# Final Technical Report 

NASA Contract NAS9-14970 June 1, 1976-May 31, 1977

# D.A. Landgrebe, Purdue University <br> Principal Investigator <br> J.D.Erickson, NASA/JSC <br> Technical Monitor 

Volume I of III

Submitted by
The Laboratory for Applications of Remote Sensing Purdue University West Lafayette, Indiana

 semination of Eath Resoureas surver prosian information and witholu hablly


NAS9-14970
T-1314/4
MA-129TA
沰 DRA NasA-CR-151454. \% $\%$ 7 7 - $0: 18$. 黑

## Final Technical Report



NASA Contract NAS9-14970 June 1, 1976 - May 31, 1977

## D. A. Landgrebe, Purdue University Principal Investigator $\because \square$ J.D.Erickson, NASA/JSC Technical Monitor



# PURDUE <br> UNIVERSTY ${ }_{\text {Lumomiony roan apuacrions }}$ OF REMOTE SENSING 

June 28, 1977

National Aeronautics and Space
Administration
Scientific and Technical
Information Facility, Cōde KS
Washington, D.C. 20546

## Dear Sirs:

The report listed below is submitted in accord with Contract NAS9-14970, Article XVI - Reports of Work and Documentation.
"Final Technical Report", Volumes I, II, and III, NASA Contract NAS9-14970, June 1, 1976 - May 31,1977 by the Laboratory for Applications of Remote Sensing.

We have also provided a space below for you to indicate that this document has been placed in the public domain in accord with Article XX - Limitations on Earth Resources Data.
. Very truly yours,
A. A. Aram Landgrebe

- Director, LARS

DAL: jh
Enclosures 1 copy of the 3 volumes

This is to certify the above document has been placed in the public domain.

Signed $\qquad$
Title $\qquad$
Table of Contents
Preface ..... ii
Table I ..... iii
RESEARCH TASKS
Volume I
2.1 Test of Boundary Finding/Per Field Classification . . 2.1-1
Volume II
2.2 Stratification of Scene Characteristics . . . . . . . 2.2-1
2.2a Stratification by Machine Clustering. . . . . . 2.2a-1
2.2b Digitization and Registration Ancillary Data. . 2.2b-1
2.2c Crop Inventory Using Full-Frame Classification. . . . . . . . . . . . . . . . 2.2c-1
Volume III
2.3 LACIE Field Measurements. ..... 2.3-1
2.4 Scanner System Parameter Selection ..... 2,4-1
2.5 Technology Transfer ..... 2.5-1
2.6 Large Area Crop Inventory Design. ..... 2.6-1
2.7 Forestry Applications Project ..... 2.7-1
2.8 Regional Applications Project ..... 2,8-1
2.9 Interpretation of Thermal Band Data ..... 2.9-1
2.10 Super Site Data Management ..... 2.10-1
2.11 Soil Classification and Survey ..... 2.11-1
2.12 Improved Analysis Techniques for Multitemporal Data ..... 2.12-1

This report provides a summary of results for the past year's effort under contract NAS9-14970. The contract called for work on a wide variety of separate and distinct but related tasks. Below is a list of the tasks as contained in the original Work Statement.

As a result of this contract a large volume of results has been generated. Technical and research reports previously submitted or presently being published are listed in Table I.

Because of the diversity present in the task list for the contract, each major subdivision of this report has been written to be relatively self-contained. We hope this will facilıtate use of the report by readers with different interests.

Many of the tasks below are continulng; as a result, the discussion presented relative to them constitutes a progress report. Task 2.1 and 2.2 are not continuing, thus the discussion relative to them is more complete.

The various tasks have been managed by various Purdue staff members during the year. It is appropriate that the contributions of these people be recognized.

## Task <br> Purdue Task Manager

2.1 Test of Boundary Finding/Per Field

Classification (ECHO) P.H. Swain
2.2 Stratification of Scene Characteristics P.H. Swain
2.2a Stratification by Machine Clustering
2.2b Digitization and Registration of Ancillary Data
2.2c Crop Inventory Using Full-Frame Classification
2.3 LACIE Field Measurements
M.E. Bauer
2.4 Scanner System Parameter Selection
2.5 Technology Transfer
2.6 Large Area Crop Inventory Design
2.7 Forestry Applications Project
P.A. Anuta
2.8 Regional Applications Project
J.C. Lindenlaub
M.E. Bauer
2.9 Interpretation of Thermal Band Data
R.P. Mroczynski
R.A. Weismiller
2.10 Super Site Data Management
2.11 Soil Classification and Survey
L.F. Silva
2.12 Improved Analysis Techniques for Multitemporal Data
L.F. Silva
M.F. Baumgardner
-
P.H. Swain

The efforts of Dr. J.D. Erickson, the contract Technical Monitor and Mr. M.C. Trichel are especially to be noted and greatfully acknowledged.

Table I. Technical and Research Reports

052977 Davis. The Focus Series: A Collection of Single-Concept Remote Sensing Educational Materials.

042777 Russell and Lindenlaub. Disseminating Technological Information on Remote Sensing to Potential Users.

120776 Vanderbilt, Silva and Bauer. A Laser Technique for Characterizing the Geometry of Plant Canopies.

091576 DeWitt and Robinson. Description and Evaluation of a Bidirectional Reflectance Factor Reflectometer.

070676 Kristof, Russell, Cary, Lube and Weismiller. Determining Land Use Patterns Through Man-Machine Analysis of LANDSAT Data - A Tutorial Simulation.

111076 Russell. Systematically Disseminating Technological Information to Potential Users.

110976 Landgrebe, Simmons and Biehl. An Empirical Study of Scanner System Parameters.

090776 Svedlow, McGillem, and Anuta. Analytical and Experimental Design and Analysis of an Optimal Processor for Image Registration.

082776 Montgomery, Baumgardner and Weismiller. An Investigation of the Relationship Between Spectral Reflectance and the Chemical, Physical, and Genetic Characteristics of Soils.

062176 Wiersma and Landgrebe. The Use of Spatial Characteristics for the Improvement of Multispectral Classification of Remotely Sensed Data.

052576 Lindenlaub and Lube. Matrix of Education of Training Materials in Remote Sensing.

051576 Bauer and Davis. Stratification of Landsat Data by Clustering.
*032576 Todd, Mausel, and Baumgardner. Urban Land Use Monitoring from Computer-Implemented Processing of Airborne Multispectral Data.
*031276 Mause1, Todd and Baumgardner. An Analysis of Metropolitan LandUse by Machine Processing of Earth Resources Technology Satellite Data.

052075 DeWitt and Robinson. Description and Operation of a Field Rated ERTS-Band Transmissometer.

* Papers reporting earlier SRT work which has just appeared in print.

OTHER REPORTS
Biehl and Simmons.
Field Measurements Data Library Catalog, Volume I, 1974-75 Crop. October 29, 1976.

Biehl and Simmons.
Field Measurements Data Library Catalog, Volume II, 1975-76 Crop. December 1, 1976.

### 2.1 Test of Boundry Finding/Per field Classification

## INTRODUCTION

One might wish to adopt a new classification technique for one of three reasons:

- It provides greater accuracy than the currently implemented technique..
- The new technique is less expensive than the current technique.
- The classification results produced by the new technique are easier to use and interpret than those currentiy produced.

The ECHO* classification algorithms developed at the Laboratory for Applications of Remote Sensing (IARS) have shown promise in all three areas 1isted above. ${ }^{1}$

Contemporary classifiers for analysis of remotely sensed data compare the spectral measurements from each feature of each point to class statistics, computing a likelihood of discriminant function associated with each class, and categorizing the point according to the class with the largest discriminant function value. Each point is classified individually on the basis of its spectral measurements alone. One premise of this technique is that the objects of interest are large in comparison to the size of a point. If this were not so, a large proportion of points would be composites of several classes, making statistical pattern classification unreliable since pre-specified categories would be inadequate to describe actual states of nature. From this premise it follows that objects are represented by arrays of points, and that a statistical dependence exists between consecutive points. Contemporary classifiers fail to exploit the statistical dependence between adjacent points when assigning classes.

[^0]Giving no consideration to the spectral response of adjacent points contributes to a "salt and pepper" effect in classification results. While the results produced may be accurate, they may be too complex, and can even be confusing when attempting to produce a useful analysis product such as a forest inventory map. ${ }^{2}$ Consequently time and effort have been invested in developing smoothing programs which "clean-up" the "salt and pepper" effects in classification results. ${ }^{2}$

The ECHO processors benefit from spatial information by first aggregating into groups points whose spectral responses are not significantly different in a statistical sense, and then applying a maximum likelihood classification rule to these homogeneous groups. Homogeneous objects are identified in a three step process. First, cells are formed by systematically partitioning the data into $N$ by $N$ sized blocks of pixels. The statistics of each cell are then compared to a homogeneity criterion. Points which do not comprise homogeneous groups are classified on a point-by-point basis, just as contemporary classifiers classify all points. Statistics of adjoining homogeneous cells are then compared to annexation thresholds. Adjoining cells which appear to belong to the same statistical population are combined into a single object.

Two separate ECHO algorithms have been developed. The first, Supervised ECHO, makes use of pre-specified class statistics to identify homogeneous objects. The second, Nonsupervised ECHO, identifies homogeneous objects without the use of class statistics. Consequently, those objects identified by the Nonsupervised algorithm may be used to aid in the training process.

## Past Work at LARS

Much of the background research on the ECHO concept was performed at LARS during 1975 and is documented in the Final Report for $1975,^{3}$ R. L. Ketting's
doctoral thesis, ${ }^{1}$ a LARS Information Note, ${ }^{4}$ and in symposium proceedings. ${ }^{5}$ The partitioning of a data set into homogeneous objects can occur in one of two general approaches. One can utilize a disjunctive approach, starting with large areas and repeatedly subdividing them until each unit appears to be homogeneous. Alternatively, one can choose a conjunctive approach where adjacent units which appear to belong to a single object are combined. One advantage to the conjunctive approach is, if second order statistics are to be used, the second order statistics of the combined area can be calculated from the second order statistics of each constituent unit. The statistics for each subdivision of the initial partition made under the disjunctive approach must be recalculated starting with the individual pixel values, however. Hence the conjunctive approach is potentially faster than the disjunctive approach.

Prior to the 1975 work, Rodd's conjunctive partitioning algorithm ${ }^{6}$ was combined with a minimum distance sample classifier and an improvement in classification accuracy over conventional classification was observed, however processing time was increased. ${ }^{4}$ Gupta and Wintz, ${ }^{7}$ added a second order statistics test to Rodd's first order test, but obtained essentially the same results at an increase in processing time. Robertson ${ }^{8}$ implemented a disjunctive partitioning algorithm with a minimum distance classifier which acheived the same classification accuracy as the conventional point-bypoint classifier with an order of magnitude increase in computer time.

During 1975 new statistical criteria were applied to the partitioning algorithm producing a "Supervised" field extraction routine to add to the Nonsupervised routine. In addition, the sample classifier utilized by ECHO was altered from a minimum distance strategy to a maximum likelihood strategy. Experimental results were obtained for two aircraft and two LANDSAT-1 data sets which indicated that ECHO algorithms improved classification accuracies for the four tested areas and that they reduced the CPU
time required to perform the classifications.
FORTRAN programs were produced implementing the Supervised and Nonsupervised algorithms in two phases. Channel selection, cell size, and cell homogeneity criteria were specified in the first phase of the Supervised processor, annexation in the second. For the Nonsupervised processor, only channel selection and cell size were specified in phase one, cell homogeneity and annexation thresholds were specified in phase two. The Nonsupervised processor did not have the capability to "split" cells; i.e. the constituent points in a singular cell had to be classified as a small sample rather than point-by-point.

Documentation of these programs consisted of commented program listings and a thesis. ${ }^{3}$

Objectives of the 1976-1977 Task
The objectives of the FY77 work were to:

1) Deliver to NASA documented FORTRAN programs implementing the ECHO algorithms, and to make improvements to those algorithms, where appropriate.
2) Systematically test the algorithms on MSS data for agricultural regions as observed by the LANDSAT satellites, aircraft scanners, and on the simulated Thematic Mapper data.
3) Provide products enabling the determination of the utility of the object maps to a LACIE Analyst Interpreter in the selection and labeling of training fields.

## DESCRIPTION OF WORK

## A. Documented Fortran Programs

1. Supervised Processor

FORTRAN Ilstings and program abstracts for the Supervised ECHO processor are presented in Appendix A. This processor requires class statistics for field extraction as well as for classification. The general flow of this algorithm is presented in Figure 1. The software is designed to function in either a single phase or a two phase mode. The two phase approach accomplishes channel selection, cell size, and cell homogeneity criteria in phase one; annexation tests take place during phase two. The intermediate results between phases are stored on magnetic tapes. The final results are written either on disk or on tape. Functioning in the two phase mode allows the analyst to vary the annexation threshold and, thus, to produce several results outputs without repeating the expensive cell homogeneity test.

The single step approach allows the analyst to specify all the necessary input parameters on a single control card deck and produce classification results without the necessity of an intermediate tape. When the analyst wishes to produce classifications with varying annexation thresholds for a particular area, it is less expensive in terms of computer time to utilize the two phase approach. When a single annexation threshold is desired, it is more efficient to use the single phase approach. The Supervised ECHO processor determines whether a single phase classification, the first phase of a two phase classification, or the second phase of a two phase classification is to be generated based on the control cards specified.


Input and output requirements for each processing mode of the Supervised processor are summarized in Table 2.1-1 and Table 2.1-2.

## TABLE 2.1-1

# Supervised ECHO <br> Single Phase Mode 

| Input | Output |
| :--- | :--- |
| Cell Width Parameter | Results File |
| Cell Homogeneity Threshold | Classification Map (optional) |
| Annexation Threshold | Singular Cell Map (optional) |
| Channel Selection |  |
| Specification of the Area |  |
| to Classify |  |
| Multispectral Image Storage |  |
| Tape (MIST) |  |
| Specification of Results |  |
| Destination |  |
| Class Statistics |  |

TABLE 2.1-2
Two Phase Mode
Supervised ECHO

## Input

Phase One
Cell Width Parameter
Cell Homogeneity Threshold
Channel Selection
Specification of the Area to Classify

Multispectral Image Storage
Tape
Intermediate Tape and File Specifications

Class Statistics

Phase Two
Intermediate Results Tape File

Annexation Threshold
Specification of Results
Destination

Intermediate Results Tape
Singular Cell Map (optional)

Figure 2
Effects of Cell Homogeneity Threshold


Figure 3


Photographs of classification results of aircraft data collected over Williams County, Kansas. Notice the "salt and pepper" in the perpoint results (above). The ECHO results (below) have both higher classification performance and lower classification variability. The ECHO classification was produced using $40 \%$ of the CPU time required to produce the perpoint classification.


## 2. Nonsupervised Processor

Output products for this task included those allowing the determination of the utility of the resulting object map to the LACIE Analyst Interpreter for identifying and labeling training fields. Since the Supervised processor utilizes class statistics for the identification of homogeneous objects, it has only limited value as a training aid. The Nonsupervised processor, on the other hand, does not require class statistics for object extraction, and thus, may prove quite beneficial as a training aid, serving, in a sense, as a spatial clustering algorithm.

The Nonsupervised processor utilizes a homogeneity test which compares $\sigma_{i j} / \mu_{i j}$ to a threshold $t_{i}$; where $i$ refers to feature or channel and $j$ refers to the cell. If the standard deviation divided by the mean for feature $i$ exceeds the user-specified threshold $t_{i}$, the cell is "singular" and elements of the cell will be classified on a point-by-point basis. Cells are annexed to fields on the basis of a two-step test, 1) that the channel variances of the field are equivalent to the channel variances of the cell and 2) that the channel means of the field are equivalent to the channel means of the cell. Should either of these criteria not be met for any channel, annexation will not take place.

