LARS Contract Report 093078

FOREST RESOURCE INFORMATION SYSTEM

Quarterly Report

for the period

1 July 1978 to 30 September 1978

Prepared for

NATIONAL AERONAUTICS and SPACE ADMINISTRATION

Johnson Space Center
Earth Observations Division
Houston, Texas 77058

Contract: NAS9-15325
Technical Monitor: R. E. Joosten/SF5

Submitted by:

The Laboratory for Applications of Remote Sensing
Purdue University
West Lafayette, Indiana 47906

Principal Investigator: R.P. Mroczynski

(E79-10175) FOREST RESOURCE INFORMATION SYSTEM Quarterly Report, 1 Jul. - 30 Sep. 1978 (Purdue Univ.) 53 p HC A04/MF A01 CSCL 02F
N79-20446
Unclas
G3/43 00175

FOREST RESOURCE INFORMATION SYSTEM

Quarterly Report

for the period

1 July 1978 to 30 September 1978

Prepared for

NATIONAL AERONAUTICS and SPACE ADMINISTRATION

Johnson Space Center
Earth Observations Division
Houston, Texas 77058

Contract: NAS9-15325
Technical Monitor: R. E. Joosten/SF5

Submitted by:

The Laboratory for Applications of Remote Sensing
Purdue University
West Lafayette, Indiana 47906

Principal Investigator: R.P. Mroczynski

Index

FRIS PROJECT SUMMARY

The Forest Resource Information System Project (FRIS) is a cooperative effort between the National Aeronautics and Space Administration (NASA) and St. Regis Paper Co.(STR). Purdue University's Laboratory for Applications of Remote Sensing (LARS), under contract to NASA, will supply technical support to the project.

FRIS is an Application System Verification and Transfer (ASVT) Project funded by NASA. The project is interdisciplinary in nature involving experties from both the public and private sectors. FRIS also represents the first ASVT to involve a large broad base forest industry (STR) in a cooperative with the government and the academic communities.

Purpose

The goal of FRIS is to demonstrate the feasibility of using computer-aided analysis of Landsat Multispectral Scanner Data to broaden and improve the existing STR Forest data base. The successful demonstration of this technology during the first half of the project will lead to the establishment by STR of an independently controled operational forest resource information system in which Landsat data is expected to make a significant contribution. FRIS can be viewed by the user community as a model of NASA's involvement in practical application and effective use of space technology.

Additionally, FRIS will serve to demonstrate the capability of Landsat MSS data and machine-assisted analysis technology to private industry by:

- Determining economic potentials,
- Providing visibility and documentation, and
- The ability to provide timely information and thus serve management needs,

The ultimate long term successfullness of FRIS be measured through future development of remote sensing technology within the forest products industry.

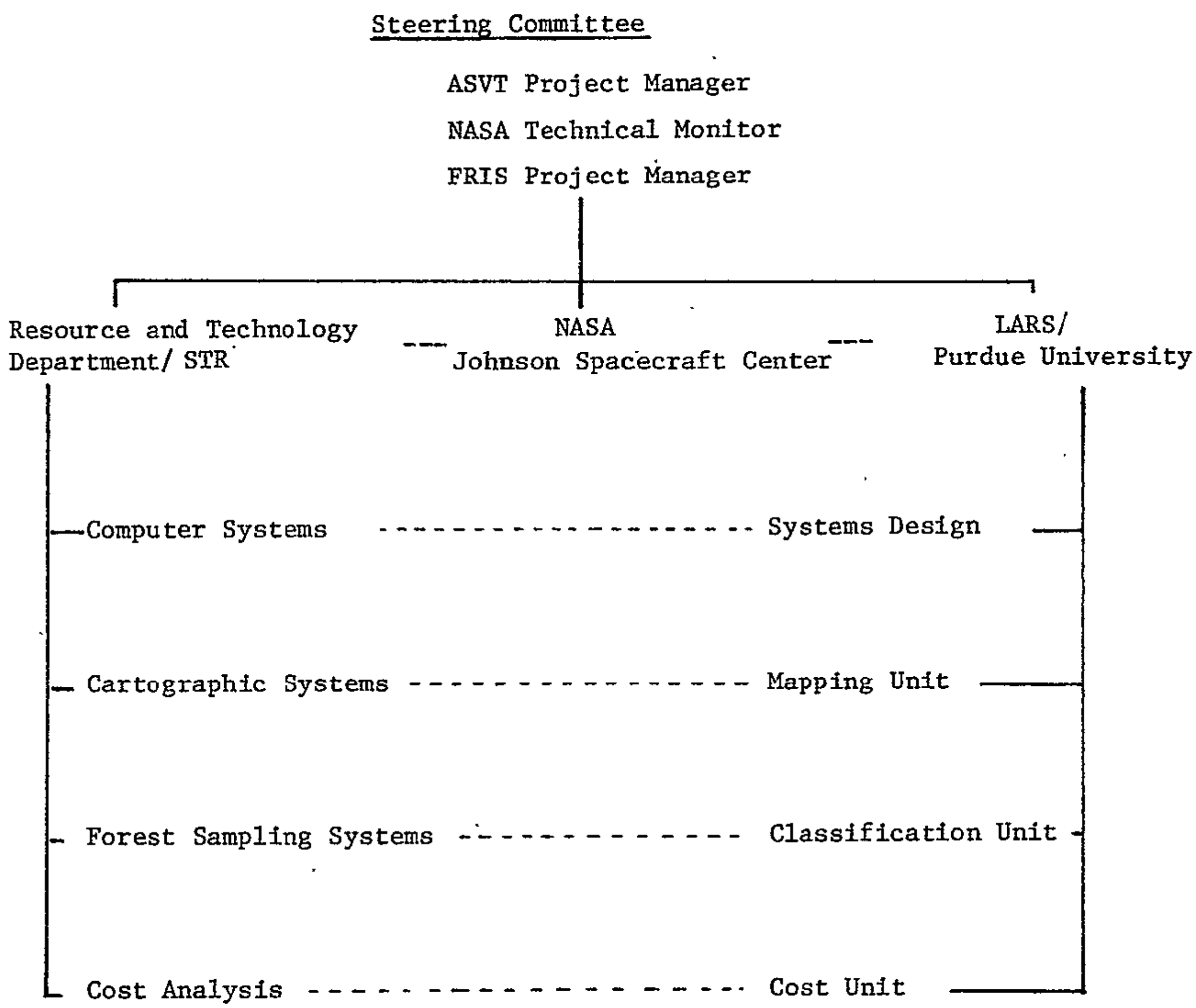
Scope

FRIS is funded as a modular or phasedproject with an anticipated duration of three years. The original project concepts were developed in 1973, and a formal project plan was submitted to NASA by STR in 1976. The project offically began in October 1977 after the signing of a cooperative agreement between NASA and STR; and after the completion of contractual arrangements with Purdue University.

Organization

The organization of FRIS is depicted in the chart that follows. Since FRIS is a cooperative involving three independent agencies, a steering committee consisting of a project manager from each institution was formed to provide for overall guidance and coordination. Operationally, both STR and LARS have project managers and project staff to insure for the timely completion of activities within the project. The NASA technical coordinator monitors project activities and provides a liasion between the STR and LARS staffs. The solid lines on the chart indicate the flow of management responsibility. The dash lines reflect the technical and scientific interchanges between operating units.

FRIS Organization



1.0 Introduction

The materials presented in this report document FRIS project staff activities for the fourth project quarter. The fourth quarter encompasses the calendar period beginning 1 July 1978 and ending 30 September 1978. This period also marks the termination of the first full year of the Forest Resource Information System (FRIS) ASVT, and the third quarter of the Phase II demonstration. The working objective of Phase II is:

> To provide St. Regis Paper Company (STR), through a demonstration of computer-aided Landsat analysis, information concerning the ecomomic feasibility and practical applicability of remote sensing technology for forest inventory.

Technical activities during this quarter occurred under one of the four Technical Working Units. They are:

1. Classification Evaluation
2. Mapping and Digitizing
3. System Design
4. Cost Evaluations

Accomplishments during this quarter included:

- Evaluation of FRIS Landsat classification performance.
- Sizeable modification to the classification evaluation technique.
- Definition of the FRIS preprocessing techniques.
- Identification of data base software which may be utilized as a component to an operational FRIS.
- Identification of the configuration for a Jacksonville remote terminal link to LARS.
- Identification of the components for the FRIS preliminary System.

- Definition of a potential framework for FRIS cost evaluations.

2.0 Working Unit Activities

All Working Units actively pursued their pre-defined goals. Generally, this Phase of the project remains on its defined timelines. The only note-able exception is in the area of data preparation. This task has lagged behind because of the magnitude and complexity of overlaying vector and gridded data sets.

The following sections contain discussions of activities conducted by each FRIS Working Unit during this quarter.

2.1 Classification Unit

2.1.1 Classification Procedure

The primary objective of this activity is to provide a demonstration of the utility of computer-aided Landsat classification techniques to industrial forest resource management. To accomplish this goal, four Test Areas have been identified from approximately 680,000 hectares (1.7 million acres) of St. Regis controlled lands in the southeast. Each area will be classified with a set of procedures that were developed during the early stages of Phase II. Through the use of pre-defined classification procedures, we will in effect have replications of classification results for four physiographic sub-provinces in which the St. Regis Paper Company controls land. Evaluation of the performance of these classification replicates will provide the project staff information needed to assess the operational feasibility of computer-aided Landsat analysis to St. Regis forest management operations in the southeast.

In order to insure that only variations in test area differences due to sub-province location and not variations in classification approach would occur, a uniform set of classification procedures were developed. A schematic of this approach is illustrated in Figure 1. The sub-routines (identified in the text by utilized *NAME) all currently exist as part of the documentation for LARSYS Version 3.1 or LARSYS DV, the image processing systems developed and used at Purdue. In its current configuration this approach is interactive, which has been a valuable asset to the technology transfer activity. For an operational application the procedures would be streamlined, and where feasible the programs optimized for the computer in which they would reside.

As a point of departure in developing an operational classification procedure for FRIS, we have identified, in outline form, a procedure for the computer-aided analysis. An early iteration of this procedure follows.

A. Data Set Generation

1. Define permanent training units. These should:
    a. be large and diverse enough to include the range of expected spectral classes; viz covertypes, within the tract.
    b. be geographically representative of the tract.
    c. represent a cross-sectional profile of the tract, both in terms of geology and vegetation.
    d. endeavor to include entire Administrative Units or similar geographically referenced areas.
    e. at the scale of the source maps, be flexible to allow for partial area replacement if required.

