04		N /N N 02		N /N N 02	
20	N /N N 01		N /N N 03		N /SS SS01		N /N N 01		N /N N 04		N /N N 02				N /N N 04		N /N N 02		
30		SS/N N 02		N /N N 02		N /N N 03		N /N N 02		N /N N 03		N /N N 02		N /N N 04		N /N N 02		SS/N N 04	
40	N /N N 02		N /N N 01		N /N N 01		N /N N 01		N /N N 04		SS/SS SS01		SS/SS SS02		N /N N 04		N /SS SS02		N /N N 02
50		N /N N 02		N /N N 02		N /N N 04		SS/N N 02		SS/SS SS01		N /N N 04		N /N N 03		SS/SS SS02		N /N N 02	
60	N /N N 02		N /N N 01		N /N N 01		SS/N N 04		N /N N 02		N /N N 04				N /N N 02		N /SS SS02		SS/SS SS02
70		SS/SS SS01		N /N N 02		SS/SS SS01		SS/N N 04		N /N N 02		N /N N 01				N /N N 04		N /N N 03	
80	SS/N N 03		SS/N N 02		N /N N 04		N /N N 04		SS/SS SS01		N /N N 04		SS/N N 04		N /N N 01		N /N N 04		N /N N 02
90		N /N N 04		SS/SS SS02		SS/SS SS01		SS/SS SS02		SS/SS SS02		SS/SS SS02		N /N N 04		N /N N 02		N /N N 02	
100	SS/N N 03		N /N N 01		N /N N 01		N /N N 04		N /N N 04		N /N N 01		SS/N N 03		N /N N 04		N /N N 04		N /N N 01
110		N /SS SS02		N /N N 02		N /N N 01		N /N N 02						N /N N 02		N /N N 04			

GROUND TRUTH VS CLASSIFIED LABELS

TYPE 2 DOT CLASSIFICATION

	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190
10	N /N N 01		N /N N 03		N /N N 02		N /N N 01		N /N N 01		N /N N 01		N /N N 01		N /N N 04		N /N N 02		N /N N 01
20		N /N N 04		N /N N 04		N /N N 02		N /N N 01		N /N N 04		M /SS SS01	M /N N 01	N /N N 01		N /N N 01		N /N N 02	M /N N 04
30	N /N N 01		N /N N 02		N /N N 02		N /N N 02		N /N N 02		N /N N 04		N /N N 02		N /N N 02		N /N N 02		N /N N 04
40		M /N N 01		N /N N 02		SS/SS SS01		N /N N 02		N /N N 04		SS/SS SS02		M /N N 01		N /SS SS02		N /N N 01	
50	N /N N 03		N /N N 01		N /N N 03		SS/N N 03		N /N N 02		N /N N 04		M /N N 04		N /N N 04		SS/SS SS02		N /N N 04
60		N /N N 02		N /N N 02		SS/N N 02		N /N N 01		N /N N 01		N /N N 02	M /SS SS01	SS/N N 04		N /N N 02		N /N N 01	
70	N /N N 02		SS/SS SS01		N /N N 03		SS/SS SS01		N /N N 02		N /N N 01		N /N N 04	M /N N 01	SS/N N 02		N /N N 02		M /N N 04
80		N /N N 04		N /N N 04		N /N N 03		N /N N 04		N /N N 03		SS/N N 03		N /N N 04		N /SS SS02		N /N N 02	
90	N /N N 01		N /N N 04		N /N N 04		N /N N 04		SS/SS SS01		N /N N 01		N /N N 02		N /N N 02		N /N N 02		N /N N 01
100		N /N N 03		SS/N N 03		N /N N 04		N /N N 04		N /N N 01		SS/SS SS01		N /SS SS02		N /N N 02		SS/SS SS02	
110	N /N N 01		N /N N 02		N /SS SS01		N /N N 02		SS/SS SS02	M /N N 04	N /N N 02	M /N N 02	N /N N 01		N /N N 01		M /N N 03	M /SS SS01	N /N N 04