The research software was designed to run in two phases. The first phase merely calculated the cell mean and covariance matrices and wrote them on tape; the second phase proceeded to perform field extraction followed by classification. This process has two disadvantages. First, although an object map could be produced, it had to be produced in phase two, the same phase that required a class statistics deck for input in order to perform the classification. Second, since only cell mean and covariance matrices were written on the intermediate tape, when singular cells were
identified in phase two, they had to be classified as small samples, a cell at a time, rather than on a point-by-point basis, because data values for individual points were not available to the phase two classifier.

The software was restructured to move the field extraction algorithm into phase one of the processor and produce an intermediate tape which contains:

1. class means and covariance matrices for each homogeneous object identified.
2. an object map containing the mean for channel $i$ of object $j$ in every pixel of object $j$ and the original data values for those pixels belonging to singular cells.
3. a tag array identifying whether a pixel falls in a singular cell (and should be classified individually), or in a homogeneous object (and hence should receive a class assignment based on the sample classification of the object to which it belongs).

The second phase of the restructured Nonsupervised processor reads the intermediate tape and the class statistics deck, performs a maximum likelihood sample classification (utilizing second order statistics where possible) on the objects identified, and a maximum likelihood point-bypoint classification on points falling in singular cells. Figure 4A presents the general processing flow for phase one (field extraction) of the Nonsupervised processor. The general processing flow for phase two (classification) is presented in Figure 4B. Inputs to outputs from the Nonsupervised ECHO processor's two phases are described in Table 2.1-3.

PHASE 1


FIGURE 4B
GENERAI FLOW OF
NONSUPERVISED ECHO
PHASE 2


## TABLE 2.1-3

## NONSUPERVISED ECHO

PROCESSOR

INPUT
Phase One:
Channel Selection
Cell Width
Cell Homogeneity Thresholds ( $\sigma_{i} / \mu_{i}$ for channel i)

Annexation Thresholds
(mean and variance)
Multispectral Image Storage Tape

Intermediate Results Tape and File Specification

OUTPUT

Intermediate Tape containing object map, object statistics and pixel tag array.

## Phase Two:

Class Statistics Results File
Intermediate Tape from
Phase One
Specification of Results
Tape and File

## B. Tests of the ECHO algorithms

1. Data Sets

The second obejctive of Task 2.1 is to test the ECHO algorithms on MSS data for agricultural regions. Data sets are to include LANDSAT, aircraft, and simulated Thematic Mapper data. Ten LANDSAT, three aircraft and eight simulated Thematic Mapper ( 2 sites and 4 resolutions) data sets were selected for the analysis. One bitemporal data set was examined separately. These data sets are summarized in Table 2.1-4

TABLE 2.1-4

Data Sets
(LANDSAT)

| CHANNEL NUMBER |  | WAVE BAND $\mu \mathrm{M}$ |  |
| :---: | :---: | :---: | :---: |
| 1,5,9 |  | .5-. 6 |  |
| 2,6,10 |  | .6-.7 |  |
| 3,7,11 |  | .7-. 8 |  |
| 4,8,12 |  | .8-1.1 |  |
| Area | Channels Used | $\begin{gathered} \text { Data } \\ \text { Collected } \end{gathered}$ | Data Set |
| Graham County, Kansas | 9, 10, 11, 12 | 5/26/74 | LACIE/SRS |
| Grant County, Kansas | 5, 6, 7, 8 | 5/19/74 | LACIE/SRS |
| Haskell County, Kansas | 9, 10, 11, 12 | 5/27/74 | LACIE/SRS |
| Kearny County, Kansas | 9, 10, 11, 12 | 5/27/74 | LACIE/SRS |
| Huntington County, Indiana | 1, 2, 3, 4, | 7/7/73 | CITARS |
| Shelby County, Indiana | 1, 2, 3, 4 | 9/7/73 | CITARS |
| White County, Indiana | 1, 2, 3, 4 | 8/21/73 | CITARS |
| Livingston County, Illinois | 1, 2, 3, 4 | 7/16/73 | CITARS |
| Fayette County, Illinois | 1, 2, 3, 4 | 8/21/73 | CITARS |
| Lee County, Illinois | 1, 2, 3, 4 | 8/5/73 | CITARS |

TABLE 2.1-4 (Continued)
Simulated Thematic Mapper

| TM CHANNEL <br> NUMBER | WAVE BAND |
| :---: | :---: |
| $\mu M$ |  |


| Area | Channels Used | Date | Resolution |
| :--- | :---: | :---: | :---: | :---: |
| Williams County, ND | $1,2,3,6,8$ | $8 / 15 / 75$ | 30 m |
| Williams County, ND | $1,2,3,6,8$ | $8 / 15 / 75$ | 40 m |
| Williams County, ND | $1,2,3,6,8$ | $8 / 15 / 75$ | 50 m |
| Williams County, ND | $1,2,3,6,8$ | $8 / 15 / 75$ | 60 m |
| Finney County, Kansas | $2,3,5,6,7$ | $7 / 6 / 75$ | 30 m |
| Finney County, Kansas | $2,3,5,6,7$ | $7 / 6 / 75$ | 40 m |
| Finney County, Kansas | $2,3,5,6,7$ | $7 / 6 / 75$ | 50 m |
| Finney County, Kansas | $2,3,5,6,7$ | $7 / 6 / 75$ | 60 m |

## TABLE 2.1-4 (Continued)

|  | Aircraft |  |
| :---: | :---: | :---: |
| Area | Wavebands Used ( $\mu \mathrm{M}$ ) | Data Collected |
| Williams County, ND | .4-.49, .59-. 64, .65-.69, | 8/15/75 |
| Finney County, Kansas | .82-.88, 1.53-1.62, 10.1-11.0 | 7/6/75 |
| Tippecanoe County, IN | . 52-. 57, .61-.70, | 8/13/71 |
|  | .72-.92, 1.0-1.4, 1.5-1.8 |  |
|  | Bitemporal |  |
| Area | Wavebands Used <br> for Each Date <br> Dates | Data Set |
| Grant County, Kansas | .5-.6, .6-.7, 5/9/74 | LACIE/SRS |
|  | .7-.8, .8-1.1 6/14/74 |  |

## 2. Training Procedures

a. LANDSAT Data Sets

The LACIE training sets were created using ground truth information provided by JSC for both test and training fields. Odd-numbered fields appearing in each subclass were used for training; even-numbered fields were used for test. For Graham and Grant counties, statistics were generated by using the STATISTICS processor. For Haskell and Kearny counties statistics were secured by clustering the training fields of each class into four subclasses which were then pooled on the basis of output from the SEPARABILITY processor. It should be noted that both training and test fields for the LACIE/SRS data are large enough to inset field boundaries approximately two pixels inside estimated field boundaries. This inset allows for any image misregistration which may occur between any two dates on the multitemporal input runs. Proportion estimates for the 1974 LACIE/SRS segments were provided in ground truth packets provided by JSC.

The CITARS training sets were originally created by a supervised procedure using ground truth information provided by the Agricultural Stabilization and Conservation Services (ASCS) to select both the training and test fields. ${ }^{9}$ Those training sets were used without change in the ЕСНО tests.

Six counties in the CITARS experiment were used as test sites. A data set free of clouds which occured late in the growing season was required for each of the six counties. Dates from July, August or early September were selected. The classification results for all of the CITARS experiments are catalogued on a series of LARS tapes. The desired data sets were located on the catalogued CITARS tapes and the statistics which had been used for the CITARS experiment were obtained by using the LARSYS PUNCHSTATISTICS processor. The pooling of classes was determined by running
the PRINTRESULTS processor and requesting training field results. This combination of requests produces a table of available classes in the statistics deck and also the informational names under which they were classified. By using this list it was possible to reconstruct the combinations of spectral classes and pooled spectral classes which were needed to reproduce the original classification.

Training and test fields for each CITARS classification were secured and appropriate control cards added to evaluate the ECHO classifications of CITARS data sets. The proportion estimates used for the CITARS evaluation are estimates of the proportions of the various classes for the entire county made by the SRS.
b. Simulated Thematic Mapper Data Sets

Training sets from the simulated Thematic Mapper tests performed at LARS in 1976 were used for ECHO analysis. These training sets were generated by selecting fields of known cover types ana ciustering each informational class separately to define subclasses.

Color infrared photographic mosaic prints were made from photographic data collected concurrently with the scanner data. Informational class information provided by ground observations was transfered to clear plastic overlays on the mosaic print. The analyst could then easily locate the corresponding fields in his cluster maps and assign the field coordinates to the informational classes.

Statistics were calculated for each training area and compared using the SEPARABILITY processor. Similar classes were combined, where indicated, and the data set was used to classify the flightline. Training areas were not excluded from the test fields since the test fields had been preselected for the entire flightline.

Two to four subclasses were found in each informational class. The Kansas flight was an exception. Because of severe line-to-line changes in signal level in the original 6 meter scanner data, the analyst was forced to create more spectral classes to account for the within-class variations due to excessive noise. This was most apparent in the 30 and 40 meter resolution data. The effect was reduced but not eliminated in the 50 and 60 meter resolution data. Alternate fields were used for training and test decks, respectively. The procedure was repeated for each of the four resolution sizes. As resolution size increased, the number of spectral classes decreased.

The entire training set selection procedure was repeated for each resolution size so that any effects on training set selection which might be caused by data resolution would be included in the analysis results. An example is the increasing difficulty and eventual impossibility of selecting samples from small, or narrow, fields as the resolution size increases.
c. Aircraft

Six Meter NASA Aircraft Scanner Data

The two 6-meter aircraft scanner data sets used to generate the simulated Thematic Mapper data set were used. The same training and test fields used for the Thematic Mapper were available in the six meter data set. Because of the vary large number of data points in the full data set, only the first two miles (one third of the total flight line) were used for the ECHO evaluations.

These data sets were not corrected for sun and scanner angle effects. To compensate for these angle effects, training fields were distributed across the width of the flight line. At least six fields in each informational class were used in the training set. Fields from informational
classes were clustered together into five spectral classes. All spectral classes from all informational classes were combined into a statistics deck and appropriate pooling was done based on SEPARABILITY results. The classification results were strongly influenced by angle effects.

Flightline 210, Corn Blight Watch Experiment
Training and test fields from the Michigan Scanner data set were located on the LARS display system using ground information collected during the Corn Blight Watch Experiment. At least six fields were selected to represent each informational class. Each informational class was then clustered to produce five spectral subclasses. All spectral classes were combined and processed with SEPARABILITY to select the subset of the original 12 channels which was to be used for classification. A subset of five channels was chosen and the training set was again processed with SEPARABILITY to determine which classes should be pooled.

## d. Bitemporal Data Set

A bitemporal data set from Grant County, Kansas was created by selecting data from a spring date and a mid-summer date. Training and test fields were available from the LACIE data and were edited to reduce the number of single line training fields by about 30 percent. Data from the two dates were clustered together within each informational class with four spectral classes requested. Classes were grouped as indicated by the CLUSTER algorithm and classes with fewer than 50 points were deleted.

## 3. Dependent Variables

There are six variables which were monitored to evaluate the ECHO algorithm:

- CPU time
- Field center pixel classification performance
- Training field classification performance
- Full field classification performance
- RMS proportion estimate exror
- Classification variability

These variables are related to the stated reasons for adopting a new classification technique listed in the introduction. The CPU time required to perform a classification is one way to measure the cost of classification. Field center pixel, full field, and training field performances and RMS proportion estimate error are all ways to evaluate the accuracy of the classifier. Classification variability is a measure of "salt and pepper effect" in classification results.

The CPU time required to execute each of the ECHO classifications has been recorded so that the effects of varying the cell homogeneity and annexation thresholds may be monitored. The CPU times required to perform the perpoint classifications have been adjusted to reflect the increased efficiency of the LARSYS perpoint classifier which is coded in assembly language. Thus, the CPU time recorded for a perpoint classification is what a FORTRAN classifier would have required to perform the classification.

The indices of classification performance were applied in several ways. Classification accuracy (identification) was evaluated utilizing field center pixel, "full field" and test field sample performance for all data sets. Proportion estimation was carried out for the LANDSAT
and Simulated Thematic Mapper data sets.
The training performance is the overall classification accuracy (number of training pixels correctly classified divided by the total number of training pixels) of the pixels used to calculate the class statistics. Field center pixel performance is the overall classification accuracy of field center pixels. Field center pixels are pixels inset at least one pixel from the field boundaries. For the registered LACIE/SRS data the field center pixels are inset at least two pixels from the field boundary. Although this procedure insures that the pixels examined are not mixture pixels, it has the unfortunate effect of eliminating smaller fields from consideration. The third measure of classification accuracy, "full field" performance, includes those pixels on the boundaries of the fields in the classification performance. The "full field" pixels were generated by expanding the field center pixel boundaries one pixel in all directions.

The RMS error of informational class proportion estimates for each flightline was found by calculating the percent of the flightine classified as a particular class and comparing it with the ground-collected estimate using equation (1).

$$
\begin{equation*}
\text { RMS Error }=\sqrt{\frac{\sum_{i=1}^{N}\left(C_{i}-C_{i}^{\prime}\right)^{2}}{N}} \tag{1}
\end{equation*}
$$

where, $N=$ number of informational classes
$C_{i}=$ percent classified as informational class i
$C^{\prime}{ }_{i}=$ percent of class $i$ estimated from ground-collected data RMS error is calculated for the LANDSAT and Thematic Mapper data runs. The Agricultural Stabilization and Conservation Service (ASCS) provided the ground truth proportion estimates for the simulated Thenatic

Mapper data set. Proportion estimates for the 1974 LACTE/SRS segments were provided in ground truth packets received from JSC. The SRS county proportion estimates were used to calculate RMS proportion error for the CITARS data set.

Average variability is a measure of the rate of change from one information class to another. It should reflect the degree to which ECHO reduces the "salt and pepper" effect which is sometimes present in perpoint classifications. Variability is calculated by systematically selecting 50 lines of the classified area, counting the number of information class changes, and dividing by the number of opportunities for class changes.

$$
\begin{equation*}
\text { Variability }=\text { NCC } /\left(50^{*}(N S-1)\right) \tag{2}
\end{equation*}
$$

Where
NCC $=$ the number of class changes over the 50 selected lines
$N S=$ the number of classified pixels/line

## 4. Results

This section outlines the results of the tests performed on the Supervised processor. Results are discussed separately for each scanner type, LANDSAT, simulated Thematic Mapper, and aircraft, followed by a summary of results for all three data types.
a. LANDSAT Results

Training and test information for the 10 data sets comprising the LANDSAT test data were drawn from the 1974 Kansas LACIE/SRS data sets and the 1973 CITARS data sets. One bitemporal and four unitemporal analyses were performed on the LACIE/SRS data sets. Six unitemporal analyses were performed on the CITARS data sets.