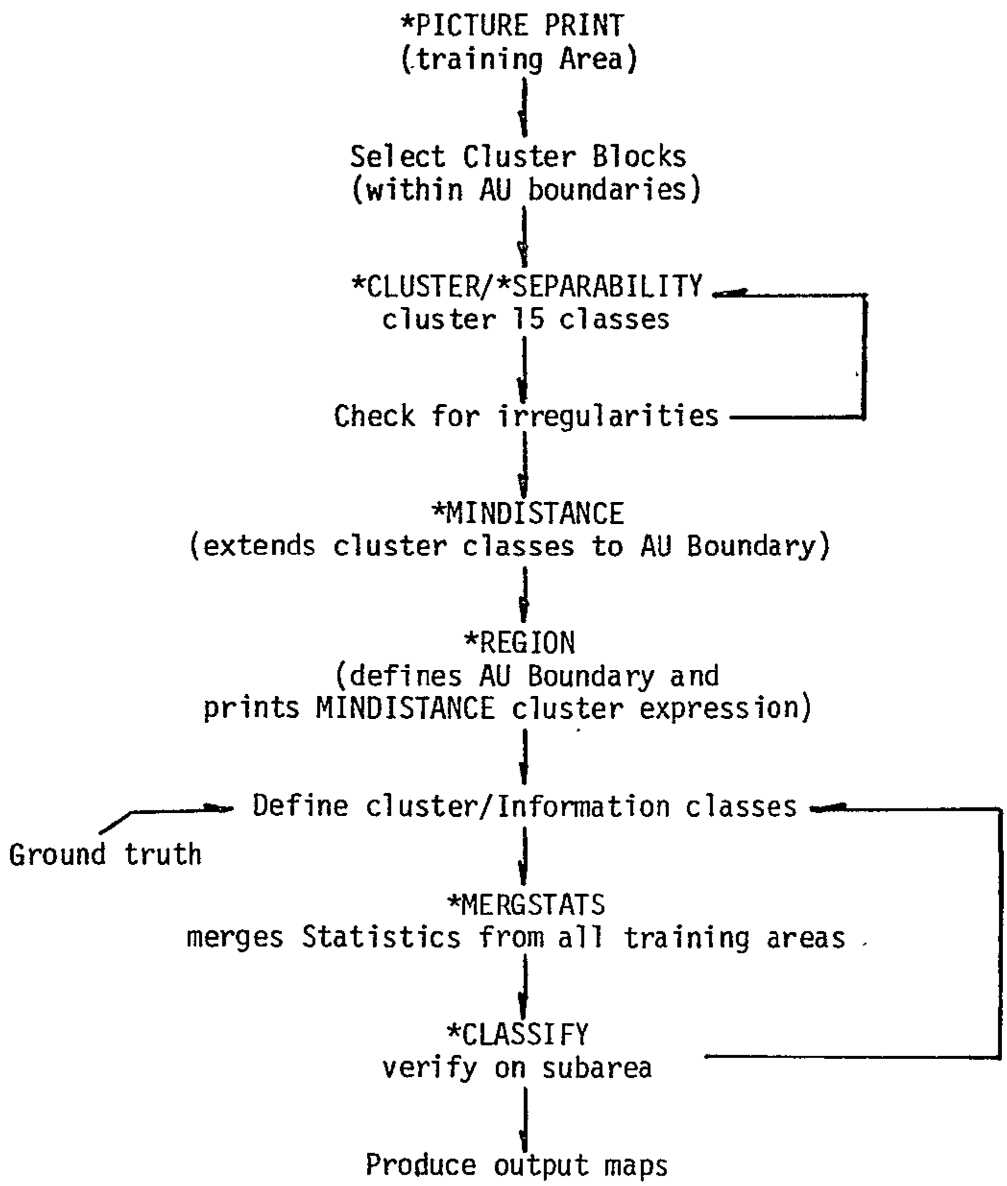


Figure 1. Flow diagram for FRIS classification procedures.

2. Clear acetate overlays should be obtained:
    a. for each Unit selected for training.
    b. updated in response to significant cultural change.
    c. permanently archived for immediate reference.
3. Boundary annotation should be made for:
    a. all Administrative Unit boundaries within each test area defined in Phase II FRIS, including the training units.
    b. all AU and Operating Area boundaries in the prime test site.

B. Classification Training Procedures as outlined in Figure 1
1. To be carried out on each training unit within each tract.
2. Generate line printer output (PICTURE PRINT) for each training unit defined in A above.
    a. For a given run (scene) line and column range with appropriate interval will be defined such that the range in both lines and columns will encompass the entire training unti.
    b. Gray scale. *PICTURE PRINT/*G DATA displays only one channel at a time. The channel best suited to locational information should be used; i.e., one of the visible channels. Optional step if area is known. Used primarily to pick cluster blocks.
    c. Unless the analyst has preference, the symbol array offered by the default option is generally satisfactory for this gray scale print-out.
3. Select cluster blocks within selected Administrative Units.
    a. Blocks will fall wholly within the boundaries of the AU in such a way as to be as inclusive as possible.

b. As many rectangular blocks will be generated as needed to properly represent the range of conditions within the unit.

c. For efficienty, Cluster blocks should range from 2500 - 4000 pixels (50 x 50 - 70 x 70) - blocks do not have to be square.

4. Cluster/Separability *CLUSTER/*SEPARABILITY

a. In clustering an arbitrary 15-classes will be designated based upon the standard size defined in 3c above. Other sizes will be considered as exceptions to this rule.

b. Separability will always be run behind Cluster as a matter of form.

*c. Analyst check point - with 15-cluster classes, little or no combining of classes is expected at this stage of the process.

o Check separability means against expected ranges in both the visible and IR for obvious irregularities.

5. Minimum Distance Classifier *MINDISTANCE - Purpose is to extend the 15-cluster classes to the boundaries of the picture-print block.

6. Region definition of Administrative Unit boundaries - *REGION

a. Defines AU within the picture-print block.

b. All area outside Unit boundaries will be null characters to be assigned by analyst.

c. By essentially clustering the entire AU in this fashion, the maximum repeat cluster classes will occur in direct relation to the map overlay. This will facilitate and help verify class definition described and performed later on in these proceedings.

7. Associate Cluster classes with information classes
    a. This process done for each training unit within the tract.
    b. Statistics deck generated and placed on temporary disk.
    c. Utilize data from SEPARABILITY to aid in identifying and combining classes.
    d. The overlayed map and associated aerial photographs should also be helpful.
8. Merge the statistics from all training Units.
    a. As decks are merged, combine like classes, checklines, with the various unit maps and photographs and other ground truth (updating) as available.
    b. Keep going through the MERGE procedure until one classification deck results.
9. Classify - *CLASSIFY
    a. If any doubt exists, classify small sub-unit to verify training.
    b. Select symbols indicative of the classification features to be emphasized.

All classification work to date has followed this approach. Since the classification task is to be operationalized and, therefore, repeatable, we forsee making modifications to the procedures. One of the first major modifications anticipated would involve the CLUSTER sub-routine. Currently only geometric blocks can input to clustering. We would envision a modification that would accept irregular areas, such as AU boundaries to the CLUSTER sub-routine. This change would eliminate the MINDISTANCE and REGION steps from the flow diagram in Figure 1. As experience is gained in performing repeat classifications we anticipate further streamlining of the classification procedures.

2.1.2 Classification Evaluations

In the process of developing a Landsat data classification procedure, we have done extensive classification work on four Administrative Units in Test Area 1. Results in the form a areal estimates of Level I cover types (pine, mixed pine/hardwood, and other) were presented in the last quarterly report. Table I summarizes the bitemporal classification Level I results, by percent area by class, for four Administrative Units.

Table 1. Percent of area by class for a 4-Channel *Bitemporal Classification of AU 264, 267, 268, 271. (STR inventory comparison shown in parentheses.)

Cover Type	Administrative Unit				
	264	267	268	271	Total
Pine	57.8	51.3	49.2	61.5	55.4
	(59.5)	(53.3)	(46.1)	(65.1)	(56.9)
Mixed P/H	36.2	46.1	49.5	27.3	39.0
	(38.0)	(43.2)	(49.4)	(32.6)	(40.1)
Non-stock	6.0	2.6	1.2	11.2	5.6
	(2.5)	(3.5)	(4.5)	(2.4)	(3.1)

*Channels Used: December 30, 1976 0.70 - 0.80 micrometers
0.80 - 1.10 micrometers
April 17, 1977 0.60 - 0.70 micrometers
0.70 - 0.80 micrometers

These estimates were calculated in the following manner:

1. Sum the number of pixels inside the Administrative Unit. Tally pixels by the class into which each pixel has been classified.

	Pine	Mixed pine/hardwood	Other	Total
Sum of Pixels classified as:	T_P	T_M	T_O	$T_P+T_M+T_O = N$

2. The value (T_P, T_M or T_O) can be associated to area by multiplying by the ratio of the total area to the total number of pixels tabulated.

$$\text{acreage of Pine} = T_P \cdot \frac{(\text{total area})}{N}$$

This is basically multiplying the total area by the percent of pixels classified as P ($\frac{T_P}{N}$). Although this method is simple and easy, it has several limitations. First, we must assume the same proportion of the various classes are on the boundaries of the area. This becomes more of a problem if we use a systematic sample rather than all the pixels. The systematic sample may cause a class of timber which grows in strips (e.g., along a branch) to be under-represented resulting in the over-representation of other classes. The second and more important problem with this method is that it is biased due to mapping, classification, and registration error. By using a systematic sample with a buffer zone, some of this error is compensated for.

A second method for calculating an area estimate is called the "unbiased estiamte" because it calculates an unbiased acreage estimate directly. Using Table 2 which we obtain from our map and Landsat inspection of an area, we calculate an error matrix E. The error e_{ij} is the ratio

$$E = \begin{matrix} e_{PP} & e_{PM} & e_{PO} \\ e_{MP} & e_{MM} & e_{MO} \\ e_{OP} & e_{OM} & e_{OO} \end{matrix} \quad \text{where } e_{ij} = \frac{a_{ij}}{A_i}$$

of the samples from class i classified into class j to the total number of samples in class i. Hence $e_{PP} = P(\hat{P}/P)$ or e_{PP} is the probability that a point is classified as pine when it really is pine.

Table 2. Classification error matrix for Landsat classification compared to management maps.

Cover Type	# of Map Samples	# Landsat pixels classified as Pine	Mixed	Other
Pine	A_P	A_{PP}	A_{PM}	A_{PO}
Mixed	A_M	A_{MP}	A_{MM}	A_{MO}
O	A_O	A_{OP}	A_{OM}	A_{OO}
	N	T_H	T_C	T_O

To obtain the unbiased estimate of area, we use our estimates (of proportions or T_i/N) in the first method, call them $\hat{p}_P$, $\hat{p}_M$ and $\hat{p}_O$ ($\hat{p}_M = \frac{T_M}{N}$) and let $\hat{P}$ be the matrix:

$$\hat{P} = \begin{matrix} \hat{P}_P \\ \hat{P}_M \\ \hat{P}_O \end{matrix}$$

To unbiase these estimates from the first method, multiply $\hat{P}$ by the inverse of the transpose of the error matrix:

$$P_{unbiased} = (E^T)^{-1}\ \hat{P}$$

It is possible for this method to give a negative value for the proportion of a cover type. When this happens the following is an alternative formula for determining P:

$$\min_{0 < P_i < 1} \| P - (E^T)^{-1}\ \hat{P} \|$$

Once $P_{unbiased}$ is determined, acreage estimates are calculated in the usual manner:

$$\text{acreage of P} = (P_{P\ unbiased})\ (\text{total area})$$

The third method is the Stratified Aerial Estimate. Using Table 2 let:

$$d_{ij} = \frac{a_{ij}}{T_i} \qquad i = P, M, 0, \cdot\ j = P, M, 0$$

$\alpha_{PP} = P(P/\hat{P})$ or α_{PP} is the probability of finding pine when the area's been classified as pine. The proportion P_i of each timber type is then found in the following manner:

$$P_i = \frac{1}{N} \sum_j \frac{a_{ij}}{T_i}\ A_j \qquad \text{where: } N = A_P + A_M + A_0$$
$$i = P, M, 0$$
$$J = P, M, 0$$

The acreage estimate is again found by:

$$\text{Acreage of P} = P_p(\text{total area})$$

This method has the same drawbacks as the "unbiased estimate." Both methods produce unbiased estimates if there is a large number of test samples and no error in identification in the test samples. Also, it must be assumed that the test samples are allocated proportional to the classes occurrance in the scene. The systematic sample design could cause problems of this type, since samples are not allocated proportionally to actual class structure.

When using all the pixels as the sample, the first method can be done easily from the computer output which generates the Landsat maps. When proceeding withethe systematic sample, however, more man hours are required. Also, the systematic sample has a smaller sample size than the sample of all pixels and hence it has less statistical accuracy (the systematic sample is a sample from the population of all pixels).

Both method I and either method II or method III can be performed and given as biased estimates (method I) and corrected Landsat estimates (method II or method III).

Method II, the "unbiased estimate" was utilized on a four AU sub-set of Test Area 1. A systematic sample was used to determine the classification accuracy of a Level 1 classification. The sample consisted of 2,033 test fields covering the 6,144 hectare (15,360 acre) training area. Each test field consisted of four pixels, approximately 1 hectare (2 acres) in size, arranged in a 2 x 2 matrix. Figure 2 shows an example of the systematic sample.