TYPE II DOT REPORTS

CONFUSION MATRIX

CLASS	SS	N	M
GT LABEL			
SS	9	6	0
N	4	80	0
M	3	10	0

ALPHA VALUE MATRIX

CLASS	SS	N	M
GT LABEL			
SS	0.563	0.063	0.0
N	0.250	0.833	0.0
M	0.188	0.104	0.0

DOT DATA PERFORMANCE SUMMARY

CATEGORY NAME	CLASSIFIED ESTIMATE	BIAS CORRECTED PROPORTION ESTIMATE	BIAS CORRECTED PROPORTION VARIANCE	RANDOM SAMPLE PROPORTION ESTIMATE	RANDOM SAMPLE PROPORTION VARIANCE	VARIANCE RATIO
SS	15.7895	14.1447	8.4640	13.3929	10.4497	0.8100
N	84.2105	74.1228	13.4839	75.0000	16.8919	0.7982
M	0.0	11.7324	9.4977	11.6071	9.2431	1.0275

SITE = .1851 DAY=   1 MONTH=   4 YEAR=   77 MASK = TY TYPE = GT

A. I. LABELS VS CLASSIFIED LABELS

TYPE 1 DOT CLASSIFICATION

	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190
10		N /N N 01		N /N N 01		N /N N 04		N /N N 02		N /N N 02		SS/N N 01		N /N N 04		N /N N 02		N /N N 02	
20	N /N N 01		N /N N 03		N /SS SS01		N /N N 01		N /N N 04		N /N N 02				N /N N 04		N /N N 02		
30		SS/N N 02		N /N N 02		N /N N 03		N /N N 02		N /N N 03		N /N N 02		N /N N 04		N /N N 02		SS/N N 04	
40	N /N N 02		N /N N 01		N /N N 01		N /N N 01		N /N N 04		SS/SS SS01		SS/SS SS02		N /N N 04		N /SS SS02		N /N N 02
50		N /N N 02		N /N N 02		N /N N 04		SS/N N 02		SS/SS SS01		N /N N 04		N /N N 03		SS/SS SS02		N /N N 02	
60	N /N N 02		N /N N 01		N /N N 01		SS/N N 04		N /N N 02		N /N N 04				N /N N 02		N /SS SS02		SS/SS SS02
70		SS/SS SS01		N /N N 02		SS/SS SS01		SS/N N 04		N /N N 02		N /N N 01				N /N N 04		N /N N 03	
80	SS/N N 03		SS/N N 02		N /N N 04		N /N N 04		N /SS SS01		N /N N 04		SS/N N 04		N /N N 01		N /N N 04		N /N N 02
90		N /N N 04		SS/SS SS02		SS/SS SS01		SS/SS SS02		SS/SS SS02		SS/SS SS02		N /N N 04		N /N N 02		N /N N 02	
100	SS/N N 03		N /N N 01		N /N N 01		N /N N 04		N /N N 04		N /N N 01		SS/N N 03		N /N N 04		N /N N 04		N /N N 01
110		N /SS SS02		N /N N 02		N /N N 01		N /N N 02						N /N N 02		N /N N 04			