## i. LACIE/SRS Results

Figures 5 through 10 present the average results for the four LACIE/SRS sites examined at a cell width of two for six cell homogenefty and four annexation parameters. Figure 5 plots the average CPU time in seconds required by the perpoint classifier (represented by the line of $P^{\prime}$ ) versus the CPU time required by the average of the four LACIE/SRS data sets at each of 24 ECHO parameter settings for 2 by 2 pixel cells. The horizontal axis is the cell homogeneity threshold. As this threshold increases, it becomes more likely a cell will be classified as a unit, less likely that a cell will be split and its constituent pixels classified individually. The dependent variable, CPU time, appears on the vertical axis. The cell to field annexation parameter for each cell homogeneity threshold appears on the plot. When two or more annexation thresholds acheive the same performance, a star appears in that position on the plot. It can be seen from Figure 5 that as the cell homogeneity parameter increases, the CPU time required to classify the area decreases. Additionally, as the annexation parameter increases from zero to four, the CPU time required for the classification is reduced. These results are expected because fewer

## GRAPH OF EPU FOX 4 AREA(S) GZAHAM SRAYT HASKELL KERZYY <br> EELL SIRE IS $2 \times 2$


GRAPH OF FCP
URAHAM GRAYI AREASS
HASKELL XEARYY

ELLL SILE IS $2 \times 2$


```
~RAPH OF FULL FLD FOR & AREAISI
    GRAHAM SRAyT MASKELL KEARYY
    EELL SILE IS 2 x2
```



```
    graph dF tratyIvg FOX 4 arcatss
        gRaham GRAVI HASKELI
        *ELL SILE IS 2 X2
```




TABLE 2.1-5
EFFECT OF ECHO PARAMETERS ON INDIVIDUAL 1974 LACIE DATA SETS

## ABBREVIATIONS:

| CPU | CENTRAL PROCESSING UNIT TIME |
| :--- | :--- |
| FCP | FIELD CENTER PIXEL PERFORMANCE |
| FULL | FULL FIELD PERFORMANCE |
| TRAIN | TRAINING FIELD PERFORMANCE |
| RMS | RMS PROPORTION ERROR |
| VAR | CLASSFICATION VARIABILITY |
| CELW | CELL WIDTH |
| HOM | CELL HOMOGENEITY THRESHOLD |
| ANN | ANNEXATION THRESHOLD |

ANALYSIS OF VARIANCE RESULTS (ENTRIES ARE SIGNIFICANCE LEVELS):
CELW HOM ANN

| CELW | CELW | HOM |
| :---: | :---: | :---: |
| X | X | X |
| HOM | ANN | ANN |

Graham County, Kansas

| CPU | .1 | .1 | 7.5 | .1 | 24.7 | 99.9 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| FCP | .1 | .1 | 18.7 | .1 | .1 | 18.9 |
| FULL | .1 | .1 | 99.9 | .1 | 1.6 | 33.6 |
| TRAIN | .1 | .1 | .1 | .1 | .1 | .2 |
| RMS | .1 | .1 | .1 | .1 | .1 | 99.9 |
| VAR | .1 | .1 | .1 | .1 | .1 | .1 |

Grant County, Kansas

| CPU | .1 | .1 | .8 | .1 | 2.2 | 99.9 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| FCP | .1 | .1 | .1 | .1 | .1 | 99.9 |
| FULL | .1 | .1 | .1 | .1 | .1 | 99.9 |
| TRAIN | .1 | .1 | .1 | .1 | .1 | 34.9 |
| RMS | .1 | .1 | 99.9 | .1 | 99.9 | 41.8 |
| VAR | .1 | .1 | .1 | .1 | .1 | 40.3 |

Haskell County, Kansas

| CPU | .1 | .1 | .2 | .1 | .1 | 33.0 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| FCP | .1 | .1 | .1 | .1 | .1 | 25.5 |
| FULL | .1 | .1 | .3 | .1 | .1 | 33.5 |
| TRAIN | .1 | .1 | .1 | .1 | .1 | 29.5 |
| RMS | .1 | .1 | .1 | .1 | .1 | 9.1 |
| VAR | .1 | .1 | .1 | .1 | .1 | 27.1 |

## TABLE 2.1-5 (Continued)

EFFECT OF ECHO PARAMETERS ON INDIVIDUAL 1974 LACIE DATA SETS

|  |  | CELW | CELW | HOM |
| :---: | :---: | :---: | :---: | :---: |
| CELW HOM | ANN | X | X | X |
|  |  | HOM | ANN | ANN |

## Kearny County, Kansas

| CPU | .1 | .1 | .1 | .1 | .4 | 99.9 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| FCP | .1 | .1 | 6.5 | .1 | .1 | 99.9 |
| FULL | .1 | .1 | 1.6 | .1 | .1 | 99.9 |
| TRAIN | .1 | .1 | .3 | .1 | .1 | 13.0 |
| RMS | .1 | .1 | .1 | .1 | 15.3 | 2.0 |
| VAR | .1 | .1 | .1 | .1 | .1 | .1 |

Table 2.1-6

EFFECT OF ECHO PARAMETERS ON FOUR LACIE DATA SETS

|  | Cell <br> Width | $\begin{gathered} \text { Cell } \\ \text { Homogeneity } \\ \text { Threshold } \\ \hline \end{gathered}$ | Annexation Threshold | ```Cell Width X Cell Homogeneity Threshold``` | Cell Width X <br> Annexation Threshold | Cell Homogeneity <br> Threshold <br> X <br> Annexation <br> Threshold |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Degrees of Freedom | 2,6 | 5,15 | 3,9 |  |  |  |
| CPU Time |  | 1 | 1 | . 1 | 99.9 | 99.9 |
| Field Center Pixel Performance | 5 | 5 |  | . 1 | . 4 | 99.9 |
| Full Field <br> Performance | 5 | 1 | 25 | . 1 | . 6 | 99.9 |
| Training Performance | 10 | 1 |  | . 1 | . 1 | 99.9 |
| RMS Error |  | 25 |  | . 1 | 99.9 | 99.9 |
| Classification Variability | 5 | 1 | 1 | . 1 | . 1 | 13.6 |

classifications are necessary when there are fewer cells which must be classified on a point-by-point basis (due to the increased cell homogeneity parameter), and fewer classifications are required when a scene is composed of a few large objects rather than many small objects (the effect of increasing the annexation parameter). Both of these effects are statistically significant at a $1 \%$ confidence level.

The statistical significance levels of the various ECHO parameters on each of the six dependent variables (CPU time, field center pixel performance, "full field" performance, training field performance, RMS proportion error, and classification variability) for each of the LACIE data sets are presented in Table 2.1-5. Table $2.1-6$ presents the statistical significance levels of the ECHO parameters on the dependent variables for the four LACIE data sets combined.

## Parameter Considerations

For the LACIE data runs, cell widths were sampled at 2,
$\sqrt{\text { average field size }} / 2+1$, and $\sqrt{\text { average fields size; }}$ hence cell widths sampled were 2,5 or 6 , and 8 or 9 . The field center pixel performances at the 2 by 2 cell size tended to be superior to those at 5 by 5 or 6 by 6 ; RMS proportion estimation errors of 2 by 2 cells were about equivalent to those of the 5 by 5 or 6 by 6 cells; and less computer time was required for the 5 by 5 or 6 by 6 sized cells than the 2 by 2 cells.

Graham County was also sampled at cell sizes 3 by 3 and 4 by 4 to observe what effects intermediate cell sizes might have. Results for the optimal values of these parameters are presented below:

Graham County, Kansas
Optimal Results

|  | CPU <br> Range <br> (Sec) | FCP <br> Performance <br> $\%$ | RMS <br> Error <br> $\%$ |
| :---: | :---: | :---: | :---: |
| $2 \times 2$ | $109-73$ | 93.9 | 0.1 |
| $3 \times 3$ | $101-59$ | 94.3 | 0.1 |
| $4 \times 4$ | $95-50$ | 93.9 | 0.0 |
| $6 \times 6$ | $98-51$ | 93.3 | 0.7 |
| Perpoint | 121 | 89.9 | 1.2 |

These results tend to indicate that cell width may be optimized for the LACIE/SRS runs when it is in the neighborhood of the fourth root of the average field size. This value is between three and four for Graham County.

For the 2 by 2 tests (performed on all four LACIE/SRS sites) the optimal cell homogeneity parameter, based on the three fields performance criteria, appears to be 78. However, on the basis of RMS error, the optimal homogeneity parameter tested is 39.

Due to the two pixel inset used to avoid mixture pixels and registration error for the LACIE/SRS field coordinates, small fields were not used for training or test purposes. Since only relatively large fields are included in the training decks, the effects that increasing the cell homogeneity parameter may have on the small fields is not reflected in the field center pixel performance results. Consequently, for the LACIE results, 39 is probably closer to the optimal homogeneity parameter than 78 . Support for this conclusion is provided by the results in Figure 8. The "full field" performance measures the performance of the large test fields, including the points which contain the field boundaries. "Full field" test fields are composed of the pixels contained in the field center pixel test fields
plus those pixels on the borders of the field ("fringe pixels"). Since for the LACIE data sets, the performance of the field center pixels improves by . $5 \%$ between cell homogeneity values 39 and 78 while the full field performance only improved by $.2 \%$, it would appear that the performance of the "fringe pixels" is degraded between homogeneity parameters 39 and 78. Those small fields not included in the test decks will be composed primarily of "fringe pixels".

It can be noted by examining Figures 6 through 9 that the effect of annexation increases as the cell homogeneity parameter increases. This result is expected since a cell A only has the opportunity to be annexed to adjoining cell $B$ when both cells $A$ and $B$ are homogeneous. The higher the cell homogeneity parameter, the more likely it becomes that both $A$ and B will be homogeneous.

Figure 10 is typical of all the variability results: as the homogeneity and annexation parameters increase, the classification variability decreases.

## LACIE/SRS Conclusions

1. CPU time and classification variablity decrease as cell homogeneity and cell to field annexation parameters increase ( $1 \%$ significance level).
2. The optimal cell width parameter may fall between two and average field size $/ 2+1$. The fourth root of the average field size may be an appropriate value.
3. For these data sets, the optimal annexation value appears to be in the neighborhood of two based on field center pixel and full field performances.
4. For a cell width of two, the optimal cell homogeneity parameter is in the neighborhood of 39 , based on RMS proportion estimate error.

## ii. CITARS Results

Figures 11 through 16 present plots for the 2 by 2 cell size Supervised ECHO results achieved from the CITARS data sets. Characteristic differences between the CITARS and the LACIE/SRS data sets include:

* The CITARS data set has a much smaller average field size than the LACIE/SRS set.
* The information classes are different. CITARS information classes are Corn, Soybeans and Other; LACIE/SRS classes are Wheat and Other.
* The ground truth proportion estimates used for the LACIE/SRS sites were for the area of the LACIE segment. The ground truth proportion estimates used for the CITARS sites are for the whole county in which the data set lies, not for that area of the county which was actually sampled.

Analysis of variance results for the six CITARS data sets are presented in Tables 2.1-7 and 2.1-8. Table $2.1-7$ presents results for each data set individually. Table 2.1-8 presents the analysis of variance results for the combined CITARS data sets.

As with the 1974 LACIE/SRS data sets, the CPU time decreased as cell homogeneity and annexation parameters increase.

However, though Figure 5 (CPU time for the LACIE/SRS data set) and Figure 11 (CITARS CPU time) have the same general shape, there is one important difference: the CPU time required to perform the ECHO classification is a higher percentage of the CPU time required to perform the perpoint classification for each cell homogeneity value. This is a reflection of the smaller field size in the CITARS data set. Since fields are smaller, more boundaries and fewer homogeneous cells exist. As a result, more cells must be classified on a point-by-point basis and more CPU time




GRAPH OF VARBLIT FDZ O AREAISI
WH-Y LE-IV SN-VI FA-Y HJ-III LYSTOY
EELL SILE IS $2 \times 2$


TABLE 2.1-7
EFFECT OF ECHO PARAMETERS ON INDIVIDUAL CITARS DATA SETS
ABBREVIATIONS:

| CPU | CENTRAL PROCESSING UNIT (COMPUTER) TIME |
| :--- | :--- |
| FCP | FIELD CENTER PIXEL PERFORMANCE |
| FULL | FULL FIELD PERFORMANCE |
| TRAIN | TRAINING FIELD PERFORMANCE |
| RMS | RMS PROPORTION ESTIMATE ERROR |
| VAR | CLASSIFICATION VARIABILITY |
| CELW | CELL WIDTH |
| HOM | CELL HOMOGENEITY THRESHOLD |
| ANN | ANNEXATION THRESHOLD |

ANALYSIS OF VARIANCE RESULTS (ENTRIES ARE SIGNIFICANCE LEVELS):

| CELW HOM ANN |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  | $X$ | $X$ | HOM |
|  |  | HOM | ANN | ANN |

Huntington County, Indiana

| CPU | .1 | .1 | .1 | .1 | .1 | 12.7 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| FCP | .7 | .1 | 99.9 | .1 | .3 | 5.9 |
| FULL | 37.4 | .1 | 37.1 | .1 | 1.3 | 20.6 |
| TRAIN | .1 | .1 | .1 | .1 | .1 | 12.9 |
| RMS | .1 | .1 | 99.9 | .1 | 32.7 | 39.2 |
| VAR | .1 | .1 | .1 | .1 | .1 | .1 |

Shelby County, Indiana

| CPU | .1 | .1 | .1 | .1 | .1 | 44.7 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| FCP | .1 | .1 | .1 | .1 | .1 | .1 |
| FULL | .1 | .1 | .1 | .1 | .1 | .1 |
| TRAIN | .1 | .4 | .1 | 27.2 | .1 | .8 |
| RMS | .1 | .1 | .1 | .1 | .1 | .1 |
| VAR | .1 | .1 | .1 | .1 | .1 | .1 |

White County, Indiana

| CPU | .1 | .1 | .1 | .1 | .4 | 99.9 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| FCP | .1 | .1 | .1 | .1 | .1 | 99.9 |
| FULL | .1 | .1 | .1 | .1 | .1 | 22.8 |
| TRAIN | .1 | .1 | .1 | .1 | .1 | 99.9 |
| RMS | .1 | .1 | .1 | .1 | .1 | 43.1 |
| VAR | .1 | .1 | .1 | 1.1 | .1 | .2 |

TABLE 2.1-7 (Continued)
EFFECT OF ECHO PARAMETERS ON INDIVIDUAL CITARS DATA SETS

|  |  | CELW | CELW | HOM |
| :---: | :---: | :---: | :---: | :---: |
| CELW |  | HOM | ANN | $X$ |
|  |  | $X$ | $X$ |  |
|  |  | HOM | ANN | ANN |

Livingston County, Illinois

| CPU | .1 | .1 | .1 | .1 | .1 | 23.0 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| FCP | .1 | 14.8 | .1 | .1 | .1 | 22.1 |
| FULL | .1 | .1 | .1 | .1 | .1 | 3.2 |
| TRAIN | .1 | .1 | .1 | .1 | .1 | 18.7 |
| RMS | .1 | .1 | .1 | .1 | 1.8 | 12.9 |
| VAR | .1 | .1 | .1 | .1 | .1 | .1 |

Fayette County, Illinois

| CPU | .1 | .1 | 1.1 | .1 | 3.2 | 99.9 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| FCP | .1 | .1 | .1 | .1 | .1 | 17.8 |
| FULL | .1 | .1 | .1 | .1 | 1.8 | 17.1 |
| TRAIN | .1 | .1 | 2.9 | 12.6 | .4 | 99.9 |
| RMS | .1 | .1 | .1 | .1 | .1 | .1 |
| VAR | .1 | .1 | .1 | .1 | .1 | .3 |

Lee County, Illinois

| CPU | .1 | .1 | .1 | .1 | .1 | 17.3 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| FCP | .1 | .1 | 34.9 | 39.4 | .1 | 99.9 |
| FULL | .1 | .1 | 99.9 | 7.5 | .1 | 99.9 |
| TRAIN | .1 | .1 | .1 | .1 | .2 | 11.0 |
| RMS | .2 | .1 | .1 | .1 | .1 | .1 |
| VAR | .1 | .1 | .1 | .1 | .1 | 1.8 |

Table 2.1-8

is required to perform those point-by-point classifications. For the smallest cell homogeneity parameter, ECHO takes more time on the average than the perpoint classifier. This is due to a great majority of the cells being split. For a cell that is split, the ECHO processor must first gather statistics and perform the homogeneity test, at some expense in CPU time. Then each point of the singular cell must be classified on a point-by-point basis which takes approximately the same CPU time as the perpoint classifier. Hence, when the great majority of cells are singular, ECHO is less efficient, in terms of CPU time, than the perpoint classifier. By examining the CITARS field center pixel performance (Figure 12) in light of the LACIE/SRS results, two conclusions can be drawn:

1. The improvement in classification accuracy the ECHO procedure provides is reduced for the CITARS data sets, as compared to the LACIE set.
2. The optimal annexation parameter is smaller for the CITARS data sets than for the LACIE data sets.