The map of systematic test fields was produced at a line printer output scale of 1:15,840. In this manner the test field map could be registered with management maps, showing forest cover types for the four Administrative

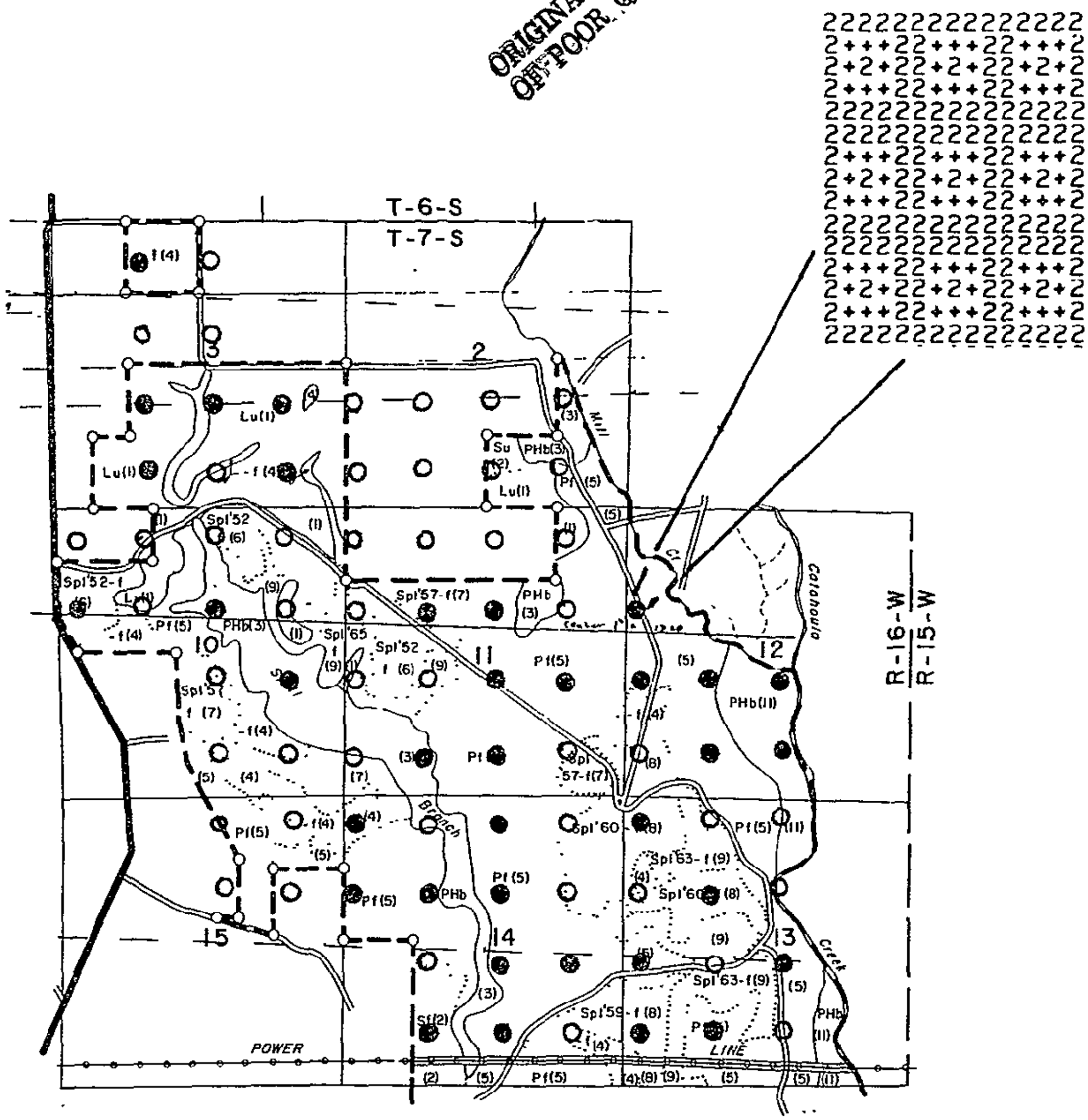


Figure 2. Example of a systematic sample that was used to evaluate classification accuracy. The circles represent potential sample points. The darkened points are those homogenous blocks, eq: insert that were used for evaluation.

Units. Only test fields containing a homogenous pixel set (4 identical pixels) were considered. This number was reduced further to eliminate test fields that fell on obvious, irreconcilable map boundaries, e.g., boundaries between pine-hardwood versus boundaries between two different pine stands. Ultimately 42.7%, or 869 fields were evaluated to assess classification performance. The results of this evaluation appear in Table 3. The overall classification performance figure of 85.4% together with the areal estimates of cover for the training area appear reasonable in view of the physical anamolies of the site.

Specifically, the evaluation presented in Table 3 is based on a comparison of Landsat classification to a forest management map of the same area. The map was developed through interpretation of color infrared aerial photographs that were exposed during December of 1975, a year before the Landsat data was collected. At best, the map represents an approximation of ground based on simplifying the photographic image. Therefore, the map is certainly not the best source of ground reference data. However, there is no better source for evaluation purposes. Direct comparison to the aerial photographs may induce additional errors due to:

a. our unfamiliarity with the site.
b. geometric distortions in the photographs which have been eliminated to an extent in the maps.
c. better positional accuracy between the classification and the map, since the Landsat data was registered to the maps.

Given that the maps are the best medium for evaluating performance, one must evaluate the results in Table 3 keeping in mind the dynamics of the ecosystem involved. Under an intensive forest management regime, one

Table 3. Classification performance of four AU Training areas in Test Area 1 based on an evaluation of Test fields.

Class	Number of 2 x 2 pixel test fields	Percent of test field area	Percentage of Test fields Classified		
			Pine	Mixed Pine/Hardwood	Non-Stocked
Pine	547	6.8	87.4	10.8	1.8
Mixed Pine/Hardwood	302	3.0	11.6	85.4	3.0
Non-Stocked	20	.3	35.0	35.0	30.0
Total	869	10.1			

Overall Test Field Accuracy 85.4%

would not expect to find a significant portion of the land area non-stocked. This is supported with information from both Tables 1 and 3, specifically in the percent of the area in the non-stocked class. However, one could anticipate, at any point in time, a higher proportion of pixels classified as "non-stocked", since these pixels would reflect clearcutting and regeneration practices which are spectrally inseparable from a "non-stocked" situation. On the maps, young regeneration is classified as pine, and clearcuts are never identified, thereby creating some difficulty during the classification evaluation. An example of this problem is shown for the "non-stocked" class in Table 3. The high proportion of test fields falling in pine and mixed/pine-hardwood can be attributed to, pine-regeneration, clearcut, or poorly stocked hardwood situations. The low proportion of "non-stocked" is simply a function of sample size, since very little truly "non-stocked" land exists in the test area.

Similar situations exist between the pine and mixed/pine-hardwood, but the examples are not a dramatic. Often the misclassification between these classes is a function of the proportion of pine or hardwood in the mix. Since the class is heterogenous a unique spectral class cannot be identified, and therefore, some classification error is likely to occur.

Future evaluations will utilize a single rather than a multiple pixel point. Proceeding in this fashion will increase the number of samples and help to decrease the man-time that is involved with the multiple pixel systematic sample.

2.2 Mapping Unit

An important part of this demonstration involves the ability to merge Landsat multi-spectral scanner data with ground reference data. The reference data is in the form of maps and forest inventory information. The ability to merge these diverse data sets is critical for the success of the demonstration. The necessity of this task is apparent when one understands that Landsat classifications unto themselves are poorly referenced in regard to the ground. The requirement to relate the Landsat classification to the ground is extremely important for industrial forest management purposes, since management activities are related to specific land parcels.

Our plan for creating this merged data set involved digitizing management maps and then using these data to create ancillary channels on the Landsat master tape. Minimally, the Administrative Unit boundaries would be digitized in vector mode as polygons. These polygons would be gridded and overlayed on the Landsat data. Landsat classification maps and acreage summaries could then be produced by Administrative Unit. Since this data was created at a map scale of four inches to the mile (1:15,840) it could be overlayed directly onto management maps to locate cultural or topographic features that had not been digitized. This is not an optimum solution for an operational system. However, the output from this procedure would graphically, but crudely, demonstrate a data base capability that would be a key part of FRIS. This capability would be expanded to be computer oriented and eventually become highly automated.

The following sections deal directly with creation of the ancillary data channels and the availability of computer data base software.

2.2.1 Map Digitizing

This section will deal with the steps involved in creating an ancillary data set, specifically maps like Figure 3, and overlaying these as a channel on the Landsat master tape. Four general steps are involved in this process; Map Preparation, Digitizing, Data Assembly, and Boundary Processing. These steps together with a short description of activity each involves is presented below:

A. Map Preparation

Management maps, such as Figure 3, are carefully checked to determine if all boundaries close (all boundary lines meet), and if areas enclosed by boundaries (polygons) are named, e.g., by forest type or numerical operating area designation. Once the maps are verified and any problems eliminated; and since the boundaries will be digitized as vectors, arch numbers, and left and right area numbers will be assigned. Also at this point in the process each map will be assigned a unique file name for use during digitizing. Figure 4 is an example of the map elements to be digitized.

B. Digitizing

During this step the actual creation of the digital map file is accomplished. A table digitizer which is interfaced to a PDP 11/34 mini-computer is utilized in this process. The map vectors are converted into a digital file stored on disk or magnetic tape in this step.

C. Data Assembly

This activity involves manipulation of the independent digitized map files to form a single file for each ownership. During this

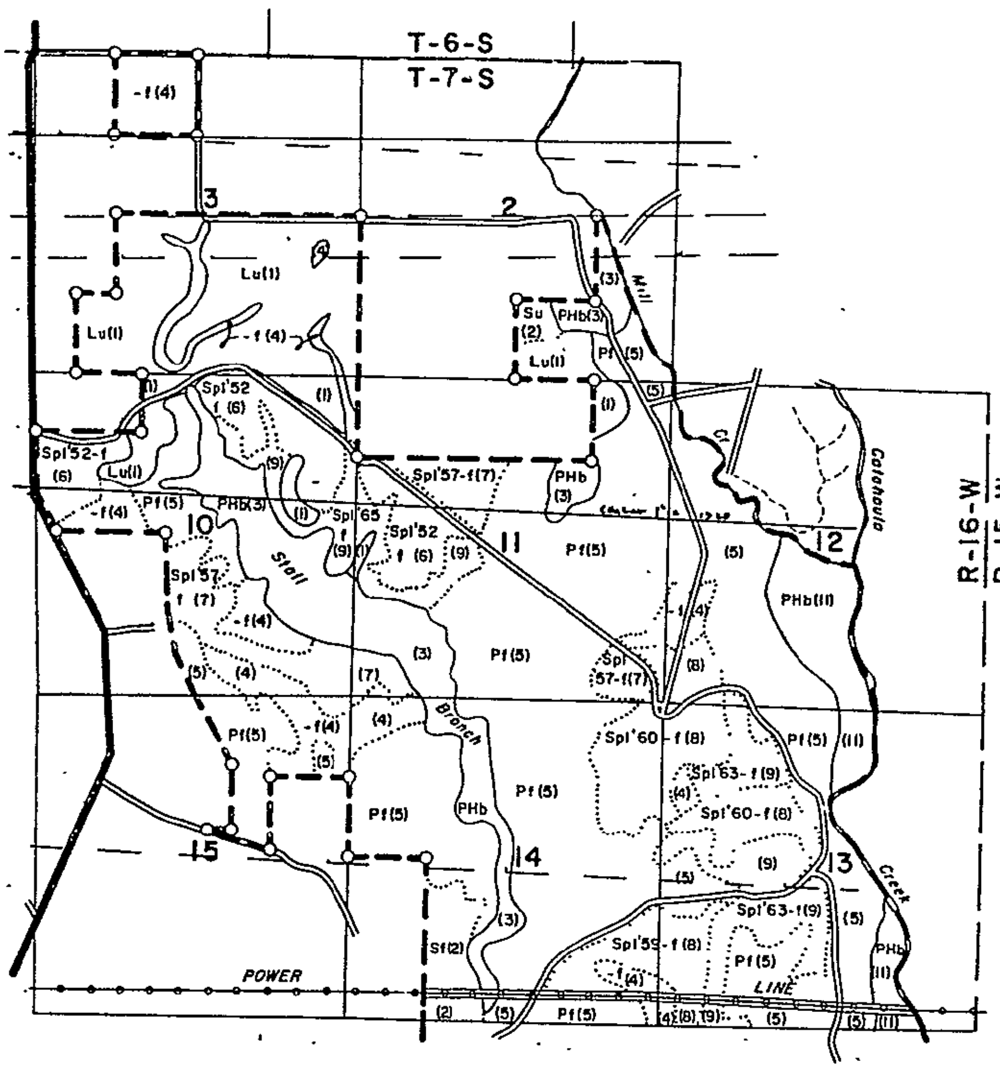


Figure 3. Example of management maps which are digitized for inclusion as channels in the master Landsat data tape.