A. I. LABELS VS CLASSIFIED LABELS

TYPE 2 DOT CLASSIFICATION

	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190
10	N /N N 01		N /N N 03		N /N N 02		N /N N 01		N /N N 01		N /N N 01		N /N N 01		N /N N 04		N /N N 02		N /N N 01
20		N /N N 04		N /N N 04		N /N N 02		N /N N 01		N /N N 04		M /SS SS01	M /N N 01	N /N N 01		N /N N 01		N /N N 02	M /N N 04
30	N /N N 01		N /N N 02		N /N N 02		N /N N 02		N /N N 02		N /N N 04		N /N N 02		N /N N 02		N /N N 02		N /N N 04
40		M /N N 01		N /N N 02		SS/SS SS01		N /N N 02		N /N N 04		SS/SS SS02		M /N N 01		N /SS SS02		N /N N 01	
50	N /N N 03		N /N N 01		N /N N 03		SS/N N 03		N /N N 02		N /N N 04		M /N N 04		N /N N 04		SS/SS SS02		N /N N 04
60		N /N N 02		N /N N 02		SS/N N 02		N /N N 01		N /N N 01		N /N N 02	M /SS SS01	SS/N N 04		N /N N 02		N /N N 01	
70	N /N N 02		SS/SS SS01		N /N N 03		SS/SS SS01		N /N N 02		N /N N 01		N /N N 04	M /N N 01	SS/N N 02		N /N N 02		M /N N 04
80		N /N N 04		N /N N 04		N /N N 03		N /N N 04		N /N N 03		SS/N N 03		N /N N 04		N /SS SS02		N /N N 02	
90	N /N N 01		N /N N 04		N /N N 04		N /N N 04		SS/SS SS01		N /N N 01		N /N N 02		N /N N 02		N /N N 02		N /N N 01
100		N /N N 03		SS/N N 03		N /N N 04		N /N N 04		N /N N 01		SS/SS SS01		N /SS SS02		N /N N 02		SS/SS SS02	
110	N /N N 01		N /N N 02		N /SS SS01		N /N N 02		SS/SS SS02	M /N N 04	N /N N 02	M /N N 02	N /N N 01		N /N N 01		M /N N 03	M /SS SS01	N /N N 04

TYPE II DOT REPORTS

CONFUSION MATRIX

CLASS	SS	N	M
AI LABEL			
SS	9	6	0
N	4	80	0
M	3	10	0

ALPHA VALUE MATRIX

CLASS	SS	N	M
AI LABEL			
SS	0.563	0.063	0.0
N	0.250	0.833	0.0
M	0.188	0.104	0.0

DOT DATA PERFORMANCE SUMMARY

CATEGORY	CLASSIFIED	BIAS CORRECTED PROPORTION		RANDOM SAMPLE PROPORTION		VARIANCE
NAME	ESTIMATE	ESTIMATE	VARIANCE	ESTIMATE	VARIANCE	RATIO
SS	15.7895	14.1447	8.4640	13.3929	10.4497	0.8100
N	84.2105	74.1228	13.4933	75.0000	16.8919	0.7982
M	0.0	11.7324	9.4977	11.6071	9.2431	1.0275

APPENDIX K

LACIE-FORMATTED DOT FILE FORMAT

APPENDIX K

LACIE-FORMATTED DOT FILE FORMAT

The LACIE-formatted dot file card images appear as follows:

Columns	Parameter	Description
1-3	DOT	The three letters DOT
4	Blank	
5	{1} {2}	Dot type
6-7	Blank	
8-9	Category name (left justified)	(The EOD-LARSYS allows two-character category names. If a one-character name is used, it should appear in column 8).
10	Blank	
11-80	$n_1, n_2, \cdots, n_N$	Dot grid numbers (integers) separated by blanks

The LACIE dot grid numbers correspond to sample and line numbers as follows:

Dot grid number	Sample	Line
1	10	10
2	20	10
⋮		
19	190	10
20	10	20
⋮		
39	10	30
⋮		
209	190	110

See figure K-1.

There is no continuation character from one card to the next; i.e., each card must have DOT in columns 1 through 3.

Limited provisions have been made for specifying off-grid dots. The dot number

$$LI \times 100000000 + SI \times 10000 + \text{LACIE number}$$

corresponds to the LACIE grid position offset by LI lines up and SI samples to the right (increasing sample numbers). This method of encoding does not correspond to the latest method used in LACIE operations and a reconciliation may be made in the future.

This format is used for EOD-LARSYS files GTUNIT, AIUNIT, and PPUNIT.