It is evident by comparing Figures 12 and 14 , that the CITARS training set is not representative of the test fields for the ECHO data sets. The performance of the training fields using ECHO is markedly superior to the performance of those same fields using the perpoint classifier. However, the field center pixel performance of the ECHO algorithm is only marginally superior in magnitude to the performance of the perpoint classifier. This indicates that on the data which the statistics truly represent, ECHO performance is superior; when the first and second order statistics fail to properly represent the classes, especially in areas with small field sizes, ECHO's performance may be no better than that of the perpoint classifier.

The average RMS error achieved by ECHO for the six CITARS data sets is inferior to the average RMS error achieved by the perpoint classifier using the same class statistics. On a flightline by flightline basis,

ECHO's RMS proportion estimate error was higher than the perpoint RMS error for three of the data sets and lower than the perpoint RMS error for the other three data sets. This may be partially due to one or more of the following:

* The RMS error values are generally high. The proportion estimates used for "truth" are the SRS estimates for the counties as a whole, not for the individual areas classified. Hence, if an area classified is not a good approximation of the county as a whole, the validity of the RMS results must be questioned.
* The training fields do not appear to be a good representation of the test fields. If the training statistics do not represent the scene as a whole, ECHO performance may suffer to a greater extent than the perpoint performance.
* When the average field size is small, there will be a high proportion of "mixture" cells (cells containing points of more than one class). Should a cell containing pixels of differing classes pass the homogeneity criteria, it is likely to be categorized into the class with the highest varfance. Thus as the cell homogeneity criteria becomes larger, the RMS error would tend to increase. (An observed result - see Figure 15.)

The effect of cell width on RMS error was not significant when all the CITARS data sets are considered at once. However, cell width produced significantly different results in terms of RMS error for each of the six data sets taken individually. The effects of increasing cell width were opposite for different data sets.

## Conclusions

* As the field size becomes smaller, the improvement in classification accuracy provided by the supervised ECHO algorithm is reduced. At the same time, due to the increased number of singular cells, the CPU time which Supervised ECHO requires increases.
* When the training set is representative of the data, ECHO's classification performance continues to be superior to that of the perpoint classifier for scenes with average fields at or below 16 to 20 pixels.
* Due to the mixture-cell, large-variance-class problem, ECHO may or may not degrade RMS proportion Estimation for areas with relatively small fields.
iii. Bitemporal Results

Figures 17 through 22 illustrate the Supervised ECHO results for the Bitemporal analysis. Figures 17a through 22a are the results of a unitemporal analysis of the same area. The training for the bitemporal analysis was done independently of the training for the unitemporal analysis. Comparing Figures 18 and 18a it can be seen that the ECHO field center pixel performance is a greater improvement over the perpoint classifer's performance in the unitemporal case. This may partially be a result of some misregistration between dates making object extraction more difficult and partially a function of the different training statistics used for the separate classifications.

With the larger number of channels in the multidate classification, the optimal homogeneity threshold is higher than it would be for the single date classification. This is expected as, for the Supervised algorithm, the homogeneity criteria approximates to a chi-squared statistic with (number of channels) times (cell width squared) degrees of freedom. For the same significance level of test, the cell homogeneity criteria should be higher for the eight channel data than for the four channel data.

Based on the field center pixel performance and the RMS proportion error results, the optimal settings of the homogeneity and annexation parameters appear to be in the neighborhood of 180 and 1 , respectively.

Note that although the multitemporal analysis does not greatly improve the field center pixel classification accuracy, it does (especially for the perpoint classifier) substantially reduce the RMS proportion estimate error. In the Bitemporal analysis the perpoint RMS error is only . $2 \%$, leaving ECHO little room for improvement. Note also that for several ECHO parameter settings in the unitemporal analysis, ECHO performs with an RMS



hiEA GRANT CELL MJOTH 2


## grafh df fult flo fot 1 matais

 GXAntB!gell sice is $2 \times 2$

grani county bitemporal fult field performance


## GRAPH OF IRAINIVG FOR I areats

 ołavia!SELL SILE IS $2 \times 2$

grant county bitemporal training field pereoryance



GRANT COUNTY BITEMPORAL RAS PROPORTIOY ESTLMATE ERROR


F16Usb 21a
grant county unitemporal rms proportion estimate error

## GRAPH DF VARELTY foz 1 areats

CXAYIB1
EEL SILE IS $2 \times 2$


proportion estimate error of under $.5 \%$ while the perpoint classifier is functioning with a 6.5\% RMS error.

In the multitemporal case, ECHO requires less CPU time than the perpoint classifier for all parameter settings. This is not true in the unitemporal case.

Bitemporal Conclusions

* The Supervised ECHO processor produced classifications superior to those produced by the perpoint classifier in terms of CPU time and field center pixel accuracy for the frame analyzed.
* The optimal homogeneity parameter for bitemporal analysis appears to be much larger than for the unitemporal case.
* The CPU time required and classification variability produced by ECHO are smaller than those of the perpoint classifier.
iv. Overall LANDSAT Results

Table 2.1-9 summarizes the effects of the ECHO parameter settings on the six dependent variables. Figures 23 through 28 plot the average ECHO performance at the various ECHO parameter settings versus the average perpoint performance for the ten LANDSAT data sets.

The cell width setting effects are significant for all dependent variables for all LANDSAT data sets considered individually, and are significant for all dependent variables except RMS proportion error when all LANDSAT data sets are considered together, using data sets as blocks.

When the LACIE and CITARS LANDSAT runs are considered separately, cell width is significant for CPU time, training performance, and classification variability for the CITARS data sets and for CPU time, field center pixel performance, "full field" performance, and classification variability in the LACIE data sets.

The statistically significant (5\% significance level) overall cell width effects for the LANDSAT data sets are:

- As cell width increases, CPU time decreases.
- As cell width increases, field center pixel, "full field" and training field performances decrease.
- As cell width increases, classification variability decreases. Though not statistically significant, the trend is for RMS error to increase as cell width increases. Two things should be noted about the cell width overall results:
* It is not uniformly true that all cell width increases, field center pixel, "full field" and training field performances decrease.

Table 2.1-9
EFFECT OF ECHO PARAMETERS ON SIX VARIABLES USING DATA SET
AS A BLOCK (ALL TEN LANDSAT DATA SETS)

|  | Cell <br> Width | Cell <br> Homogeneity Threshold | Annexation Threshold | Cell Width <br> X <br> Cell Homogeneity <br> Threshold | ```Cell Width X Annexation Threshold``` | Cell Homogeneity Threshold X <br> Annexation Threshold |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Degrees of Freedom | 2,18 | 5,45 | 3,27 |  |  |  |
| CPU Time | 5 | 1 | 1 | . 1 | 99.9 | 99.9 |
| Field Center Pixel Performance | 1 | 1 | 25 | . 1 | . 1 | 31.5 |
| Full Field <br> Performance | 1 | 1 | 25 | . 1 | . 1 | 99.9 |
| Training Performance | 1 | 1 | 10 | . 1 | . 1 | 1.1 |
| RMS Error | 25 | 10 | 25 | . 1 | 3.2 | 99.9 |
| Classification Variabılity | 1 | 1 | 1 | . 1 | . 1 | . 1 |

[^1]

| CRAPH OF FCP | Fox 10 | $\underset{\substack{\text { areals } \\ \text { hasxele }}}{ }$ |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| טганам | grayt | hasxelt | kearny | -v | Le-tv | SH-YI | FA-v | HO HO | Livston |

Eell size is $2 \times 2$


landsat fule filld performance



EELL SIIE $152 \times 2$




* Since not all cell widths were sampled between two and the $\sqrt{\text { average field size }} / 2+1$, it may be true that the optimal cell width for a data set lies between these two parameters. Indications from the Graham County analysis are that the optimal cell width may be in the neighborhood of the fourth root of the average field size.

The cell homogeneity threshold is significant for all dependent variables both when all LANDSAT data sets are considered individually and when all LANDSAT data sets are considered together.

The cell homogeneity threshold is significant for CPU time, "full field" performance, training performance, and classification variability for the combined CITARS data sets and is significant for CPU time, field center pixel performance, "full field" performance, training performance and classification variability in the combined LACIE data sets.

The statistically significant overall effects of the cell homogeneity parameter are:

- As the homogeneity parameter increases, CPU time and variability decrease.
- As the homogeneity parameter increases, field center pixel, "full field" and training performances decrease. Figures 24 through 26 indicate, however, that for the first two or three homogeneity threshold values, as the homogeneity threshold increases, the performances also increase.

As the homogeneity parameter increases, RMS error shows a trend towards increasing (Figure 27) though this trendis not significant at a $5 \%$ confidence level.

The annexation parameter is significant in $86 \%$ of the 60 cases in the LANDSAT data sets. It is significant only for CPU time and classification variability when the CITARS or the LACIE or all the LANDSAT data sets are
considered together. This indicates that the proper annexation parameter is highly data dependent. As the annexation threshold increases, CPU time and classification variability tend to decrease.

The interaction of cell width and the cell homogeneity threshold is significant in $95 \%$ of the 60 cases of 6 dependent variables and 10 LANDSAT data sets. This interaction is significant for all dependent variables when the CITARS or the LACIE or all LANDSAT data sets are considered together.

The interaction of cell width and annexation parameter is significant in $93 \%$ of the 60 LANDSAT cases. When all the LANDSAT data sets are considered together, the cell width-annexation interaction is significant for all dependent variables except CPU time. This is also true for the CITARS data sets. However, for the LACIE data sets, the interaction is not significant for RMS proportion error.

The cell homogeneity-annexation interaction was significant in $30 \%$ of the 60 individual LANDSAT data set ANOVAs, being significant in $90 \%$ of the data sets for variability and $50 \%$ of the data sets for RMS proportion error. When the 10 LANDSAT data sets are considered together, the cell homogeneity-annexation interaction is significant for training performance and variability. When considering only the CITARS or only the LACIE data sets, the interaction is significant for variability alone.

ECHO-Perpoint LANDSAT Comparison
When the optimal parameter setting for each dependent variable was selected from the ECHO results and compared to the perpoint results for the 10 LANDSAT data sets the following results were obtained:

- ECHO was faster than the perpoint classifier (.1\% confidence level).
- ECHO's field center pixel, "full field", and training field performances were superior to those of the perpoint classifier (1.3, 5.7, . $1 \%$ confidence levels, respectively).
- ECHO had a lower RMS proportion estimate error (4.4\% confidence 1eve1).
- ECHO had less classification variability than the perpoint classifier (.1\% confidence level).
b. Simulated Thematic Mapper Results

The analyses of variance for the effects of the ECHO parameters on the eight simulated Thematic Mapper data sets (two sites, four resolutions) are presented in Table 2.1-10. Results for the eight data sets considered together are presented in Table 2.1-11.

Cell width is significant for all dependent variables and all Thematic Mapper data sets individually. When the Thematic Mapper data sets are considered together, using site as a block and resolution as a fixed factor, cell width is significant only for training performance and classification variability. As cell width increases, classification variability decreases. As cell width increases, training performance decreases. Trends which。 can be noted, but which are not statistically significant at the sample size used include: as cell width increases, CPU time, field center pixel performance, and "full field" performance decrease. It may be true that a cell width between two and the square root of the average field size divided by two would be superior to those cell sizes tested in terms of CPU time, RMS error, and field center pixel, "full field" and training field performances.

The cell homogeneity parameter is significant for all dependent variables in all Thematic Mapper data sets individually. When the data sets are considered together, the cell homogeneity parameter is significant for CPU time, "full field" performance, training performance and variability. As the cell homogeneity parameter increases, the CPU time and classification variability decreases (1\% significance level). "Full field" and training field performance decrease over the range of homogeneity threshold tested (1\% confidence level). The graph indicates that RMS proportion error increases as the cell homogeneity criteria increases. However, this is not a consistent trend in the eight data sets and is not statistically significant.

TABLE 2.1-10

EFFECT OF ECHO PARAMETERS ON INDIVIDUAL SIMULATED THEMATIC MAPPER DATA SETS ABBREVIATIONS:

| CPU | CENTRAL PROCESSING UNIT (COMPUTER) TIME |
| :--- | :--- |
| FCP | FIELD CENTER PIXEL PERFORMANCE |
| FULL | FULL FIELD PERFORMANCE |
| TRAIN | TRAINING FIELD PERFORMANCE |
| RMS | RMS PROPORTION ERROR |
| VAR | CLASSIFICATION VARIABILITY |
| CELW | CELI WIDTH |
| HOM | CELL HOMOGENEITY THRESHOLD |
| ANN | ANNEXATION THRESHOLD |


|  |  |  | CELW | CELW |
| :---: | :---: | :---: | :---: | :---: |
| CELW HOM | ANN | $X$ | $X$ | HOM |
|  |  | HOM | ANN | ANN |

Williams County, North Dakota - 30 meter resolution

| CPU | .1 | .1 | 99.9 | .1 | 99.9 | 31.8 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| FCP | .1 | .1 | .1 | .1 | .1 | 20.5 |
| FULL | .1 | .1 | .1 | .1 | .1 | 28.7 |
| TRAIN | .1 | .1 | 4.9 | .1 | 2.3 | 99.9 |
| RMS | .1 | 1.0 | 4.4 | .1 | 4.1 | 44.7 |
| VAR | .1 | .1 | .1 | .1 | .1 | 46.3 |

Williams County, North Dakota - 40 meter resolution

| CPU | .1 | .1 | 99.9 | 1.6 | 99.9 | 99.9 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| FCP | .1 | .1 | .1 | .1 | .1 | 99.9 |
| FULL | 1.9 | .1 | 8.9 | .1 | 3.0 | 40.5 |
| TRAIN | .1 | .1 | 99.9 | .1 | 99.9 | 99.9 |
| RMS | .1 | .1 | 99.9 | .1 | 99.9 | 27.1 |
| VAR | .1 | .1 | 99.9 | .5 | 99.9 | 46.0 |

Williams County, North Dakota - 50 meter resolution

| CPU | .1 | .1 | 99.9 | .1 | 99.9 | 99.9 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| FCP | .1 | .1 | 99.9 | .1 | 9.1 | 99.9 |
| FULL | .1 | .1 | 99.9 | .1 | 36.3 | 99.9 |
| TRAIN | .1 | .1 | 5.6 | .1 | 42.7 | 34.2 |
| RMS | .1 | .1 | .1 | .1 | .8 | .1 |
| VAR | .1 | .1 | .1 | .1 | .1 | 15.5 |

TABLE 2.1-10 (Continued)
EFFECT OF ECHO PARAMETERS ON INDIVIDUAL SIMULATED THEMATIC MAPPER DATA SETS CELW HOM ANN

| CELW | CELW | HOM |
| :---: | :---: | :---: |
| $X$ | $X$ | $X$ |
| HOM | ANN | ANN |

Williams County, North Dakota - 60 meter resolution

| CPU | .1 | .1 | 99.9 | .1 | 99.9 | 99.9 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| FCP | .1 | .1 | 19.8 | .1 | 8.7 | 99.9 |
| FULL | .1 | .1 | 2.1 | .1 | .7 | 99.9 |
| TRAIN | .1 | .1 | 14.6 | .1 | 99.9 | 99.9 |
| RMS | .3 | .1 | 99.9 | 9.2 | 26.4 | 99.9 |
| VAR | .1 | .1 | .1 | .1 | .1 | 11.6 |

Finney County, Kansas - 30 meter resolution

| CPU | .1 | .1 | 39.7 | .1 | 15.0 | 3.4 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| FCP | .1 | .1 | .1 | .1 | .1 | 29.4 |
| FULI | .1 | .1 | .1 | .1 | .1 | 38.6 |
| TRAIN | .1 | .1 | .1 | .1 | .1 | 99.9 |
| RMS | .1 | .1 | 16.7 | .1 | 44.4 | 24.9 |
| VAR | .1 | .1 | .1 | .1 | .1 | 37.5 |