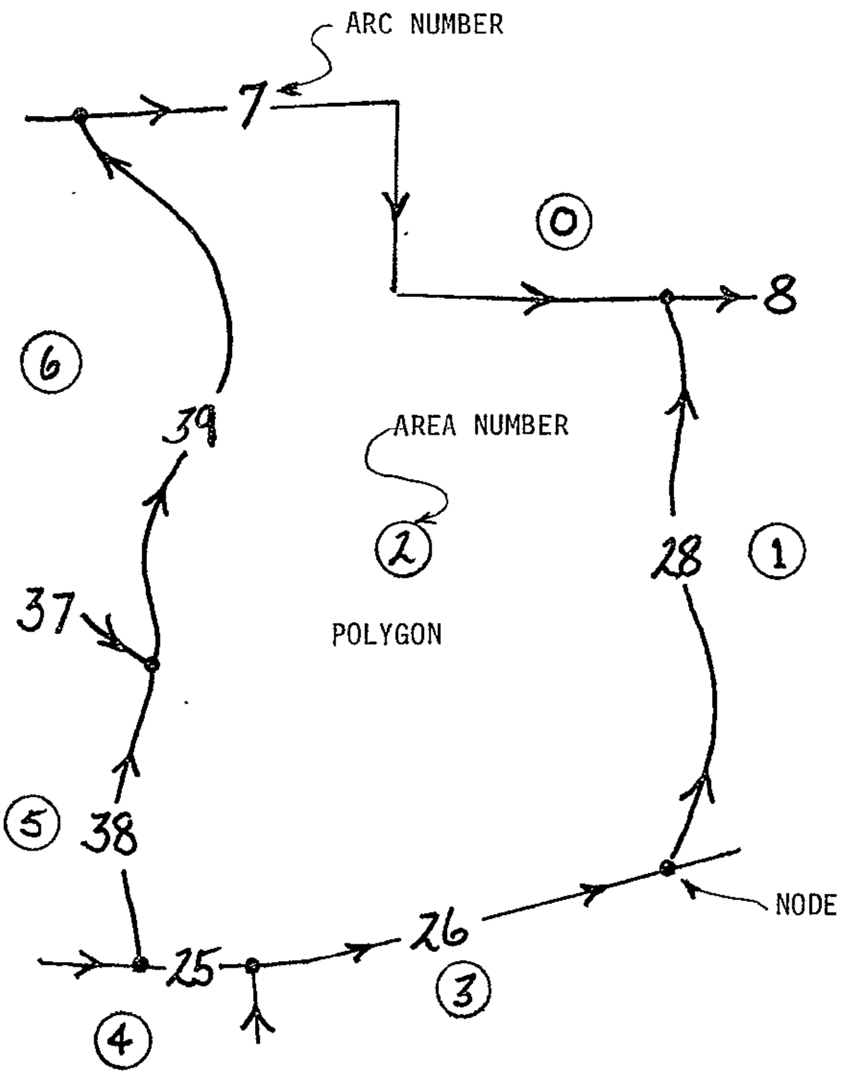


Figure 4. An example of the map elements that are recorded on the digital file.

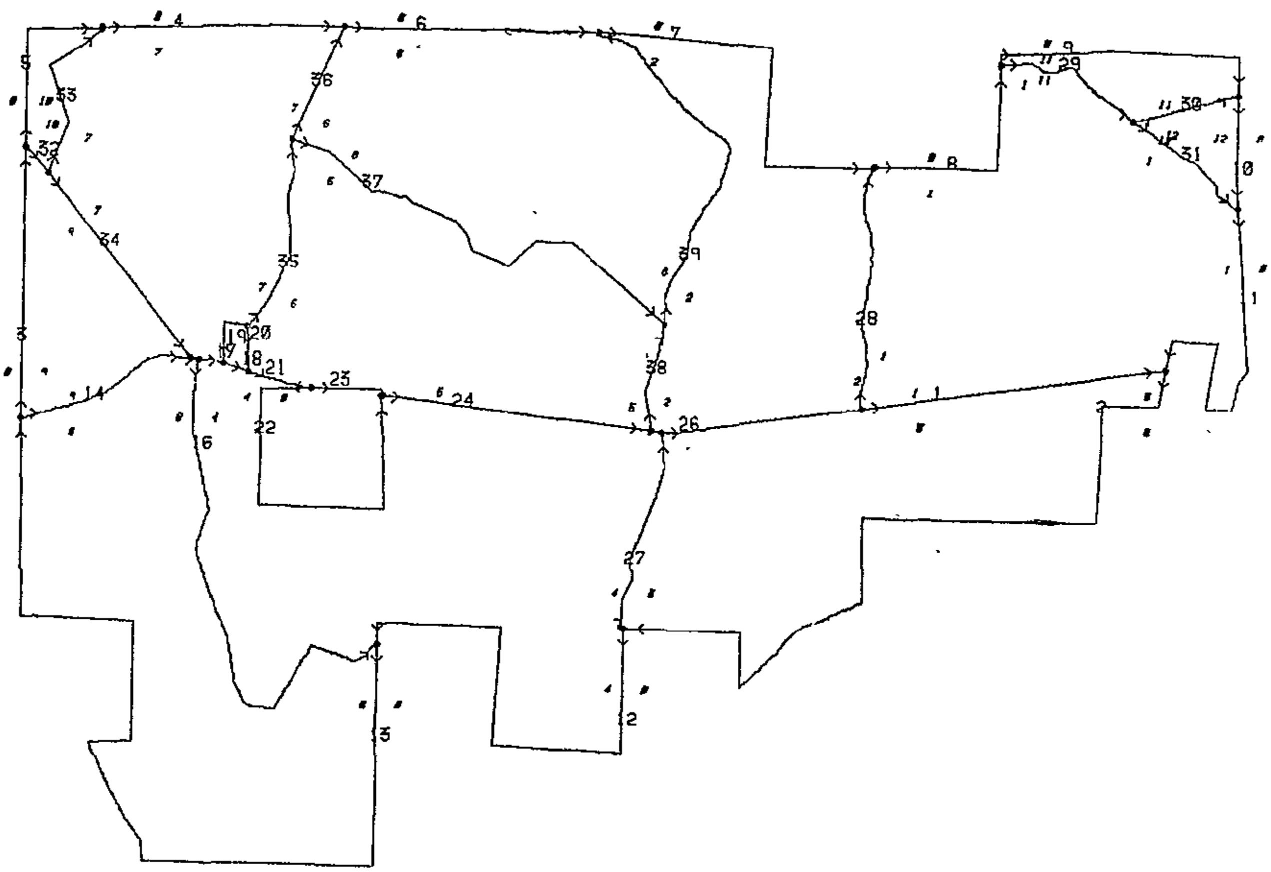
operation maps are rotated as needed to attain proper fit, arcs are edited to insure that arc nodes properly meet, area numbers are verified, and arc and area numbers are re-sequenced to eliminate duplicate numbers. All these operations are performed on a back-up tape so that none of the original digitized data will be lost or modified. Output of this operation is in the form of a map, Figure 5.

D. Boundary Processing

The boundary process converts the digitized map vectors to the Landsat raster format. Also during this process the boundaries are checked for errors and edited, as necessary, for corrections. The output from this step is illustrated in Figure 6.

Prior to executing the boundary software, the Landsat Computer Compatible Tape (CCT) had been reformatted and a coarse geometric correction performed on the relevant portion of the Landsat scene. Control points, which are identifiable on both the digitized maps and in the Landsat data have been identified. These points are used to transform the coarsely corrected Landsat data to a precision registered data set. The registration step precisely relates the Landsat data to the map data base. The nominal registration error of the Landsat data measured against the map control points is .5 pixel RMS.

Currently this is the process used for preparing Landsat II data. Since Landsat III data will be received in a different format, we assume that the preprocessing sequence for the spectral data would be modified. However, we have insufficient information regarding the Landsat III format at this time to expressly define what system changes would be made.



ARC	LEFT	RIGHT	ARC	LEFT	RIGHT
15	7	8	17	7	4
18	0	4	19	7	0
20	5	0	21	5	4
23	5	0	25	2	4
26	2	3	32	10	1

Figure 5. An example of a digitized map file. Arc numbers and area designators should correspond to the table in the lower right of the map.

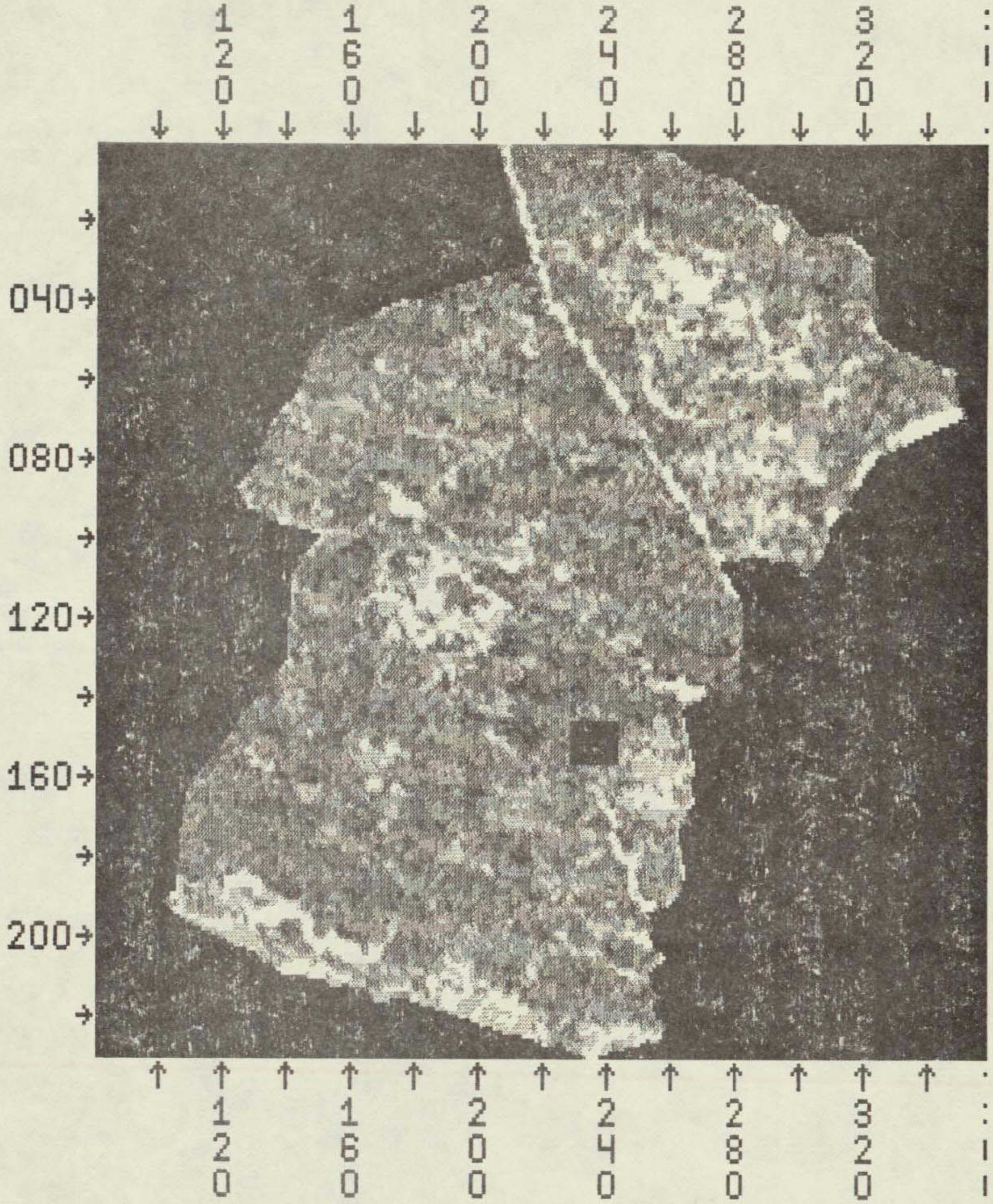


Figure 6. An example of digitized map data overlaid on a Landsat data channel.