356

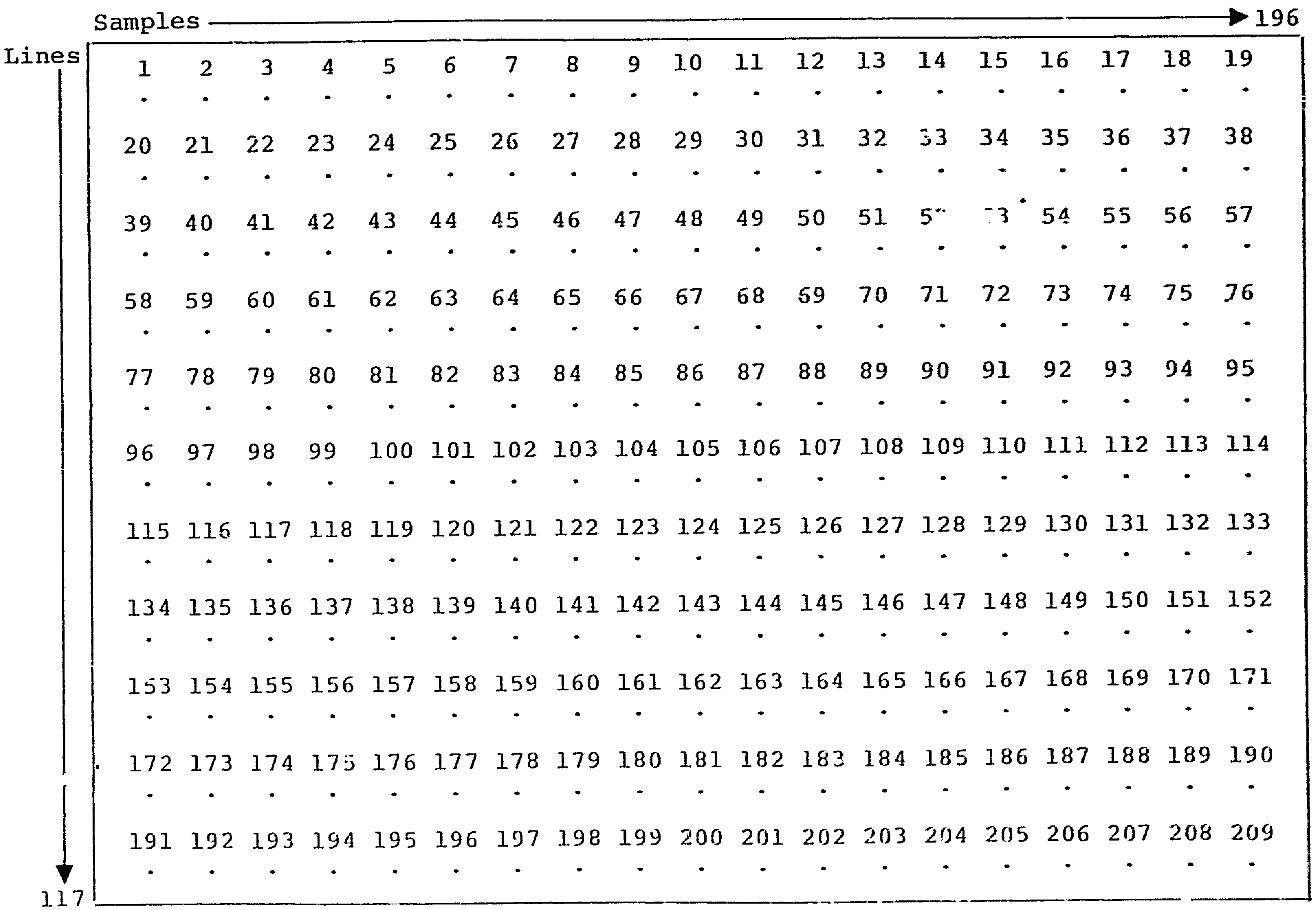


Figure K-1.– Dots in LACIE Procedure 1.