Finney County, Kansas - 40 meter resolution

| CPU | .1 | .1 | 10.4 | .1 | 1.5 | 9.4 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| FCP | .1 | .1 | .1 | .1 | .1 | 46.2 |
| FULL | .1 | .1 | .1 | .1 | .1 | 47.4 |
| TRAIN | .1 | .1 | .1 | .1 | .1 | 99.9 |
| RMS | .1 | .1 | 7.8 | .1 | 4.4 | 99.9 |
| VAR | .1 | .1 | .1 | .1 | .1 | 43.4 |

Finney County, Kansas - 50 meter resolution

| CPU | .1 | .1 | 99.9 | .1 | 99.9 | 99.9 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| FCP | .1 | .1 | 9.6 | .1 | 2.7 | 99.9 |
| FULL | .1 | .1 | .3 | .1 | 17.3 | 29.7 |
| TRAIN | .1 | .1 | 7.5 | .1 | 8.7 | 99.9 |
| RMS | .1 | .1 | .1 | .1 | .1 | 2.3 |
| VAR | .1 | .1 | .1 | .1 | .1 | 44.0 |

Finney County, Kansas - 60 meter resolution

| CPU | .1 | .1 | 99.9 | .1 | 99.9 | 99.9 |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| FCP | .1 | .1 | .1 | .1 | .1 | 47.0 |
| FULL | .1 | .1 | .7 | .1 | .1 | 99.9 |
| TRAIN | .1 | .1 | .7 | .1 | .2 | 99.9 |
| RMS | .1 | .1 | 16.9 | .1 | 13.4 | 47.3 |
| VAR | .1 | .1 | .1 | .1 | .1 | 99.9 |

Table 2.1-11
EFFECT OF ECHO ON SIX VARIABLES USING SITE AS A BLOCK AND RESOLUTION AS A FIXED FACTOR
(Eight Thematic Mapper Data Sets)

|  | Ce11 <br> Width | Cell Homogeneity Threshold | Annexation Threshold | $\qquad$ | ```Cell Width X Annexation Threshold``` | ```Cell Homogeneity Threshold X Ammexation Threshold``` |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Degrees of Freedom | 2,2 | 5,5 | 3,3 |  |  |  |
| CPU Time | 10 | 1 |  | . 1 | 99 | 99 |
| Field Center Pixel Performance | 25 |  | 25 | . 1 | 99 | 99 |
| Full Field <br> Performance | 25 | 1 | 25 | . 1 | 99 | 99 |
| Training Performance | 5 | 1 |  | . 1 | 99 | 99 |
| RMS Error |  | 25 | 10 | . 1 | 99 | 99 |
| Classification Variability | 1 | 1 | 1 | . 1 | . 1 | 99 |









The annexation parameter is significant in $63 \%$ of the 48 cases involving 6 dependent variables and 2 sites and 4 resolutions of thematic mapper data. It is never significant for CPU time in the individual data sets. When the Thematic Mapper data sets are considered together, the annexation parameter is signlficant only for variability. As the annexation parameter increases, classification variability decreases.

The interaction of cell width and cell homogeneity is significant for a11 dependent variables both when the Thematic Mapper data sets are considered together or individually.

The interaction of cell width and annexation is also significant in $63 \%$ of the 48 cases in the Thematic Mapper data sets. It is significant only for variability when all the Thematic Mapper data sets are considered together.

The interaction of cell homogeneity and annexation is significant in $8 \%$ of the 48 cases of the individual Thematic Mapper data sets and is not significant for any dependent variable when the Thematic Mapper data sets are considered together.

Comparison of the ECHO and the Perpoint Classifier
F-values for the analysis of variance comparing the effects of classification method (ECHO versus perpoint) and resolution size (30, 40, 50 and 60 meters) to the six variables for the optimum parameter setting of the Supervised ECHO classifier for each of the dependent variables is presented in Table 2.1-12.

Table 2.1-12

|  | Method | Resolution |
| :--- | :---: | :---: |
| CPU Time | $1204.39 *$ | $9.86 * *$ |
| Field Center Pixel Performance | $147.63 * * *$ | 1.25 |
| Full Field Performance | $74.21 * * *$ | 1.47 |
| Training Field Performance | $15.21+$ | .62 |
| RMS Proportion Error | 2.62 | .21 |
| Classification Variability | 13547.44** | $7.23 * * *$ |
| Degrees of Freedom | $(3,3)$ |  |
| $* \quad$ significant at $1 \%$ confidence level |  |  |
| $\# *$ significant at $5 \%$ confidence level |  |  |
| $* * *$ significant at a $10 \%$ confidence level |  |  |
| $+\quad$ significant at a $25 \%$ confidence level |  |  |

Classification variability is reduced as resolution is increased. This result is expected since the field which is 10 pixels wide in the 30 meter data is only 5 pixels wide in the 60 meter data. The CPU time required to perform the classification is increased as the resolution increases. Results based on classification method are hampered by having only $(1,1)$ degrees of freedom. ECHO requires less CPU time and has less classification variability than the perpoint classifier at a $1 \%$ confidence level. ECHO field center pixel and "full field" performance appear to be superior to the perpoint classifier at a $10 \%$ confidence level.

More repetitions are necessary before it can be deomonstrated that ECHO RMS error is significantly less than that of the perpoint classifier for the Thematic Mapper data.
c. Aircraft Results

Tables 2.1-13 and 2.1-14 summarize the analysis of variance results for the effect of the ECHO parameters on the three aircraft data sets taken individually and taken as a block. Figures 35 through 39 present the CPU time, field center pixel performance, "full field" performance, training field performance, and classification variability. RMS values were not calculated because ground truth proportion estimates were not available for all three aircraft sites. The aircraft data sets were sampled only at the 2 by 2 cell size.

As the cell homogeneity parameter increases, field center pixel and training field performance increase ( $5 \%$ confidence levels). As with the LANDSAT and simulated Thematic Mapper data sets, classification variability decreases as annexation and cell selection threshold increase.

The effect of the annexation parameter increases as the cell homogeneity parameter increases for training field performance, classification variability and CPU time.

ECHO - Perpoint Comparison
Though only three aircraft data sets were tested, the Supervised ECHO processor produced results which had significantly less variability than the results produced by the perpoint classifier ( $5 \%$ confidence level). The comparison between the perpoint and the ECHO processor is approaching statistical significance for the other four variables. The following statements can be made for the aircraft data sets at a $15 \%$ confidence level.

* ECHO requires less computer time than the perpoint classifier.
* ECHO has higher field center pixel, "full field" and training performances than the perpoint classifier.

For the aircraft data sets, the cell homogeneity value appears to maximize performance for a value in the neighborhood of 70 and an annexation threshold of 4 seems to be optimal.

Table 2.1-13
INDIVIDUAL AIRCRAFT DATA SETS (Entries are Significance Levels)

Finney County, North Dakota*

|  |  | Cell Homogeneity <br> Threshold |
| :--- | :---: | :---: |
|  |  |  |
| CPU Time | .2 | .1 |
| Field Center Pixel |  |  |
| $\quad$ Performance | .1 | .1 |
| Full Field Performance | .1 | .1 |
| Training Field Performance | .1 | .1 |
| Classification Variability | .1 | .1 |

## Williams County, Kansas*

## Annexation Threshold

Cell Homogeneity Threshold

$$
11.2
$$

. 1
Field Center Pixel
Performance
. 2
Full Field Performance
. 1
.5
.1
Trainıng Field Performance
Classification Variabilıty
. 1

Tippecanoe County, Indiana*

|  |  | Cell Homogeneity |
| :--- | :---: | ---: |

[^2]Table 2.1-14

RESULTS FOR ECHO CLASSIFICATIONS OF AIRCRAFT DATA, USING ALL THREE DATA SETS

|  | Cell <br> Homogeneity | Annexation Threshold | Interaction |
| :---: | :---: | :---: | :---: |
| Degrees of Freedom | 5,10 | 3,6 | 15,30 |
| Field Center Pixel Performance | 3.92 *** | $2.516^{\circ}$ | $1.609^{\circ}$ |
| Full Field Performance | 0.92 | 2.88 | 1. $75 * * *$ |
| Training Field Performance | 6.84* | 4.64*** | 5.47* |
| Classification Variability | 18.80* | 63.33* | 21.78* |
| CPU Time | $1.81{ }^{\circ}$ | $1.78{ }^{\circ}$ | 3.20* |
| * sig $1 \%$ |  |  |  |
| $\begin{aligned} & * * \text { sig } 5 \% \\ & * * * \operatorname{sig} 10 \% \end{aligned}$ |  |  |  |
| - sig $25 \%$ | Table entries are F values.$\qquad$ Sig. level indicated by symbols. |  |  |



| CRAPH OF FCP | FOR 3 | AREALSI |
| :---: | :---: | :---: |
| FL3PIO FLI700 | FLZLIO |  |

CELL SILE IS $<\times 2$


CRAPH OF JHALYING FOR 3 AREA(S)
FL37LO FLITOO FL2IO
EELL SILE $152 \times 2$


```
graph of yarbliy for 3 areatsi FL3710 FillaO FL210
```

EELL SILE IS $2 \times 2$


Figure 39
Aircraft Classification Vaxiability
$\qquad$


## C. Products for LACIE/AI Evaluation

Magnetic computer tapes containing examples of object maps produced by the Nonsupervised field extraction algorithm are provided in Universal format. Object map tapes are multispectral image storage tapes which have had the data values altered for pixels falling within objects identified by the algorithm. The mean response for the pixels comprising the object replaces the individual response for each constituent pixel of the object.

## SUMMARY

The ECHO processor developed at LARS has been implemented in documented FORTRAN programs.

The Supervised ECHO processor has been tested over LANDSAT, simulated Thematic Mapper and aircraft data sets. Evaluations of those tests indicate that:

* ECHO is less expensive than the perpoint classífication algorithm (statistically significant for LANDSAT and simulated Thematic Mapper data sets at a $1 \%$ confidence level and at a $15 \%$ confidence level for the aircraft data sets),
* ECHO provides better field center pixel and "full field" classification performance than the perpoint classifier ( $5 \%$ confidence level on LANDSAT data, $10 \%$ confidence level on simulated Thematic Mapper data, $15 \%$ confidence level on aircraft data).
* ECHO RMS proportion error is significantly less than that of the perpoint classifier at optimal parameter settings for the LANDSAT data. Proportion estimation for the Thematic Mapper data set was not significantly different for the two classification methods.
* ECHO classification results are less variable than the classification results of the perpoint classifier. This result is statistically significant at a $5 \%$ confidence level for all data sets.

The effects of the three ECHO parameters on six variables were monitored. In general, cell homogeneity had the strongest effect on the six variables. Cell width had a strong effect on the variables when the data sets were considered individually. However, when the data sets were considered together, the effect of cell width was not significant for RMS error and was less significant than the cell homogeneity parameter for field center pixel and "full field" performances, indicating the effects of cell width are sometimes opposite in direction for differing data sets. The annexation threshold had a significant effect on classification variability and CPU time but a weak effect on the other variables.

## CONCLUSIONS

ECHO successfully exploits the redundancy of states characteristics of sampled imagery of ground scenes to achieve better classification accuracy, reduce the number of classifications required, and reduce the variability of the classification results. The information required to produce ECHO classifications are cell size, cell homogeneity, and cell-to-field annexation parameters, input data, and a class-conditional marginal density statistics deck.

Future research should be directed towards developing methods for utilizing the information produced by the ECHO processors to aid in the training process. For example, the singular cell map produced by the Supervised ECHO processor (see Figure 40) provides information on training adequacy. Singular cells are represented on this map by $0^{\prime} s$. For the Supervised processor, cells are identified as being singular when the likelihood of the cell belonging to the most likely of the available classes falls below a threshold. Therefore, when a cell is categorized as "singular", it is either because it contains pixels from more than one class, or because the spectral class of the field is not represented in the available class statistics. Large groups of contiguous singular cells will occur when one or more spectral classes have been omitted. The singular cell map may indicate where additional training statistics should be collected.

The intermediate tape produced by the Nonsupervised ECHO processor contains the statistics for each homogeneous object identified by the field extraction routine. It is possible that this spectral-spatial cluster will contain all the information an analyst will need to produce a class statistics deck superior to what would be produced by traditional training procedures for a smaller investment of analyst time than is now required.


Figure 40
Example of the Supervised ECHO Singular Cell Map

Lastly, background research should be performed to determine what training techniques are optimal for a field classifier. ECHO performance may be enhanced by utilizing training techniques developed specifically for it, rather than techniques which are optimal for the perpoint classifier.

1. Kettig, R. L. 1975.

Computer Classification of Remotely Sensed Multispectral Image
Data by Extraction and Classification of Homogeneous Objects.
Ph.D. Thesis, School of Electrical Engineering, Purdue University, West Lafayette, Indiana.
2. Davis, W. A. and Peet, F. G. 1977.

A Method of Smoothing Digital Thematic Maps.
Remote Sensing of Environment, Volume 6, pp. 45-49
3. Landgrebe, D. A. 1975.

Final Report, NASA Contract NAS9-14016, June 1, 1974-May 31, 1975. pp. X-1 - X-11.
4. Kettig, R. L. and Landgrebe, D. A. 1973.

Automatic Boundary Finding and Sample Classification of Remotely Sensed Multispectral Data.

LARS Information Note 041773, Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Indiana.
5. Kettig, R. L. and Landgrebe, D. A. 1975.

Classification of Multispectral Image Data by Extraction and Classification of Homogeneous Objects.

Proceedings Symposium on Machine Processing of Remotely Sensed Data Purdue University, IEEE Catalog No. 75CH1009-0-C.
6. Rodd, E. M. 1972.

Closed Boundary Field Selection in Multispectral Digital Images.
IBM Publication No. 320.2420.
7. Gupta, J. N. and Wintz, P. A. 1973.

Closed Boundary Finding, Feature Selection, and Classification Approach
to Multi-Image Modeling,
LARS Information Note 062773, Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Indiana.
8. Robertson, T. V. 1973.

Extraction and Classification of Objects in Multispectral Images. Proceedings of the Conference on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, Indiana, Section 3B, pp. 27-34.
9. Bauer, M. E., Cary, T. K., Davis, B. J., and Swain, P. H. 1975. Crop Identification Technology Assessment for Remote Sensing (CITARS),

Volume VI: Data Processing at the Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Indiana.

National Aeronautics and Space Administration, Houston, Texas.
JSC 09389.

APPENDIX A
TASK 2.1

PROGRAM ABSTRACTS AND LISTINGS FOR TIIE SUPERVISED ECHO PROCESSOR

LARS Program Abstract $\qquad$
MODULE IDENTIFICATION
Module Name:_ SECSUP Function Name:________ SECSUP
Purpose: Supervised ECHO supervisor routine.
System/Language:_CMS/FORTRAN

MODULE ABSTRACT

SECSUP is the supervisor routine for the supervised ECHO classifier.

1. Module Usage

SECSUP
Calling sequence:
CALL SECSUP
SECSUP is the supervisor routine for the supervised ECHO classifier. It is called with no arguments.
2. Internal Description

The program sECSUP basically controls the program sequencing for the many subprograms involved in the supervised ECHO classifier. Two subroutines are called by SECSUP: SECRDR and SECINT.

SECRDR is the control card reader for reading the user input data cards and is called first. SECINT is the initialization routine for the SECHO function and is called to handle utilization of all input/output operations as well as classification processing. For more detailed descriptions of these modules, see their program abstracts.

Two common blocks are used in SECSUP. GLOCOM, the main LARSYS common block and SECCOM, the common block for the supervised ECHO function are included.
3. Input Description

Not applicable
4. Output Description

Two messages are produced and written to unit TYPEWR, the console.

SUPERVISED ECHO FUNCTION REQUESTED
Signifies beginning of the function.