Flow charts for the map data processing are given in Figure 7a, b, and c. The flow charts provide specific information with regards to the device on which the process occurs and where back-up data is stored. We have estimated tha map data processing procedure requires the following resource allocation:

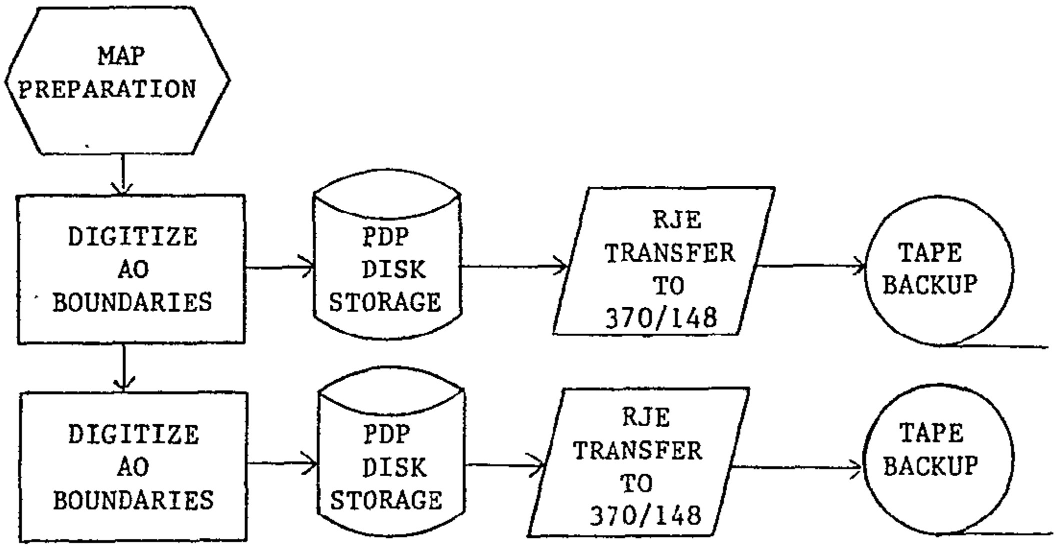
Process	Resources Required Percent
Digitizing	30
Data Edit/Assembly	60
Boundary Processing	10

Once the map processing is complete we are ready to classify the data set. The classified data can then be manipulated within a data base structure to provide maps and acreage statistics by Administrative Unit. Programs available at LARS for this manipulation are defined in the next sub-section.

2.2.2 Data Base Interaction Systems

A number of data base manipulation programs are available at LARS which have the potential for providing the needs of the Mapping Unit for the FRIS project. These programs are on the LARS 370/148 Computer system and can be relatively readily accessed. In addition, there are numerous data base systems available from external sources which are well documented and have extensive capabilities. Such a large scale system will be required for an operational capability by a FRIS resident at St. Regis. A review of the internal and some attractive external systems are presented and recommendations are presented for Phase II implementation. This discussion concerns data base interaction systems only; input and output approaches are discussed separately.

MAP PREPARATION/DIGITIZATION



DATA ASSEMBLY

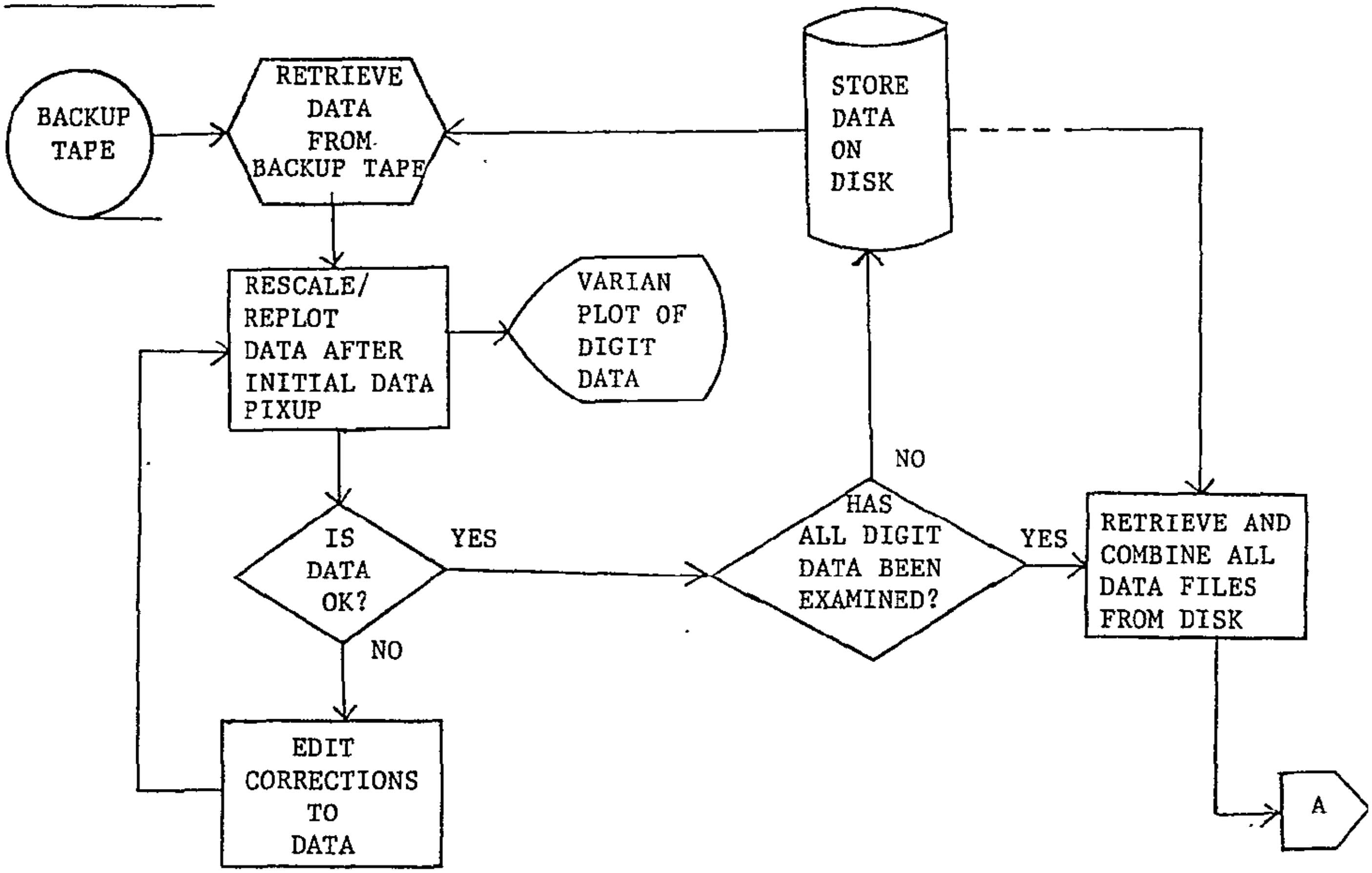


Figure 7. Flow charts of the preprocessing steps necessary prior to classification. a) Map Preparation/Digitization, b) Data Assembly, c) Boundary Processing.

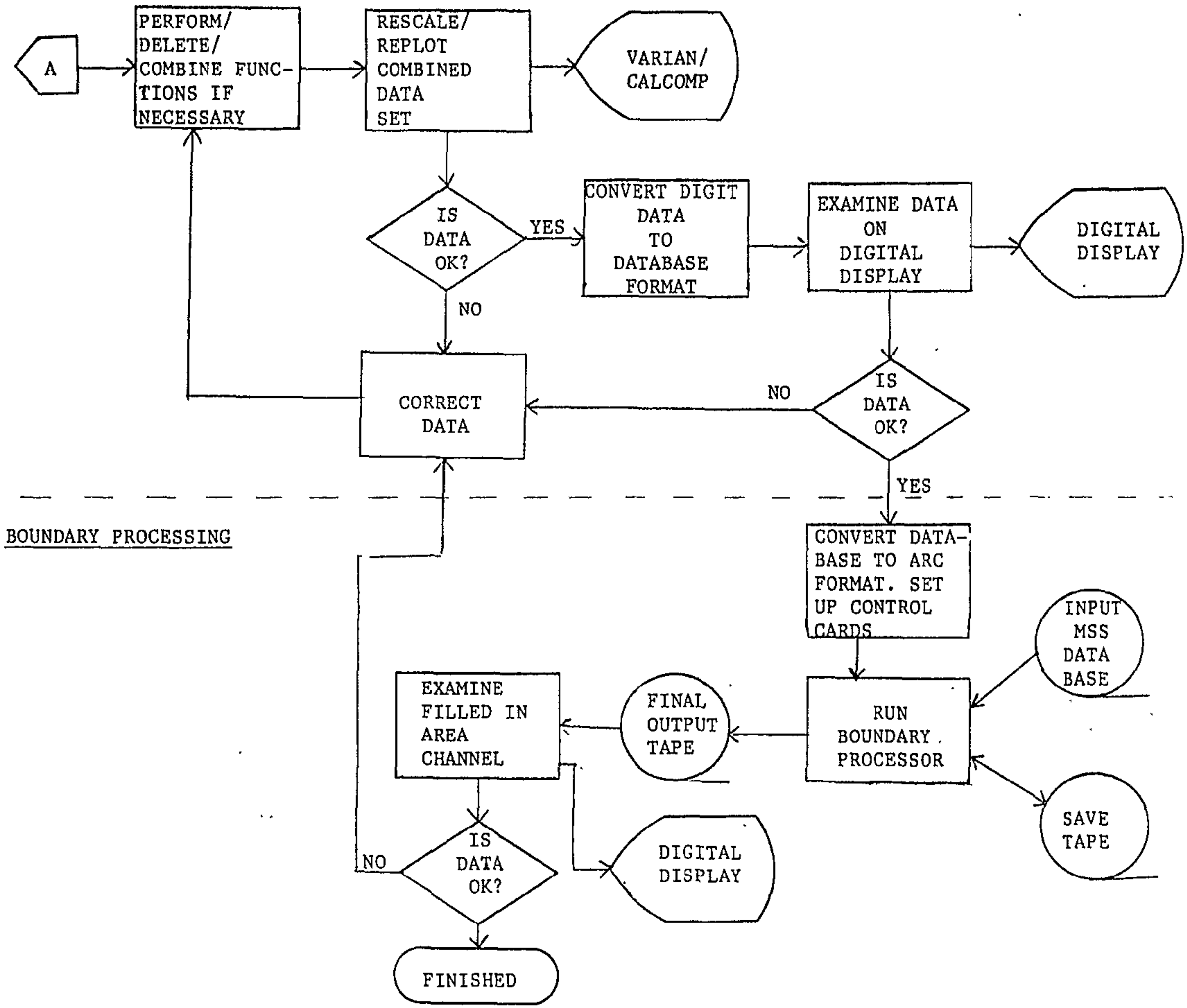


Figure 7. continued

DB Interaction Programs at LARS

1. GRID: This program was imported to LARS from Harvard Univ. in 1971 to relate variables in a Tippecanoe County. The data base that was created contained around 40 variables. The data must be put into a rectangular grid array (1/10 km square in this example) and stored on tape. The program reads the tape and creates maps based on criteria supplied by the user. The program cannot produce count tables. The user interaction features of this program are cumbersome and newer, easier to use versions are available but not at LARS. This program has the ability to relate a large number of variables; however, this is not needed in Phase II. The program is poorly documented and for these reasons GRID is likely not a candidate for use in FRIS.