SUPERVISED ECHO FUNCTION COMPLETED Signifies end of the function.
5. Supplemental Information SECCOM
6. Flowchart



LARS Program Abstract $\qquad$
MODULE IDENTIFICATION


MODULE ABSTRACT

SECCOM is the common block for the supervised ECHO classifier.

## PURDUE UNIVERSITY

Laboratory for Applications of Remote Sensing 1220 Potter Drive West Lafayette, Indiana 47906

## SECCOM

| BUFROZ | INTEGER*4 Number of rows of cells in buffer. |
| :---: | :---: |
| CSELEC | REAL*4 Cell selection parameter. |
| $\operatorname{CSET}(3,30)$ | REAL*4 Channel calibration code vector |
|  | from statistics. |
| CSET3 $(3,30)$ | REAL*4 Channel calibration code vector |
|  | from channels card. |
| THRES | REAL* 4 Annexation t'hreshold. |
| CELSIZ | INTEGER* 4 Number of data points in a |
|  | cell. \| |
| CELWTH | INTEGER*4 Number of rows in cell. |
| INFO (16) | INTEGER*4 Array for storing field de- |
|  | scription information and areas to |
|  | be classified. |
| INPUT | INTEGER*4 Unit number for input tape. |
| INTTAP | INTEGER*4 Tape number of intermediate |
|  | tape. |
| INTFIL | INTEGER*4 File number of intermediate |
|  | tape. |
| JPTS | INTEGER*4 Number of cells horizontally |
|  | in data area. |
| LINES | INTEGER* 4 Number of rows of cells in |
|  | area to be processed. |
| NOCLS | INTEGER* 4 Number of classes in statis- |
|  | tics deck. |


| NOFET | INTEGER* 4 Number of channels in statistics deck. |
| :---: | :---: |
| NOFET3 | INTEGER*4 Number of channels to be |
|  | used in classification. |
| NOPOOL | INTEGER* 4 Number of pools to be used |
|  | in classification. |
| NWORD | INTEGER*4 Integer array word counter. |
| NWORD2 | INTEGER*4 Integer word counter. |
| NVR | INTEGER* 4 Number of channels on data |
|  | tape. |
| OUTPUT | INTEGER* 4 Unit number, of output file. |
| POLNAM $(2,60)$ | INTEGER*4 Names of statistics classes. |
| PREFIX (2) | INTEGER* 4 Prefix information used in |
|  | writing results. |
| RQTAPE | INTEGER*4 Tape number of the results |
|  | tape. . |
| RQFILE | INTEGER* 4 File number of the results |
|  | file. |
| SYMCNT | INTEGER*4 Number of user supplied |
|  | symbols. |
| VARS23 | INTEGER* 4 Number of words to store |
|  | half of a triangular covariance matrix |
|  | ```for classes (channels * (channels +``` |
|  | 1)) $/ 2$. |
| CSEL (30) | INTEGER*2 Channel select vector from |
|  | statistics. |


| CSLL 3 (30) | INTEGER*2 Channel select vector from channels card. |
| :---: | :---: |
| FETVEC (30) | INTEGLR*2 Array of channel numbers from statistics deck. |
| rctve $3(30)$ | INTEGLR*2 Array of channel numbers Irom channels card. |
| POLPTR $(2,60)$ | INTEGER*2 Array of classification pools and pooling information. |
| POLSTK (60) | INTEGER*2 Array of stacked class numbers organized by pool request. |
| CDFLAG | LOGICAL*I Logical variable signalling input of statistics on cards in control card deck. |
| CLSSMAP | LOGICAL*I Logical variable requesting production of classification map. |
| $\operatorname{cserct}(3,30)$ | LOGICAL* 1 Logical array indicating storage of calibration values by user. |
| OBJHAP | LOGICAI,*l Logical variable requesting production of singular cell map in phase 2 (annexation). |
| PIIASLI | LOGICAL*I Logical variable requesting inftial cell processing is to be carried out. |
| PHASL2 | LOGICAI,*I Logical variable indicating annexation of cells is to be carried out. |
| POLNMI (C0) | LOGICAL*1 Logacal array signalling that the name for POOL (I) has been stored. |

PRSTAT

SYM (60)
SYMMTX (60)

LOGICAL*l Logical variable indicating request to print statistics information. LOGICAL*l Array of user supplied symbols. LOGICAL*1 Array of 'default symbols.


LARS Program Abstract $\qquad$
MODULE IDENTIFICATION
Module Name: SECRDR Function Name: SECSUP
Purpose: Control card reader for Supervised ECHO Function
System/Language: CMS/FORTRAN
Author:
P. D. Alenduff Date: 3/30/77

Latest Revisor: $\qquad$ Date:

MODULE ABSTRACT
SECRDR is the control card reader for the Supervised ECHO function. After user requests are read, many different checks are made to detect any control card errors. The input or output classification or intermediate tapes are then mounted on the appropriate devices, and an options list is printed to document user selections.

1. Module Usage

SECRDR
Calling sequence:
CALL SECRDR

- SECRDR is called without parameters to read user option selections and detect any errors on specified combinations of options. Upon error detection, either corrected data is requested or the function is terminated following an information message detailing the error. Output and/or intermediate tapes and the multispectral image storage tape are mounted and a list of selected options is produced.


## 2. Internal Description

SECRDR first performs initialization on all variables stored in SECCOM common block to set up any default options and clear any previously selected options. Control card decoding is heavily dependent upon LARSYS control card processing routınes CTLWRD and BCDVAL. CTLWRD is called with a list of acceptable keywords to read input control cards. The number of the matching keyword is returned to SFCRDR and a branch is made to the code that will appropriately decode the remainder of the control card. Other special purpose LARSYS routanes used for control card decoding include CHANEL for decoding the CHANNELS card, and POLSCN for decoding the CLASSES card.

After encountering a DATA or END card, a check is made for a request of at least one channel on the CHANNELS card. The cell width parameter is then checked to be greater than or equal to 2. The cell selection option must also be non-negative, and annexation threshold TIIRES must be zero or positive. If both
an object map and classification map are reques'ted, only an object map is produced. At this point, tape and disk parameters are checked. If OPTIONS INTERMEDIATE (start processing with the intermediate tape) is selected, a tape number must be supplied. If an intermediate tape number is supplied, a file number must also be specified.

Upon specification of OPTIONS INTERMEDIATE, processing begins with the intermediate tape and a results file is produced. If, however, no OPTIONS INTERMEDIATE request is made and both an intermediate tape number and a results location is specıfied, a conflict is recognized and EXECUTION is terminated. If a results location is needed and none is specified, more data is requested. A file number must be specified when a tape location is selected. Specificatıons of both tape and disk options for classification results information causes termination.

After the error checking sequence is finished, tapes are mounted on the appropriate devices by calls to MTAPE. Initialization, if requested, is performed on the results or intermediate tape. A list of the selected options is then produced. Both GLOCOM and SECCOM common blocks are used by SECRDR.

## 3. Input Description

Input data is read from unit READIN in GLOCOM which is elther CRDRDR, the card reader, or KEYBD, the console keyboard, when cards are typed in by the user. This selection is accomplished by the LARSYS system.
4. Output Description

Output is written to both PRNTR, the printer, and TYPEWR,
an object map and classification map are requested, only an object map is produced. At this point, tape and disk parameters are checked. If OPTIONS INTERMEDIATE (start processing with the intermediate tape) is selected, a tape number must be supplied. If an intermediate tape number is supplied, a file number must also be specified.

Upon specification of OPTIONS INTERMEDIATE, processing begins with the intermediate tape and a results file is produced. If, however, no OPTIONS INTERMEDIATE request is made and both an intermediate tape number and a results location is specified, a conflict is recognized and EXECUTION is terminated. If a results location is needed and none is specified, more data is requested. A file number must be specified when a tape location is selected. Specifications of both tape and disk options for classification results information causes termination.

After the error checking sequence is finished, tapes are mounted on the appropriate devices by calls to MTAPE. Initiallzation, if requested, is performed on the results or intermediate tape. A list of the selected options is then produced. Both GLOCOM and SECCOM common blocks are used by SECRDR.

## 3. Input Description

Input data is read from unit READIN in GLOCOM which is either CRDRDR, the card reader, or KEYBD, the console keyboard, when cards are typed in by the user. This selection is accomplished by the LARSYS system.
4. Output Description

Output is written to both PRNTR, the printex, and TYPEWR,
the console devicc. $A$ varıety of error messages is produced by SECRDR, and brıef list follows:

LRROR IN TAPE OR FILE SPECIFICATION-TYPE CORRECT CARD
A non-numeric character was entered as either a tape or file number.

ZITROR IIN PARAALTER VALUE SPFCIFICATION-TYPE CORRECT CARD
A non-numeric character was specified as the cell size, cell selection, or annexation value.

YOU HAV: ENTERED X SYMEOLS. THE MAXIMUM ALLONED IS 60. EXCESS GYIWOLS WILL NOT BE USLD.

The maximum number of symbols that can be stored is 60. Only the first 60 can be used.

CELL SIRE HUST BE GRLATLR THAN OR EQUAL TO THO-DEFAULT OF 2 ASGUIILD-TYPE CORRECT CARD

The CELL SIZE (X) parameter cannot be less than 2. A corrected card must be supplied.

CELL SELECTION PARAMLTER MUST LE GREATER THAY OR LQUAL TO ZEROTYPC CORREC' CARD The CFLL SELECT (X) entry cannot be negative, a new card 1s requested.

ANNEXATION THRUSIOLD MUST BE NON-NEGATIVE. TYPE A CORRECTED ANNLYATIOIN TIIRESHOLD CARD. The THRES specifled is negative. The user is requested to correct the error.

BOTH SIIGULAR CELL AND CLASSIFICATION MAPS RLQUESTED. ONLY ODJECT IAP WILL BE PRODUCED

Only one map can be produced during annexation. If both are requested, only the singular cell map will be printed.

OBIGINAL PAGE IS POOR
NO INTERMEDIATE TAPE SUPPLIED FOR OPTIONS INTERMEDIATE. TYPE IN INTERMEDIATE CARD

When requesting that processing start from an intermediate tape, an intermediate tape number must be specified. INTERMEDIATE FILE OR INITIALIZE MUST BE SPECIFIED. TYPE ADDITIONAL INTERMEDIATE CARD

When producing an intermediate tape, some indication of file number must be given.

BOTH INTERMEDIATE TAPE AND RESULTS LOCATION SPECIFIED WITHOUT OPTIONS INTERMEDIATE-JOB TERMINATED

It is impossible to determine which part of the classification should be performed when this set of options is supplied. Execution is terminated.

EITHER INTERMEDIATE TAPE OR RESULTS LOCATION MUST BE SPECIFIED WITHOUT OPTIONS INTERMEDIATE. TYPE IN ADDITIONAL CARD

Some output file is needed for processing. Additional information is requested.

EITHER RESULTS FILE OR INITIALIZE MUST BE REQUESTED. TYPE IN ADDITIONAL RESULTS CARD

When producing a results tape, some indication of file number must be given.

NO RESULTS DESTINATION SPECIFIED-TYPE IN RESULTS CARD
Either tape and file or disk must be specified when results will be produced. Additional information is requested. BOTH RESULTS TAPE PARAMETERS AND DISK SPECIFIED-FUNCTION TERMINATED

Either tape or disk can be selected for results but not both.

BOTH FILE AND INITIALIZE REQUESTED FOR INTERMEDIATE TAPEFILE REQUEST IGNORED

The intermediate output tape is initialized. Only file 1 can be initialized.

BOTH FILE AND INITIALIZE REQUESTED FOR RESULTS TAPE-FILE REQUEST IGNORED

The results tape is initialized. Only file l can be initialized.

In addition to these messages, the list of options selected is printed on the unit PRNTR.
5. Supplemental Information

Common blocks GLOCOM and SECCOM are used by SECRDR.







| Fortara | level | 20.7 | stcride | OATE $=17129$ | 15018006 |  | PAGE 0007 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| file secror |  |  |  |  |  |  |  |
|  | 4100 | R9File |  |  |  | supat570 |  |
|  |  |  |  |  |  | Supotisi |  |
|  | ¢ int | aneota |  |  |  |  |  |
|  | 4200 | ${ }^{1 / 5}$ | E2) 60 |  |  | Sutiousti |  |
|  |  | \% | -2. ${ }^{\text {PH }}$ |  |  | Stupoticio |  |
|  |  | Cftrin |  |  |  | Stupaitio |  |
|  |  | $\mathrm{CO}_{60} \mathrm{CHO}_{5}$ | () |  |  | Ş4potioi |  |
|  | ${ }_{c}^{\text {c opt }}$ |  |  |  |  | Stupoutio |  |
| $\begin{aligned} & 0268 \\ & 0268 \\ & 0270 \\ & 0271 \end{aligned}$ | ${ }_{4} 300$ | IF 1 | InTFll |  |  | Suratis |  |
|  |  | H1 | SELECTED | ING SUPERYISEO |  | Stuptitio |  |
| ${ }^{2} 8273$ |  | 12 |  | conerio |  | Stuch |  |
|  | 3101 | forka | Stits | Caxo deck'l |  |  |  |
| - ${ }^{02}$ | 94302 | Fonk | S ${ }^{\text {S }}$ | T5 MAPI: |  | Stile |  |
|  | 94303 | Tor | Chareck |  |  |  |  |
| 027 | 84304 | Formar | W174t | ssimg of raen*) |  | Swiaitic |  |
| ${ }^{212}$ ¢1 | 94305 | Ifogn | CELT' ${ }^{\text {anNe }}$ | produce class | don resust | 约 |  |
| $\begin{aligned} & 0292 \\ & 028 \\ & 028 \\ & 0288 \\ & 0285 \end{aligned}$ |  |  |  |  |  | క゙upotitio |  |
|  | 94306 |  | tispectrat | rics usen |  |  |  |

LARS Program Abstract $\qquad$
MODULE IDENTIFICATION
Module Name: SECINT Function Name:___ SECSUP
Purpose: Initiator routine for Supervised ECHO classifier.
System/Language: CMS/Fortran
Author: P.D. Alenduff $\quad$ Date: $4 / 27 / 77$
Latest Revisor: $\qquad$ Date: $\qquad$

## MODULE ABSTRACT

SECINT carries out required initialization of the rest of variables used by SECSUP and finishes reading any data cards as well as the statistics to be used with the processor. In addition to the array allocation performed, some option information is printed and a loop is entered that carries out the desired stages of classification for each area to be classified.

1. Module Usage

SECINT
CALL SECIN'
The program SECINT is called with no arguments. Any varıables to be used or changed are contained in common blocks GLOCOM and SECCOM, both required by SECINT. AIl information required to perform the classification is gathered by SLCINT from the appropriate source (ie., statistics and areas to be classified) and array allocation is carried out using variables supplied to common block SECCOM by the control card reader SECRDR. Statistics are reduced by calling REDSAV to save statistics for only classes and channels requested and statistics information is printed by calling SECPRT. Initial records are then written to the output device. A loop is entered and, for each area to be classified, a sequence of tests is performed. If multispectral data is required, the correct run is requested and existance of requested channels is confirmed. Further allocation is performed and SECHOl is called to perform the needed processing for the area.

## 2. Internal Description

The first information handled by SECINT is the statistics file. If processing is starting from an intermediate tape, records 1 and 2 are read from the intermediate tape to reestablish information used in initial processing such as channels selected and calibration used, and classes selected and pooling requested. The statistics file is then read from the
tape and written onto the unit SDATA, the statistics file on disk. The next card is checked for an "end" card and if any area description cards are required, the function is terminated.

Otherwise, CLASSX, the storage disk file for all classification areas for this file, is rewound and processing loops until all areas are stored on disk. Some initial allocation is then performed as array storage is set up for the covariance and mean matrices for the entire statistics file on disk. REDSTA transfers the file into memory. CLSCHK and FETCHK then check for fatal errors in selection of pooling and channels. The addresses of reduced arrays for storing only needed information are then computed and REDSAV reduces the information to that required.

Initial records are then produced and written to OUTPUT. The entire statistics-file from SDATA is then transfered to the output file, and SECPRT is called to print the statistics information, if requested, and to produce the record type 4 with covariance and means matrices of the classes. A loop is then started and executed once for each area to be classified. If processing begins with the data type, the correct input run $1 s$ requested and calibration information is established. If processing begins with an intermediate type, all areas on the tape in this file are used. Field size and the allocation of all other arrays used by SECHOl are calculated. Then a record type 5, area 1 dentification record, is produced and SECHOI is called once for each area to be processed. When finished, the

Iinal record is written to the output device.
3. Input Description

Input cell information for SECINT is read from CPYOUT if an intermediate tape is requested as the starting point of processing. The format is identical to that of a classification results file. SDATA, the statistics file on disk is also read to be copied into the output file. CLASSX, the areas to be classified as stored on disk, is also used to store the list of areas to be processed.

## 4. Output Description

A copy of the statistics file is written to SDATA from the intermediate tape is used as input. CLASSX is written as the list of areas to be classified. The PRNTR is used for messages and output information. The file OUTPUT is the result of processing with the same format as a classification results file.
5. Supplemental Information

This program uses common blocks GLOCOM and SECCOM. See LARSYS Systems Manual for classification results file format.


CALL REDSTA
TO READ
STATS INTO
CORE


RE-ALLOCATE \& ADD NEW ARRAY STORAGE






| $\begin{aligned} & 0024 \\ & 0025 \\ & 0026 \end{aligned}$ | CALL GTSERL(SERIAL) <br> IF (PHASE1) GO TO 200 |
| :---: | :---: |
|  |  |
| 0027 | REAO (INPUT, END $=9000$ ) 1 , $1,1,1,1,1, S E R I A L, M O D F L G$ |
| 0028 0029 |  |
|  |  |
|  |  |
|  |  |
|  | 4.(PCLSTK (1), 1=1,NCCLS4), (CLSDAT (1), 1=3,5), CLS |
| 0030 | 105 GEAD (İPUT, END = 9000)I, I, NOCLS4, NCFET3, NOFLD, NO |
|  |  |
|  |  |
|  |  |
|  |  |
| 0032 | ${ }^{\text {CEE MTH }}=$ (KORKIIT |
| 0033 | JHRES $=$ WORX (2) |
|  | $C$ Reaor to ccoy from input tape to disk |
| 0034 | 110 REHIND SDATA |
| 0035 | PEADIINPUY, END=9000)I,I,CARDI, CRDSEO |
| 0036 | EINFLG ${ }^{\text {\% }}$ O |
| 0037 | IF ICARDI(11) EC. IONES EINFLG m 1 |
| 0039 |  |
| 0040 | 115 REAO IMPUT, EHO=9000) 1, I, CARDI, CROSEO |
| 0641 | 9115 KRITESSOAIA,9125ICARDİCRDSEQ |
| 0043 |  |
| 0044 | G0 to 300 (ras |
|  | C CHECK STATS ON DISK |
| 0045 | 200 CHKFLG $=$ FALSE. |
| 0046 | UNIT = CROROR |
| 0047 | IF GNCT.COFLAGS UNIT = SDATA |
| 0049 | REMUN SOATA. CRDRDR. UR. - NUT-CHXUUT) 6010210 |
| 0050 | REAO(SCATA, 9200 ) |
| O651 |  |
| 0053 | CHKFLGz. TRUE. |
|  | ¢ STAIS ON CARDS OR DISK |
| 0054 | C 210 CALL SIATILNIT, NOCLS, NOFLD, MOFET, NUM, GINFLG) |
|  | C Stats are adw cn disk - transfer to core |
| 0055 | C 300 REWIND CLASSX |
| 0056 | IF (PHASEZ-AND. - NOT. PHASE1) AREAFG $=1$ |
| 0057 |  |
| 0058 |  |
|  | ¢ read area cards |
| 0060 | 310 IF (AREAFG.E9. 11 G0 T0, 350 |
| 0061 | CALL LAREAC ISNFO, CODE, READIN:0,0) |
| 0062 |  |
| 0064 | IF (CHKOUT) Go Io 310 ERPNT373. ${ }^{\circ}$ |
| 0065 0086 | MRITE(CLASSXI INFO |
|  | C CONTINUE PROCESSING |
| 0067 | 350 YARSI2 = MOFET* (MOFET + $1 / 1 / 2$ |
| 0068 | COVARI = |
| 0069 | AVARI COVARI - (VARSIL*NOCLS+1)/2 |

[^3]





LARS Program Abstract $\qquad$
MODULE IDENTIFICATION
Module Name: MTAPE Function Name: FUNCTIQNAI_SUPPOBT Purpose: Mounts and positions results or intermediate tapes

System/Language: CMS/FORTRAN
Author: P. D. Alenduff Date:_ $4 / 8 / 77$
Latest Revisor: $\qquad$ Date: $\qquad$

MODULE ABSTRACT

MTAPE mounts and positions the results or intermediate tape
(or a tape to be used as output for copying results files).

PURDUE UNIVERSITY
Laboratory for Applications of Remote Sensing
1220 Potter Drive
West Lafayette, Indiana 47906

## 1. Module Usage

MTAPE
CALL MTAPE (RQTAPE, RQFILE, MODE)
Input Arguments:
RQTAPE

RQFILE

MODE
I*4 Tape number of requested tape. A tape number of 0 is a request for a scratch tape.

I*4 File number of requested file.
If RQFILE is $=0$, then the tape will be inltialized by writing a record type 1 on the results tape with filetype $=0$.

I*4 Flag indicating usage of MMTAPE. MODE $=-1$ indicates MMTAPE has been called to mount and position a tape to be used for copying results files onto. Mode $=0$ indicates that a results tape is being mounted for reading a results file. In this case, the tape is mounted ring out. Also, if MODE $=0, \operatorname{RQFILE}=0$ is invalid and will cause an error when an attempt is made to write on the tape. MODE $=1$ indicates a tape is being mounted for writing a new results file (or continuing a suspended classification). MODE $=2$ indicates a tape is to be used as an intermediate tape for reading an intermediate results file. The difference
between $M O D E=-1$ and $M O D E=+1$ is the DSRN used for the tape. For MODE $=$ -1 and -2 DSRN is CPYOUT and for MODE $=+1$, DSRN is MAPTAP. (DSRN is MAPTAP for MODE $=0$ ).

Output Arguments:
RQTAPE $I * 4$ When MODE $=0$, set to -1 if requested tape file was full and user decided to use disk for results. Otherwise, remains unchanged.

I*4 When MODE $=1$, set to -1 if requested tape file was full and user decided to use disk for results. Otherwise, sends back current file position of tape.

MTAPE checks the validity of the tape by reading the record type 1 from the tape and verifying the tape and file number as well as checking fro the correct type of file. Any attempt to overwirte an existing file causes MTAPE to ask the user (via the typewriter) if he wishes to overwrite the file, respecify a new results card, or terminate the function. Note, however, that if a request has been made to initialize a tape, no checking is performed on previous contents.
2. Internal Description

See output description.
3. Input Description

The record type 1 of the results tape is read for each
file up to and including the file needed. That is, if file
4 is requested the record type 1 is read from files 1-4.
4. Output Descrıption

The following information messages are issued under the circumstances listed. The term filetype means the filetype code from record type 1 of a results file (the program uses variable CHECK for this number).

I0041 is typed when a tape has been mounted and before MTAPE positions it. This message is not typed when the tape is being initialized or when the correct type number was already mounted.

IOO43 is typed when MODE $= \pm 1$ and file type of the requested file $=0$.

IOO44 is typed when MODE $= \pm 1$ and filetype of the requested file $=1$ and the restart flag from GLOCOM (RESTRT) is not $=1$.

I0045 is typed when the tape is correctly positioned. This is not typed when initializing a tape.

After I0043 and I0044, the user is asked whether he wishes to overwrite the file, respecify a new results card with a new tape and/or file or disk option, or terminate the function. IOl00 is typed to allow entry of the new results card. This occurs when the user requests to respecify the results card.

IOl0l is typed to confirm usage of disk for results and occurs whenever disk is specified on the results card.

The following error messages are typed under the conditions listed.

E361 is written when the tape is being filed forward and a file is encountered with filetype other than zero before the requested file is reached and MODE $=0$.

E362 is written when the circumstances for E361 occurs and MODE $=1$. It is also written when MODE $=1$ and the filetype of the file requested is $=-1$.

E363 is written if the RESTRT flag is $=1$ and the filetype of the requested file is not $=1$.

E364 is written when MODE $=1$ and the filetype of the file requested $=1$.

E365 is written when an EOF is read on the results file. This should never occur with valid results files.

For message texts refer to the User's Manual.
5. Supplemental Information

This section deals with the handing of tapes by MMTAPE.
Input:
If a tape is mounted on the device and it is the incorrect tape number (as noted from the appropriate status words in GLOCOM), TOPRU is called to unload the tape before the correct tape is mounted. If the correct tape is mounted, MTAPE will check for the ring in if MODE $= \pm 1$. If the ring is not in, the tape is unloaded and MOUNT is called to mount the tape with the ring in. If the correct tape is mounted, MTAPE assumes that the file number (as recorded in GLOCOM) is correct and moves the tape backwards or forwards to find the requested file. Output:

The tape is mounted with ring in for MODE $= \pm 1$ and with ring out for $M O D E=0$ and -2 .

The tape is left positioned at the beginning of the requested file. When the tape is initialized a TOPRW is used to do this.






LARS Program Abstract $\qquad$
MODULE IDENTIFICATION
Module Name: SECHOl_ Function Name:__ SECSUP
Purpose: Performs Supervised ECHO Processing on area.
System/Language:
Author: P. D. Alenduff Date: 4/11/77

Latest Revisor: Date: $\qquad$

MODULE ABSTRACT

This routine is called once for each area to be classified and performs all classification on each line of data and produces the output file.

PURDUE UNIVERSITY
Laboratory for Applications of Remote Sensing
1220 Potter Drive
West Lafayette, Indiana 47906

1. Module Usage

SECHOI
CALL SECHO1 (ICLSCV, VECPRD, SCAPRD, DETCOV, CONST, BDATA, RDATA, PIXCOR, PIXVAL, CELCOR, CELSUM, CELIKE, CELCLS, PIXCLS, STACK, OPEN, CLOSED, FLDSIZ, FLDLIK, PRTBUF, BUFFER)

Input Arguments:
ICLSCV
(VARSZ3, NOPOOL)
VECPRD
(NOFET3, NOPOOL)

SCAPRD (NOPOOL)

DETCOV (NOPOOL)

CONST (NOPOOL)

R*4 An array dimensioned (channels * (channels +1 )/2) * number of classes, containing the upper half of the triangular covariance matrices for each of the classes used for classification.

R*4 An array dimensioned channels * classes containing the vector product of the classes for each channel from the inverse covariance matrix and the determinant of the covariance matrix.

R*4 An array dimensioned "classes" for the scalar product of the mean and covariance matrices for each class. R*4 An array dimensioned "classes" words for the determinant of the covariance matrix for each class. R*4 An array dimensioned "classes" words containing the constant term used consisting of an expression with

|  | determinant and scalar product for each class. |
| :---: | :---: |
| BDATA <br> ((NOFET3*ID(6))/2) | I*2 The array used for reading the byte information of data values from data tape. |
| RDATA <br> (NVR, CELWTH) | R*4 The array dimensioned "(ID(6)-6) <br> * cell width" words for reading actual data values from tape and calibrating them. |
| $\begin{aligned} & \text { PIXCOR } \\ & \text { (VARSZ } 3 \text { * CELSIZ) } \end{aligned}$ | R*4 The array dimensioned (channels * $($ channels +1$) / 2$ ) by (cell width) ${ }^{2}$ words for storing the cross product terms of the data values of a cell (correlation term). |
| PIXVAL <br> (NOFET3 *CELSIZ) | R*4 An array dimensioned "channels * (cell width) ${ }^{2 n}$ for storing actual pixel data values for each channel for 1 cell. |
| CELCOR (VARSZ3) | R*4 An array dimensioned (channels * <br> (channels + 1))/2 words for saving <br> the cross product of the pixel values. |
| CELSUM (NOFET3) | R*4 An array dimensioned "channels" words used to store the sum of data values for each pixel in the cell for a given channel number. |
| CELIKE <br> (NWORD, JPTS) | R*4 An array dimensioned "NWORD by columns of cells" for storing the cell likelihoods or pixel values for each |


FLDLIK
(NOPOOL, JPTS)
PRTBUF (JPTS)

| BUFFER |
| :--- |
| (BUFROZ * JPTS) |

- R*4 The array of likelihoods dimensioned columns of cells by classes words used to store the likelihood of each field for each class. I*2 The output print buffer dimensioned "columns of cells" halfwords. I*2 The buffer array used for annexation of neighboring rows of cells dimensioned BUFROZ * JPTS.

SECHO1 performs processing of an area to be classified. If only an intermediate tape is to be produced, cell information is gathered and written to tape. Otherwise all information is used in annexation and classification, and the classification results are written to the output device. Any maps are also produced.

## 2. Internal Description

Initially, column headers are printed if any map is to be produced. All buffer pointers and cell arrays are initialized so that the arrays OPEN and CLOSE indicate that no annexation and classification has taken place. Then a loop is entered and executed once for each row of cells (two or more data lines). The routine FILBUF is called to gather statistics on one row of cells, either directly from processing the data tape, or from reading the intermediate tape. If no annexation is to be performed, the information is written to the output device and this loop is executed until this entire area is completed and a return

1s executed. Otherwise, a series of calls to LIKRAT for a comparison of the likelihood of the cell being annexed to the cell above, left, and right is performed. The cell is annexed to the most likely cell if the comparison with the annexation threshold is successful. This is continued for all cells in the row. When finished, any fields formed from adjoining cells that are closed and no longer continued are classified and then any new fields are opened in that statistics from the initial cell are copied into the array storing field statistics. Then LIKRAT is called to classify closed areas. A print buffer is then filled with symbols for either the cell map or classification amp and printed out. Statistics are handled for remaining open fields and the row is written to the output tape. This is continued until the area has been processed.

## 3. Input Description

Not applicable

## 4. Output Description

The unit PRNTR receives the map and column header produced if a map is to be printed.

The unit OUTPUT is used as the output unit for either intermediate tape - cell information or classified cell lines. These lines follow the format of regular classification records.
5. Supplemental Information

This program uses common blocks GLOCOM and SECCOM.
6. FLOWCHART









## LARS Program Abstract

$\qquad$
MODULE IDENTIFICATION
Module Name: GATHER . Function Name: SUPPORT
Purpose: Gather cell statistics information for processing.
System/Language:_CMS/FORTRAN
Author: P. D. Alenduff_ Date:_4/12/77

Latest Revisor: $\qquad$ Date: $\qquad$

MODULE ABSTRACT
GATHER computes information about cell statistics for actual classification by SECHO.

## 1. Module Usage

## GATHER

CALL GATHER (RDATA, PIXVAL, PIXCOR, CELSUM, CELCOR, CELWTH, VECSIZ, MTXSIZ, NSR, NVR)

## Input Arguments:

RDATA (NSR, NVR) $\quad$ ** Calibrated data values from data tape with NSR samples and NVR channels. I*4 Number of data points in width of cell.

I*4 Number of channels to be used in classification.

I*4 Number of entries in the triangular covariance matrix. Equal to (channels * (channels + 1))/2.

I*4 Number of data samples per line in RDATA.

I*4 Number of channels of data in RDATA array.

Output Arguments:

PIXVAL (NOFET3, CELWTH*CELWTH)

PIXCOR (MTXSIZ, CELWTH*CELWTH)

CELSUM (VECSIZ)

R*4 An array of actual data values in the cell dimensioned channels by (cell width) ${ }^{2}$.