2. REGION: This program reads a classification tape and a data tape with an ancillary data channel to be used as a mask and outputs all classification points coinciding with a given parameter value in the ancillary data channel. Tables are produced. It can compare only two channels in one job.

3. COMBINE: A system developed for a Landsat study which relates classification results and topographic data. One program merges the topographic tape and the classification file to form a multispectral image storage tape with both types of data on it. The COMBINE program then associates given classes and ranges of topographic variables to produce maps of the result classes. No tables are produced. This program could be used to relate three (the ones in this application was elevation, slope, aspect, but any three variables could be used) variables to a classification map. Tables could be produced using the program COUNT described below.

4. COUNT: A program which reads a classification tape and a channel from a

MIST (Multispectral Image Storage Tape) tape and produces a table of the number of occurrances of each class in each value of the ancillary variable channel. Directly useable for FRIS classifications. Labels can be attached to the count tables.

These programs can be readily modified to produce OA and AU selective maps and counts from FRIS classifications. This assumes that AU's and OA's are registered to the Landsat data files.

DB Programs Available Externally

Numerous data base systems exist which have been developed by industry, government and academic agencies, several systems have been considered and a brief description is presented here:

1. COMLUP: This system is a batch-oriented data base input interaction and output system offered by the University of Massachusetts and developed by the U.S. Forest Service. This program takes digitized arcs as inputs, converts this data to grid cell format, carries out processing in grid format and produces output maps in grid or polygon format. Tabular output is also produced. A cell capacity of 500,000 is specified. The complete package includes CRT interactive interrogation features and the cost to universities is $900.

2. COMARC: The COMARC Design Systems Co. provides a polygon oriented data base system as a customer service. This means that the object programs are installed on the users computer and maintenance and up-dating must be handled by COMARC. Input arc processing, polygon interaction and output mapping are included and tabular output is provided. Grid data is handled by converting it to polygon format. Hardware including a computer is included.

3. M & S Computing: M & S Corp. provides a polygon system very similar to COMARC's with hardware and software as a package. The system is polygon oriented and is primarily used for architectural planning and drafting automation. It is not equipped to interact with grid format remote sensing data.
4. ODYSSEY: A data base system developed by Harvard which is well documented and designed for transferability. It is polygon oriented and does data base interaction in polygon mode. Its capability to utilize grid remote sensing data is unknown, but being investigated.

These are the systems given particular attention by the LARS FRIS staff. A tabular review of 24 systems is presented in Table 4 which includes the four discussed. The only ones which are listed as highly transferable (transferability codes 1 or 2) are the COMARC, Grid II, AUTOMAP, LCDMS and ODYSSEY systems.

Consideration of the requirements of Phase II and the programming tasks involved in importing any large system has led to the conclusion that the basic needs of the Phase II demonstration can be met by modification of current LARS software. The REGION or COMBINE system can be modified to relate OA, AU and classification results to provide tabulations and maps suitable for the purposes of FRIS. Phase III would include selection of complete data base system, acquisition and installation at St. Regis for large areas, multiple variable data base manipulation. Program analysis and modification planning was caused out in the quarter and modification and testing of the programs will be carried out in the next quarter.

Table 4. Summary of available geographic information systems from: Computer Science Corporation, 1978 Geographic Information System Survey Interim Report prepared to Contract NAS 5-24350.

STANDARD REPORTING FORM CATEGORIES	ORGANIZATION AND SYSTEM ACRONYM			
	W. E. GATES & ASSOCIATES ADAPT	ALABAMA DEVELOPMENT OFFICE ARIS	BENDIX AEROSPACE	COMPUTERVISION CADDS 3/MAP
PROGRAMMING BASIS & OPERATING INFORMATION				
OPERATIVE COMPUTERS	IBM 370, AMDAHL V6	IBM 370/158	PDP11, PDP10, IBM/370, Z80	NOVA 1200, DCC 118 COMPUTERVISION CGP-100
PROGRAMMING LANGUAGE	FORTRAN IV	FORTRAN IV, PL/1	FORTRAN IV, MACRO ASSEMBLY	FORTRAN, APL, ASSEMBLER
MODE OF USAGE	BATCH	BATCH & INTER-ACTIVE	BATCH & INTER-ACTIVE	INTERACTIVE
MEMORY SIZE	256-350K	128-150K	192K	8K & 24K
WORD SIZE (BITS)	32	32	16	16
GEOGRAPHIC DATA TYPE				
INPUT				
LINE	YES	YES	YES	YES
CELL	YES	NO	YES	NO
TABULAR	YES	NO	YES	NO
POLYGON	YES	YES	YES	NO
ANALYSIS				
CELL	–	YES	YES	NO
POLYGON	–	NO	YES	YES
TABULAR	YES	NO	YES	NO
CELL & POLYGON	YES	NO	YES	NO
DATA ENTRY & DATA OUTPUT PRODUCTS				
ENTRY				
AUTOMATIC	YES	YES	YES	YES
SEMI-AUTOMATIC	NO	NO	YES	NO
MANUAL	YES	YES	YES	YES
OUTPUT PRODUCTS				
GRAPHIC	YES	YES	YES	YES
TABULAR	YES	NO	YES	ANALYSIS REPORTS
DIGITAL	NO	YES	YES	NO
ANALYTIC CAPABILITIES				
COMPOSITE MAPPING	YES	YES	YES	–
POLYGON OVERLAY	YES	NO	POLYGON INTER-SECTION	–
CELLULAR	YES	YES	YES	–
ABILITY TO VARY SCALE	YES	NO	YES	YES
ABILITY TO VARY RESOLUTION	YES	YES	–	–
AREA MEASURE	YES	YES	YES	YES
SIMULATION AND/OR MODELING	YES	YES	YES	NO
BOOLEAN COMBINATIONS	YES	YES	YES	NO
CORRELATION	YES	NO	YES	NO
REGRESSION	YES	NO	YES	NO
INTERPRETIVE MAPS	YES	YES	YES	YES
DATA STORAGE				
STRUCTURE				
DIRECT ACCESS	YES	NO	YES	YES
SEQUENTIAL	NO	YES	YES	YES
OTHER	N/A	N/A	N/A	N/A
ORGANIZATION				
HIERARCHICAL	NO	YES	YES	YES
POINTER	YES	NO	YES	YES
RELATIONAL	YES	NO	YES	YES
INTERFACE WITH CLASSIFIED				
LANDSAT DATA				
EXPERIMENTALLY	–	NO	YES	NO
OPERATIONALLY	YES	NO	YES	NO
ACQUISITION CONDITIONS	LEASE SPECIFIC USER BY PROJECT	–	SPECIFIC USERS ONLY	AS PART OF SUPPLIED SYSTEM
STATUS OF RELEASE	TESTED & EVALU-ATED	–	PARTIALLY TESTED & EVALUATED	TESTED & EVALU-ATED
WILLINGNESS TO ADAPT TO HP3000	YES	NO	NO	NO
WITHIN COST THRESHOLD	BASELINE SYSTEM – YES	–	N/A	NO
WITHIN TIME THRESHOLD	TO BE DETERMINED	–	N/A	FOR TURNKEY SYSTEM – YES
CUSTOMER SUPPORT	MAINTENANCE/ CONSULTATION	–	MAINTENANCE/ CONSULTATION	MAINTENANCE
TRANSFERABILITY*	4	4	5	5

*Note The data presented in this table has been supplied by the individual organizations and has not been verified by Computer Sciences Corporation or NASA.

Table 4. continued

STANDARD REPORTING FORM CATEGORIES	ORGANIZATION AND SYSTEM ACRONYM			
	U.S. FORESTRY SERVICE COMLUP	COMPREHENSIVE PLANNING ORGANIZATION	UNIVERSITY OF GEORGIA CONGRID	COMARC CRIS
PROGRAMMING BASIS & OPERATIONS INFORMATION				
OPERATIVE COMPUTERS	CDC 3100	BURROUGHS B6700	BURROUGHS, IBM 360/370 UNIVAC 1100	DATA GENERAL ECLIPSE
PROGRAMMING LANGUAGE	FORTRAN	FORTRAN (BASIC — 4061)	FORTRAN	FORTRAN IV & V
MODE OF USAGE	BATCH	BATCH	BATCH & INTER ACTIVE	INTERACTIVE & REALTIME
MEMORY SIZE	32K	VARIES WITH PROGRAM	250K	128—512K
WORD SIZE (BITS)	6	VARIES WITH PROGRAM	4	16
GEOGRAPHIC DATA TYPE				
INPUT				
LINE	YES	YES	NO	YES
CELL	NO	NO	YES	YES
TABULAR	NO	NO	NO	YES
POLYGON	NO	YES	NO	YES
ANALYSIS				
CELL	YES	—	YES	YES
POLYGON	NO	—	NO	YES
TABULAR	NO	YES	NO	YES
CELL & POLYGON	NO	YES	NO	YES
DATA ENTRY & DATA OUTPUT PRODUCTS				
ENTRY				
AUTOMATIC	NO	YES	YES	YES
SEMI AUTOMATIC	NO	NO	NO	NO
MANUAL	YES	YES	YES	YES
OUTPUT PRODUCTS				
GRAPHIC	YES	YES	YES	YES
TABULAR	NO	YES	YES	YES
DIGITAL	YES	YES	NO	YES
ANALYTIC CAPABILITIES				
COMPOSITE MAPPING	YES	YES	YES	YES
POLYGON OVERLAY	POLYGON INTER-SECTION	YES	—	YES
CELLULAR	—	—	YES	YES
ABILITY TO VARY SCALE	YES	YES	NO	YES
ABILITY TO VARY RESOLUTION	—	—	—	YES
AREA MEASURE	YES	YES	YES	YES
SIMULATION AND/OR MODELING	NO	YES	YES	YES
BOOLEAN COMBINATIONS	YES	YES	YES	YES
CORRELATION	NO	NO	NO	YES
REGRESSION	NO	NO	NO	YES
INTERPRETIVE MAPS	YES	YES	YES	YES
DATA STORAGE				
STRUCTURE				
DIRECT ACCESS	NO	YES	YES	NO
SEQUENTIAL	YES	YES	NO	NO
OTHER	N/A	NO	N/A	RANDOM
ORGANIZATION				
HIERARCHICAL	YES	YES	GRID MATRIX NO	NO
POINTER	NO	YES	NO	NO
RELATIONAL	NO	NO	NO	YES
INTERFACE WITH CLASSIFIED LANDSAT DATA				
EXPERIMENTALLY	NO	NO	YES	YES
OPERATIONALLY	NO	NO	—	NO
ACQUISITION CONDITIONS	HANDLING/MAILING	HANDLING/MAILING ONLY	FREE OF CHARGE	LEASE
STATUS OF RELEASE	TESTED & EVALU-ATED	TESTED & EVALU-ATED	TESTED	TESTED
WILLINGNESS TO ADAPT TO HP3000	NO	YES	YES	YES
WITHIN COST THRESHOLD	YES	TO BE DETERMINED	TO BE DETERMINED	YES
WITHIN TIME THRESHOLD	YES	TO BE DETERMINED	TO BE DETERMINED	YES
CUSTOMER SUPPORT	NONE	—	CONSULTATION	MAINTENANCE/ CONSULTATION
TRANSFERABILITY*	5	3	4	2

†Note The data presented in this table has been supplied by the individual organizations and has not been verified by Computer Sciences Corporation or NASA.