R*4 Cross product terms for all data points in a cell, dimensioned (cell width) ${ }^{2}$ by MTXSIZ.

R*4 Sum of data values in the cell for each channel of data (dimensioned channels).

## CELCOR (MTXSIZ)

R*4 Correlation matrix for data values in cell consisting of a sum of products of data values in cell.

GATHER is called to return information concerning the data characteristics of one cell of data.
2. Internal Description

GATHER loops for each channel of data and each pixel within the cell to produce the data array returned.
3. Input Description

Not applicable
4. Output Description

Not applicable
5. Supplemental Information

Not applicable

## 6. Flowchart




LARS Program Abstract $\qquad$
MODULE IDENTIFICATION
Module Name: FILBUF
Function Name: SECSUP
Purpose: Retrieve cell statistics
System/Language: CMS/FORTRAN
Author: P. D. Alenduff Date: 4/18/77
Latest Revisor: $\qquad$ Date: $\qquad$

MODULE ABSTRACT

FILBUF returns a set of buffers with cell information from either direct processing from the data tape or reading from the intermediate tape.

## 1. Module Usage

FILBUF
CALL FILBUF (ICLSCV, VECPRD, SCAPRD, SETCOV, CONST, BDATA, RDATA, PIXCOR, PIXVAL, CELCOR, CELSUM, CELIKE, CELCLS, PIXCLS, CURLIN, *)

Input Arguments:

ICLSCV
(VARSZ3, NOPOOL)

VECPRD
(NOFET3, NOPOOL)

SCAPRD (NOPOOL)

DETCOV (NOPOOL)

CONST (NOPOOL)

BDATA
(NVR, ID(6))

RDATA
(NVR, CELWTH)

R*4 Class covariance matrices from statistics dimensioned classes by (channels * (channels + 1))/2. R*4 Vector product matrix that is inverse of class mean matrix transposed times the covariance matrix for the class, dimensioned channels by classes.

R*4 Scalar product of the classes for
all statistics classes.
R*4 Determinant of the covariance matrix for each class.

R*4 Constant array computed from in (2*PI*class covariance matrix) for each class.

L*l Buffer array used by GADLIN for' reading data tape dimensioned channels on tape by samples per line on tape. R*4 Array used to store calibrated data values from tape for one row of cells (multiple data lines).
PIXCOR
(VARSZ $3, \mathrm{CELWTH}^{2}$ )

PIXVAL
(NOFET3, CELWTH $^{2}$ )

CELCOR (VARSZ3)

## Output Arguments:

CELSUM (NOFET3)

CELIKE
(NWORD, JPTS)

CELCLS (JPTS)

PIXCLS
(NWORDS, JPTS)

CURLIN

R*4 Array used to store cross product data values in one cell for a correlation type indication.

R*4 Array used to store pixel data values for a cell.
$R * 4$ Sum of the PIXCOR values for a cell being processed.

R*4 The sum of all data points in a cell for each channel.

R*4 Likelihood values for each class for each cell in a row of cells. I*2 Class numbers for each cell in a row of cells.

I*2 Class numbers of pixels for all cells in a row of cells.

I*4 Current line number requested.

FILBUF is called to fill up CELCLS and CELIKE for annexation or writing to the intermediate tape. By either reading from the intermediate tape for annexation processing, or reading from the data tape and carrying out processing for the area, cell information can be produced.

## 2. Internal Description

FILBUF produces arrays CELCLS and CELIKE from either reading the intermediate tape that is in a classification format or from reading the input data tape via GADLIN. The raw information is then processed by producing cell statistics with GATHER,
establishing cell classes with SAMCLS, testing for cell homogeneity, and outputing a row of cell information. This routine uses common blocks GLOCOM and SECCOM.
3. Input Description

The only input is possible input from the intermediate tape which stores the same format used by classification tapes and stores cell statistics for later annexation processing.
4. Output Description

Not applicable.
5. Supplemental Information

Not applicable.



## K HDYUDUGIBILITY OR THE ORIGINAL. PAG居 IS POOR




LARS Program Abstract $\qquad$
MODULE IDENTIFICATION
Module Name: STATS 2 Function Name: SECSUP
Purpose: Carry out annexation tests and processing
System/Language: CMS/FORTRAN
Author: P. D. Alenduff Date: 4/20/77
Latest Revisor: $\qquad$ Date: $\qquad$

MODULE ABSTRACT

STATS2 contains entry points LIKRAT, ADD2, and BEGFID. LIKRAT produces a likelihood test by subtracting log - likelihoods and returns a resultant likelihood value. ADD2 adds a cell to the statistics of a field. BEGFLD initializes field statistics from a fields statistics.
1.' Module Usage
LIKRAT
CALL LIKRAT (STAT, CELIKE, FLDLIK, AUXLIK, CELCLS, FLDCLS,
AUXCLS:
Input Arguments:
CELIKE (NOPOOL) R*4 Log likelihood values for this cell
for each class in the statistics.
R*4 Log likelihood values of all cells
annexed to the field for each class
in the statistics.
I*2 Class number of cell being con-
sidered.
I*2 Class number of field being tested.
Output Arguments:
AUXLIK (NOPOOL) R*4 Resultant log likelihood of an-
nexation of cell to given field for
each class.
AUXCLS I*2 Resultant class of annexation
test for this cell.
R*4 Resulting statistic concerning
annexation performance for this cell
to the field being considered.
LIKRAT produces a resulting likelihood, class and sta-
tistic if the given cell is annexed to the supplied field.
ADD 2
CALL ADD2 (FLDLIK, AUXLIK, FLDCLS, AUXCLS, FLDSIZ)

Input Arguments:
AUXLIK (NOPOOL) R*4 Likelihood values formed with temporary annexation of cell to field.

AUXCLS I*2 Class number of cell when annexation of cell to field was successful. I*4 Number of pixels included in the given field.

Output Arguments:
FLDLIK (NOPOOL) R*4 Likelihood values of new fiéld formed.

I*2 Class of field formed.
I*4 (see above).
ADD2 adds the cell supplied to the field information supplied. Temporary likelihood values are assigned as new likelihoods for field.

BEGFLD
CALL BEGFLD (CELCLS, FLDCLS, CELIKE, FLDLIK, FLDSIZ)
Input Arguments:
CELCLS I*2 See above.
CELIKE (NOPOOL) R*4 See above.
Output Arguments:
FLDCLS I*2 See above.
FLDLIK (NOPOOL) R*4 See above.
FLDSIZ I*4 See above.
BEGFLD initialize FLDCLS, FLDLIK, and FLDSIZ to be equal to CELCLS, CELIKE, and CELSIZ.
2. Internal Description

See above.
3. Input Description

Not applicable
4. Output Description

Not applicable
5. Supplemental Information

Not applicable
6. Flowchart

LIKRAT


ADD2




[^4]$\qquad$
MODULE IDENTIFICATION
Module Name: SAMCLS Function Name: ..... SECSUP
Purpose: Classify cell using maximum likelihood
System/Language: ..... CMS/FORTRAN
Author: P. D. Alenduff ..... Date: 4/14/77Latest Revisor:
$\qquad$ Date: $\qquad$ 1

MODULE ABSTRACT

SAMCLS performs a maximum likelihood classification by calling LOGLIK and finđing the maximum likelihood class for a cell.

## 1. Module Usage

## SAMCIS

CALL SAMCLS (CLASS, CHISQR, LIKELY, COVMTX, VECPRD, SCAPRD, CONST, SAMCOR, SAMSUM, SAMPTS, NOFET3, MTXSIZ, NOPOOL, RETLIK)

## Input Arguments:

COVMTX
(VARSZ3, NOPOOL)

VECPRD (NOFET3, NOPOOL)

SCAPRD (NOPOOI)

CONST (NOPOOL)

SAMSUM (NOFET3)

SAMPTS

NOFET3

MTXSIZ

R*4 An array dimensioned (channels * (channels +1$) / 2$ ) * number of classes, containing the upper half of the triangular covariance matrices for each of the classes used for classification.

R*4 An array dimensioned channels * classes containing the vector product of the classes for each channel from the inverse covariance matrix and the determinant of the covariance matrix.
$R^{*} 4$ An array dimensioned classes for the scalar product of the mean and covariance matrices for each class. R*4 An array dimensioned "classes" words containing the constant term used consisting of an expression. R*4 Sum of data values for all points in cell for given channel.

I*4 Number of points per cell.
I*4 Number of channels.
I*4 Number of words needed to store
half of triangular covariance matrix.
NOPOOL
RETLIK
I*4 Number of classification pools. L*4 Indication of whether or not likelihoods are to be returned in LIKELY.

Output Arguments:

| CLASS | $I * 4$ Class number of cell. |
| :--- | :--- |
| CHISQR | $R * 4$ Resultant Chi-square value from |
| LIKELY (NOPOOL) | LOGLIK and classification. |
|  | $R * 4$ Log likelihood of cell belonging |
|  | to pools. |

## 2. Internal Description

SAMCLS Calls LOGLIK once for each class and is returned a $\log$ likelihood for the cell for each class. If the likelihood is greater than the previously stored, the new likelihood is saved with the class number and, when done, this class and value are returned.
3. Input Description

Not applicable.
4. Output Description

Not applicable.
5. Supplemental Information

Not applicable.
6. Flowchart



LARS Program Abstract $\qquad$
MODULE IDENTIFICATION
Module Name: LOGLIK Function Name:_SECSUP
Purpose: Produce Log likelihood of a cell
System/Language: CMS/FORTRAN
Author: P. D. Alenduff Date: 4/13/77
Latest Revisor: Date: $\qquad$

MODULE ABSTRACT

LOGLIK produces a log likelihood for a class in the input statistics.

## 1. Module Usage

LOGLIK
CALL LOGLIK (LIKELY, CHISQR, ICLSCV, VECPRD, SCAPRD, CONST, DIMEN, SAMCOR, SAMSUM, SAMPTS)

Input Arguments:

ICLSCV (MTXSIZ)

VECPRD (DIMEN)

SCAPRD

CONST
DIMEN

SAMCOR (MTXSIZ)

SAMSUM (DIMEN)

SAMPTS

Output Arguments:
LIKELY
CHISQR

R*4 Class covariance matrix from statistics, dimensioned channels * (channels +1 )/2. R*4 Vector product of class covariance matrix determinant and means dimensioned "channels" words.

R*4 The scalar product term for the class being considered.

R*4 The constant term derived in SECPRT.
I*4 Number of channels of data in cell information.

R*4 The cross product terms from the cells data values dimensioned as ICLSCV.

R*4 The sum of the data values for each channel for the cell.

I*4 Number of data points in a cell, actually $(\text { cell width })^{2}$.

R*4 Log likelihood value for the cell. R*4 Chi-square value returned for the cell.

LOGLIK is called to return a $\log$ likelihood and a Chisquare value for the cell being considered.
2. Internal Description

The Chi-square value is calculated by the equation: $C H I S Q R=\operatorname{tr}\left(C_{j}{ }^{-1} \sum_{i}^{m} Y_{i} Y_{i}{ }^{\prime}\right)-2 M_{j} C_{j}{ }^{-1} \underset{i}{m} Y_{i}+m M_{j}^{\prime} C_{j}^{-1} M_{j}$ and the $\log$ likelihood is calculated by:

LIKELY $=-.5\left(\mathrm{~m} \cdot \ln \left|2 \pi C_{j}\right|+\right.$ CHISQR $)$
Where:
$C_{j}$ is the class covariance matrix for class $j$.
$M_{j}$ is the mean vector for class $j$.
$Y_{i}$ is the channel response vector for pixel $i$ of the cell.
$m$ is the number of pixels in the cell.
3. Input Description

Not applicable.
4. Output Description

Not applicable.
5. Supplemental Information

Not applicable.


```
FDRTRAN IV G levEl 20.7


LARS Program Abstract \(\qquad\)
MODULE IDENTIFICATION
Module Name: SECRD Function Name: SECSUP
Purpose: Read data arrays from intermediate tape.
System/Language: CMS/Fortran
Author: P. D. Alenduff Date: 4/12/77
Latest Revisor: Date: \(\qquad\)

MODULE ABSTRACT

SECRD reads data arrays from the intermediate tape for annexation processing by SECHO.

\author{
PURDUE UNIVERSITY \\ Laboratory for Applications of Remote Sensing \\ 1220 Potter Drive \\ West Lafayette, Indiana 47906
}
1. Module Usage

SECRD
CALL SECRD (COVMTX, AVEMTX)
Input Arguments:

COVMTX
(VARSZ3, NOPOOL)

AVEMTX
(NOPET3, NOPOOL)

R*4 An array dimensioned (channels
* (channels + 1))/2 fullwords for storing the covariance matrices of classes used in classification. R*4 An array dimensioned channels by classes words for storing the class means of classes used in the classification.

Output Arguments:
Same as above.
SECRD is called to read one reocrd from the intermediate tape. COVMTX and AVEMTX are read in an unformatted fashion from the unit INPUT. This program uses common blocks GLOCOM and SECCOM.
2. Internal Description

See above.
3. Input Description

SECRD reads from unit INPUT a record type 4 from the classification format file.
4. Output Description

Not applicable.
```

SECRD-3

```
5. Supplemental Information

See LARSYS Systems Manual for format of classification results file.

\section*{6. Flowchart}



LARS Program Abstract \(\qquad\)
MODULE IDENTIFICATION
Module Name: SECPRT Function Name:_ SECSUP
Purpose: Statistics printing and output routine for SECHO
System/Language: CMS/FORTRAN
Author: P. D. Alenduff Date: 4/12/77
Latest Revisor:
Date: \(\qquad\)

MODULE ABSTRACT
SECPRT produces tabular statistics for user information as well as carrying out minor data array processing.
each channel. Then the information concerning classes used is printed and WRTMTX is called to print the correlation matrix. The determinant of the covariance matrices is formed by a call to SAMINV, and a check is made to determine if any determinants were zero or negative. Then SMMULT is called to produce the array SCAPRD, the scalar product, and a calculation is performed to produce the constant array. Execution is terminated if any of the determinants were zero or negative. This routine uses common blocks GLOCOM and SECCOM.
3. Input Description

Not applicable.
4. Output Description

A record type 4 of the classification format output file is written to unit ourpur.

Also, statistics information is produced on unit PRNTR.
5. Supplemental Information

Not applicable.

\section*{1.}

Module Usage
SECPRT
CALI SECPRT (COVMTX, AVEMTX, SCAPRD, DETCOV, CONST, WORK)
Input Arguments:
COVMTX
(VARSZ 3, NOPOOL)

AVEMTX
(NOFET3, NOPOOL)
Output Arguments:
SCAPRD (NOPOOL) R*4 Scalar product for each class.
DETCOV (NOPOOL)

CONST (NOPOOL)

WORK (NOFET3 * (NOFET3 + 2))

\author{
R*4 Covariance matrix for each class. \\ R*4 Mean matrix for each class. \\ R*4 Determinant of the covariance matrix for each class. R*4 Constant formed for computation in classification for each class. \\ R*4 Buffer array for inverting covariance matrix at least channels by channels +2 words.
}

SECPRT is called to print the statistics information to be used for classification, if requested, and produce the record type 4 of a classification results format file for SECHO. Also the determinant, scalar product, and constant arrays, DETCOV, SCAPRD, and CONST, are produced for later usage.
2. Internal Description

Header information is printed on the Unit PRNTR followed by the writing of record type 4 on the unit OUTPUT. This classification format record contains COVMTX and AVEMTX, the covariance and mean matrices used in the classification for
6. Flowchart



```


[^0]:    *ECHO stands for Extraction and Classification of Homogeneous Objects

[^1]:    Entries are Significance Levels

[^2]:    * Degrees of freedom for the Annexation Threshold test are 3,15 and Degrees of freedom for the Cell Homogeneity Threshold are 5,15.

[^3]:    

[^4]:    