Table 4. continued

STANDARD REPORTING FORM CATEGORIES	ORGANIZATION AND SYSTEM ACRONYM			
	EARTH SYSTEMS RESEARCH, INC EPPL/4	EARTH RESOURCES LABORATORY GDBS	DAMES & MOORE GIMS	U S G S GIRAS
PROGRAMMING BASIS & OPERATING INFORMATION				
OPERATIVE COMPUTERS	CDC CYBER 74	VARIAN V–75	CDC6000, IBM 370, UNIVAC 1108 PDP-11	IBM 370
PROGRAMMING LANGUAGE	FORTRAN	FORTRAN IV & ASSEMBLER	FORTRAN IV	FORTRAN IV
MODE OF USAGE	INTERACTIVE & BATCH	BATCH & INTER-ACTIVE	BATCH	BATCH
MEMORY SIZE	40 KW	64K	64K	–
WORD SIZE (BITS)	60	16	32	32
GEOGRAPHIC DATA TYPE				
INPUT				
LINE	NO	NO	YES	NO
CELL	YES	YES	YES	NO
TABULAR	YES	NO	YES	NO
POLYGON	NO	YES	YES	YES
ANALYSIS				
CELL	YES	–	–	–
POLYGON	NO	–	–	–
TABULAR	YES	NO	YES	NO
CELL & POLYGON	NO	YES	YES	YES
DATA ENTRY & DATA OUTPUT PRODUCTS				
ENTRY				
AUTOMATIC	YES	YES	YES	NO
SEMI AUTOMATIC	NO	YES	NO	YES
MANUAL	NO	YES	YES	YES
OUTPUT PRODUCTS				
GRAPHIC	YES	YES	YES	YES
TABULAR	YES	YES	NO	YES
DIGITAL	YES	–	NO	YES
ANALYTIC CAPABILITIES				
COMPOSITE MAPPING	YES	–	YES	YES
POLYGON OVERLAY	–	–	YES	YES
CELLULAR	YES	–	NO	YES
ABILITY TO VARY SCALE	YES	YES	YES	YES
ABILITY TO VARY RESOLUTION	–	–	–	YES
AREA MEASURE	YES	ACREAGE PROGRAM	YES	YES
SIMULATION AND/OR MODELING	YES	NO	YES	NO
BOOLEAN COMBINATIONS	YES	–	YES	NO
CORRELATION	YES	NO	YES	NO
REGRESSION	YES	NO	NO	NO
INTERPRETIVE MAPS	YES	YES	YES	NO
DATA STORAGE				
STRUCTURE				
DIRECT ACCESS	YES	YES	NO	YES
SEQUENTIAL	NO	YES	YES	YES
OTHER	N/A	N/A	N/A	N/A
ORGANIZATION				
HIERARCHICAL	NO	NO	NO	NO
POINTER	YES	YES	NO	YES
RELATIONAL	NO	YES	YES	NO
INTERFACE WITH CLASSIFIED LANDSAT DATA				
EXPERIMENTALLY	–	–	NO	NO
OPERATIONALLY	YES	YES	NO	NO
ACQUISITION CONDITIONS	HANDLING/MAILING	–	–	SPECIFIC USERS ONLY
STATUS OF RELEASE	TESTED	–	TESTED	TESTED
WILLINGNESS TO ADAPT TO HP3000	YES	–	YES	NO
WITHIN COST THRESHOLD	YES	BASELINE SYSTEM – YES	UNKNOWN	BASELINE SYSTEM – YES
WITHIN TIME THRESHOLD	NO	–	YES	–
CUSTOMER SUPPORT	MAINTENANCE & CONSULTATION	–	MAINTENANCE & CONSULTATION	–
TRANSFERABILITY*	4	3	5	5

*Note The data presented in this table has been supplied by the individual organizations and has not been verified by Computer Sciences Corporation or NASA.

Table 4. continued

STANDARD REPORTING FORM CATEGORIES	ORGANIZATION AND SYSTEM ACRONYM			
	ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE			COMPUMATRIX, INC. EPIC
	GRID II	PIOS	AUTOMAP/GRIPS	
PROGRAMMING BASIS & OPERATING INFORMATION				
OPERATIVE COMPUTERS	HP3000, VARIAN TOSBAC ICL, IBM, PRIME, UNIVAC CDC	IBM 360, IBM 370, PRIME	BURROUGHS ICL, IBM, PRIME UNIVAC, CDC TOSBAC VARIAN, HP3000	HONEYWELL 6000, IBM 360, 370'
PROGRAMMING LANGUAGE	FORTRAN IV	FORTRAN IV	FORTRAN IV	BASIC, ALGOL, FORTRAN IV
MODE OF USAGE	BATCH & INTER-ACTIVE	BATCH & INTER-ACTIVE	BATCH & INTER-ACTIVE	INTERACTIVE
MEMORY SIZE	64K	64K	32–50K	VARIABLE (16K)
WORD SIZE (BITS)	16 OR LARGER	16 OR LARGER	16 OR LARGER	VARIABLE
GEOGRAPHIC DATA TYPE				
INPUT				
LINE	YES	YES	YES	YES
CELL	YES	YES	YES	YES
TABULAR	YES	NO	YES	YES
POLYGON	YES	YES	YES	YES
ANALYSIS				
CELL	YES	NO	–	–
POLYGON	NO	YES	–	–
TABULAR	YES	YES	YES	NO
CELL & POLYGON	NO	NO	YES	YES
DATA ENTRY & DATA OUTPUT PRODUCTS				
ENTRY				
AUTOMATIC	YES	YES	YES	YES
SEMI AUTOMATIC	YES	YES	YES	NO
MANUAL	YES	YES	YES	YES
OUTPUT PRODUCTS				
GRAPHIC	YES	YES	YES	YES
TABULAR	YES	YES	YES	YES
DIGITAL	YES	YES	YES	YES
ANALYTIC CAPABILITIES				
COMPOSITE MAPPING	YES	YES	YES	YES
POLYGON OVERLAY	NO	YES	YES	POLYGON INTER-SECTION
CELLULAR	YES	NO	YES	–
ABILITY TO VARY SCALE	NO	YES	YES	YES
ABILITY TO VARY RESOLUTION	YES	YES	YES	–
AREA MEASURE	YES	YES	YES	YES
SIMULATION AND/OR MODELING	YES	YES	YES	YES
BOOLEAN COMBINATIONS	YES	YES	NO	YES
CORRELATION	NO	YES	NO	YES
REGRESSION	NO	YES	NO	YES
INTERPRETIVE MAPS	YES	YES	YES	NO
DATA STORAGE				
STRUCTURE				
DIRECT ACCESS	NO	NO	NO	USER OPTION
SEQUENTIAL	YES	YES	YES	USER OPTION
OTHER	N/A	N/A	N/A	USER OPTION
ORGANIZATION				
HIERARCHICAL	NO	NO	NO	USER OPTION
POINTER	YES	YES	NO	USER OPTION
RELATIONAL	NO	YES	YES	USER OPTION
INTERFACE WITH CLASSIFIED LANDSAT DATA				
EXPERIMENTALLY	YES	NO	NO	YES
OPERATIONALLY	NO	NO	NO	NO
ACQUISITION CONDITIONS	LEASE	LEASE	LEASE	LEASE
STATUS OF RELEASE	TESTED & EVALU ATED	TESTED & EVALU-ATED	TESTED & EVALU-ATED	–
WILLINGNESS TO ADAPT TO HP3000	HAS BEEN ADAPTED	YES	HAS BEEN ADAPTED	YES
WITHIN COST THRESHOLD	YES	YES	YES	BASELINE SYSTEM – YES
WITHIN TIME THRESHOLD	YES	YES	YES	YES
CUSTOMER SUPPORT	MAINTENANCE/ CONSULTATION	MAINTENANCE/ CONSULTATION	MAINTENANCE/ CONSULTATION	MAINTENANCE/ CONSULTATION
TRANSFERABILITY*	1	3	1	3

†Note The data presented in this table has been supplied by the individual organizations and has not been verified by Computer Sciences Corporation or NASA

Table 4. continued

STANDARD REPORTING FORM CATEGORIES	ORGANIZATION AND SYSTEM ACRONYM			
	JET PROPULSION LABORATORY IBIS	M&S COMPUTING IGDS	H. DELL FOSTER IGS/330	EARTH SATELLITE CORP LCDMS
PROGRAMMING BASIS & OPERATING INFORMATION				
OPERATIVE COMPUTERS	IBM 360/65, 75, 91 IBM 370	PDP 11/34, PDP 11/70	DATA GENERAL 1200 ECLIPSE NOVA	PDP 11/45
PROGRAMMING LANGUAGE	FORTRAN IV/ IBM ASSEMBLER	FORTRAN & MACRO-11	FORTRAN/DATA GENERAL ASSEMBLY	FORTRAN IV, PLUS (DEC), MACRO-11
MODE OF USAGE	BATCH & INTER ACTIVE	INTERACTIVE	INTERACTIVE	INTERACTIVE
MEMORY SIZE	200K	256K	64K	64K
WORD SIZE (BITS)	32	16	16	16
GEOGRAPHIC DATA TYPE				
INPUT				
LINE	NO	YES	YES	NO
CELL	YES	NO	YES	YES
TABULAR	NO	NO	NO	YES
POLYGON	YES	YES	NO	YES
ANALYSIS				
CELL	NO	NO	–	YES
POLYGON	YES	YES	–	NO
TABULAR	YES	NO	YES	NO
CELL & POLYGON	NO	NO	YES	NO
ENTRY				
AUTOMATIC	YES	YES	NO	YES
SEMI-AUTOMATIC	YES	NO	NO	YES
MANUAL	NO	YES	YES	YES
OUTPUT PRODUCTS				
GRAPHIC	YES	YES	YES	YES
TABULAR	YES	YES	YES	YES
DIGITAL	YES	NO	NO	NO
ANALYTIC CAPABILITIES				
COMPOSITE MAPPING	YES	YES	–	–
POLYGON OVERLAY	–	POLYGON INTER SECTION	POLYGON INTER-SECTION	–
CELLULAR	–	NO	–	–
ABILITY TO VARY SCALE	YES	YES	YES	YES
ABILITY TO VARY RESOLUTION	–	YES	–	–
AREA MEASURE	YES	YES	YES	YES
SIMULATION AND/OR MODELING	NO	NO	NO	NO
BOOLEAN COMBINATIONS	NO	YES	NO	NO
CORRELATION	YES	NO	NO	NO
REGRESSION	NO	NO	NO	NO
INTERPRETIVE MAPS	NO	NO	NO	NO
DATA STORAGE				
STRUCTURE				
DIRECT ACCESS	YES	NO	YES	YES
SEQUENTIAL	YES	YES	NO	NO
OTHER	N/A	N/A	N/A	N/A
ORGANIZATION				
HIERARCHICAL	NO	NO	NO	NO
POINTER	NO	NO	YES	YES
RELATIONAL	NO	NO	NO	NO
INTERFACE WITH CLASSIFIED LANDSAT DATA				
EXPERIMENTALLY	–	NO	NO	–
OPERATONALLY	YES	NO	NO	YES
ACQUISITION CONDITIONS	ROYALTY PAYMENT	LEASE/LICENSE SALE	SPECIFIC USERS ONLY	–
STATUS OF RELEASE	TESTED & EVALU-ATED	TESTED	TESTED	TESTED
WILLINGNESS TO ADAPT TO HP3000	YES	NO	NO	YES
WITHIN COST THRESHOLD	N/A	NO	FOR TURNKEY SYSTEM – YES	YES
WITHIN TIME THRESHOLD	N/A	YES	FOR TURNKEY SYSTEM – YES	YES
CUSTOMER SUPPORT	CONSULTATION	MAINTENANCE CONSULTATION	MAINTENANCE/ CONSULTATION	MAINTENANCE/ CONSULTATION
TRANSFERABILITY*	4	5	5	2

†Note The data presented in this table has been supplied by the individual organizations and has not been verified by Computer Sciences Corporation or NASA

Table 4. continued

STANDARD REPORTING FORM CATEGORIES	ORGANIZATION AND SYSTEM ACRONYM			
	HARVARD UNIVERSITY ODYSSEY	OAK RIDGE NATIONAL LABORATORY ORRMIS	BATELLE PTOLEMY	SOUTH DAKOTA STATE PLANNING BUREAU SDAGIS
PROGRAMMING BASIS & OPERATING INFORMATION				
OPERATIVE COMPUTERS	PDP-10	IBM 360, PDP 10	PDP 11/35	IBM 370/145 (TEK 4051)
PROGRAMMING LANGUAGE	FORTRAN IV	FORTRAN (SOME IBM ASSEMBLY)	FORTRAN	FORTRAN IBM ASSEMBLER, BASIC
MODE OF USAGE	BATCH & INTERACTIVE	BATCH & INTERACTIVE	INTERACTIVE	BATCH & INTERACTIVE
MEMORY SIZE	45K	120–2000K	48K	18000K
WORD SIZE (BITS)	32	16 OR 32	16	32
GEOGRAPHIC DATA TYPE				
INPUT				
LINE	YES	YES	YES	YES
CELL	NO	YES	NO	YES
TABULAR	YES	YES	NO	NO
POLYGON	YES	YES	YES	YES
ANALYSIS				
CELL	NO	–	NO	–
POLYGON	YES	–	YES	–
TABULAR	YES	–	NO	NO
CELL & POLYGON	NO	YES	NO	YES
DATA ENTRY & DATA OUTPUT PRODUCTS				
ENTRY				
AUTOMATIC	YES	YES	NO	YES
SEMI-AUTOMATIC	YES	YES	NO	NO
MANUAL	YES	YES	YES	YES
OUTPUT PRODUCTS				
GRAPHIC	YES	YES	YES	YES
TABULAR	YES	YES	YES	YES
DIGITAL	YES	YES	NO	YES
ANALYTIC CAPABILITIES				
COMPOSITE MAPPING	YES	YES	YES	YES
POLYGON OVERLAY	YES	POLYGON INTERSECTION	POLYGON INTERSECTION	POLYGON INTERSECTION
CELLULAR	–	YES	NO	–
ABILITY TO VARY SCALE	YES	YES	YES	YES
ABILITY TO VARY RESOLUTION	YES	YES	–	–
AREA MEASURE	YES	YES	YES	YES
SIMULATION AND/OR MODELING	NO	YES	NO	YES
BOOLEAN COMBINATIONS	YES	YES	YES	YES
CORRELATION	NO	YES	NO	YES
REGRESSION	NO	YES	NO	YES
INTERPRETIVE MAPS	YES	YES	NO	YES
DATA STORAGE				
STRUCTURE				
DIRECT ACCESS	NO	YES	YES	NO
SEQUENTIAL	YES	YES	YES	YES
OTHER	N/A	N/A	NO	–
ORGANIZATION				
HIERARCHICAL	YES	YES	YES	NO
POINTER	YES	YES	YES	NO
RELATIONAL	NO	YES	NO	NO
INTERFACE WITH CLASSIFIED LANDSAT DATA				
EXPERIMENTALLY	NO	–	NO	–
OPERATIONALLY	NO	YES	NO	YES
ACQUISITION CONDITIONS	RENEWABLE LEASE		CONTRACT TO CONVERT TO HP	LEASE
STATUS OF RELEASE	PARTIALLY TESTED AND EVALUATED	PARTIALLY TESTED AND EVALUATED	IN HOUSE RESEARCH TOOL	TESTED
WILLINGNESS TO ADAPT TO HP3000	YES		YES	NO
WITHIN COST THRESHOLD	YES		NO	BASELINE SYSTEM-YES
WITHIN TIME THRESHOLD	YES		NO	BASELINE SYSTEM-YES
CUSTOMER SUPPORT	MAINTENANCE/CONSULTATION		MAINTENANCE/CONSULTATION	MAINTENANCE/CONSULTATION
TRANSFERABILITY*	2	4	3	4

†Note The data presented in this table has been supplied by the individual organizations and has not been verified by Computer Sciences Corporation or NASA

2.3 Systems Unit

Activity for this unit has been directed toward two tasks; 1) a remote terminal installation at Jacksonville, and 2) planning a preliminary FRIS System. The following sub-sections deal specifically with these items.

2.3.1 JAX Remote Terminal

During this quarter agreement on the design for the initial remote terminal configuration between Purdue/LARS and St. Regis in Jacksonville, Florida was reached. This configuration is a modified version of previous options considered. Since St. Regis already has an IBM 3776 remote job entry terminal, it will be used to communicate with the Purdue/LARS computer at scheduled times or when not connected to the St. Regis National Computer Center in Dallas, Texas. This terminal has a card reader, dual-drive diskette storage and a printer. Job control cards for the Purdue/LARS computer could be entered into a file on the diskette storage or key-punched on cards. These control cards could then be submitted to the computer from the IBM 3776 terminal by designating the appropriate batch machine parameters on the initial cards. However, primary use of this terminal is anticipated to be for receiving printer files on the IBM 3776 printer.

Preparation of most job control files and initiation of job execution will usually take place from a Decwriter LA36 typewriter terminal. Both terminals will communicate with the Purdue/LARS computer via a telephone line and two 4800 bps modems, one at each location. The Decwriter terminal will operate through a secondary (reverse) channel in the modem at 110 bps. The telephone line is currently scheduled for installation by October 27, 1978. This may have to be changed if some used modems are not located soon. The modem companies are quoting three to four months deliver time on new modems. We were hoping to obtain used ICC modems from Racal/Milgo but they no longer have in stock the 4800 bps modem with a secondary channel which we need. Recently we learned that one of our present remote terminals will disconnect within a month and the availability of these modems is being ascertained. Should all attempts to locate a used modem fail, the terminal installation will be rescheduled for the first week of January, 1979.

Figure 8 illustrates the terminal hardware configuration we are working toward. St. Regis is responsible for providing the two terminals and a modem selector switch to connect the IBM 3776 batch terminal to the desired computer. Purdue/LARS is ordering the telephone line, two modems and ports into the IBM 3705 communications controller at LARS.

2.3.2 Preliminary System Design

Preliminary system design work began in ernest during this quarterly period. Within the project structure a system design group has been identified. This group has the task of addressing the FRIS computer requirements. The group is composed of personnel from; St. Regis Corporate Offices, The Corporations Computer Center, Southern Timberlands Division, and LARS.

The group's first meeting was at the St. Regis National Computer Center in Dallas, Texas. The day-and-a-half session was held during the latter part of August. The purpose of this meeting was to:

A. Acquaint the National Computer Center with FRIS and its impact on the St. Regis data processing activity.

B. Acquaint staffs within each organization that would be involved in the System Transfer phase.

C. Review the options relating to the JAX-LARS remote terminal link.

D. Identify actions relative to development of a preliminary system design and establish a time table.

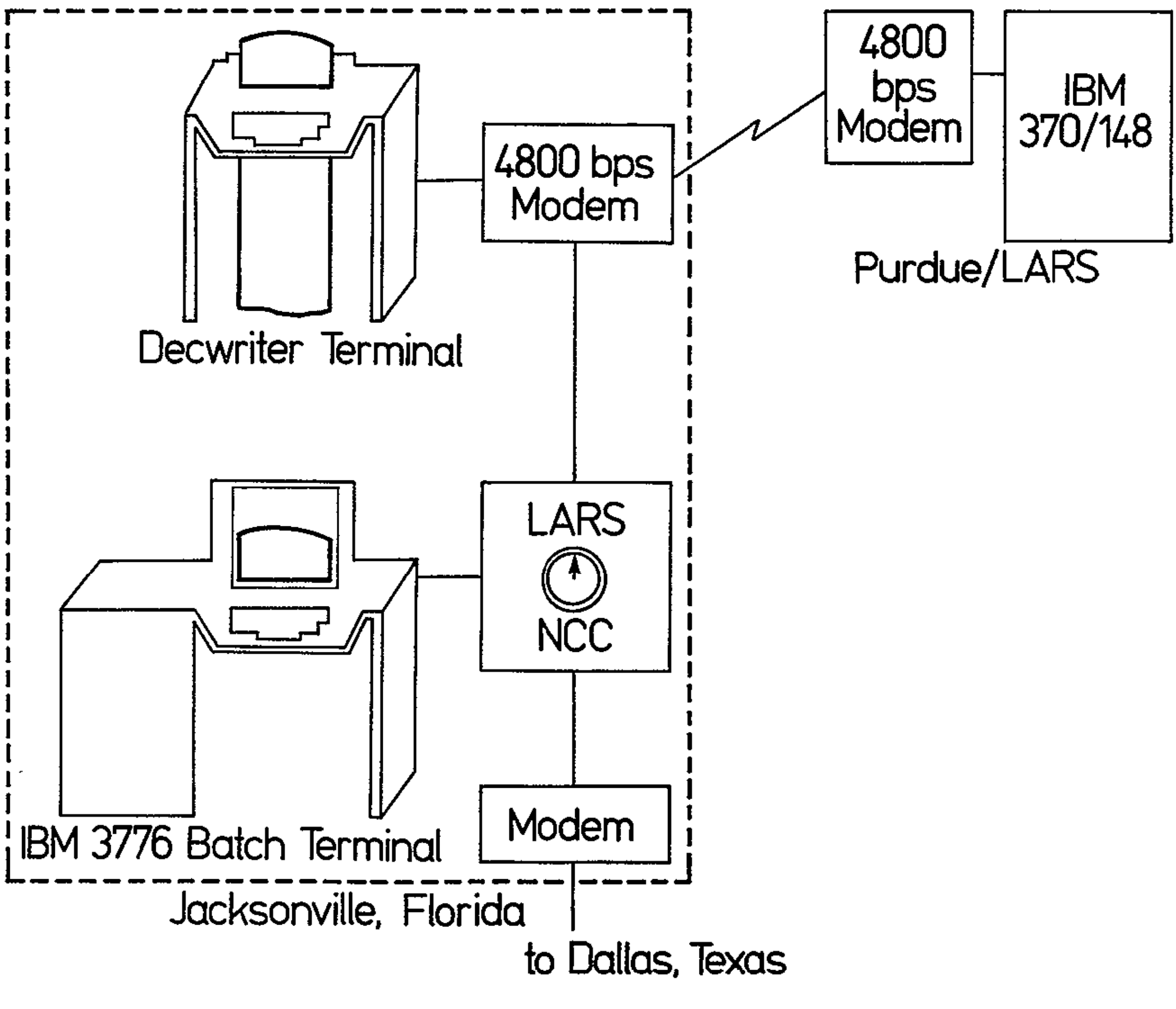


Figure 8. Jacksonville remote terminal hardware configuration.

A number of briefings were given; covering the FRIS Project, the physical basis of remote sensing, the future outlook for computing within St. Regis, and the computational requirements necessary to support LARSYS. There was also a detailed discussion on the various considerations necessary to implement a data base. General considerations revolve about; 1) the form of the data input, 2) types of data manipulation desired, and 3) the types of output products needed.

Growing from the above discussion a committee was formed to develop the FRIS Preliminary System Design. The primary responsibility given to this committee was to assess the various data base and image processing software that is commerically available that would meet the FRIS objectives. As much information as possible would be collected and presented to the group on 1 November 1978 in order to explore alternatives and costs. This information is a prerequisite to help develop an implementation schedule which will be necessary in order to move into the Phase III System Transfer task.

Prior to the 1 November meeting, LARS Staff would develop a number of straw-man system proposals. These proposals would range across a broad gament of capabilities from nothing more than a remote job entry station upwards to a corporate remote sensing facility.

Items which would be considered during the development of these straw-man proposals would include:

A. Communications Network

- identify locations between which information would be expected to flow.

B. Resource Requirements

- identify the system components which include:

  Hardware

  Software

  Man-power

C. Costs

- financial requirements to include both start-up and operational costs.

D. Documentation

- define the level of software and user documentation necessary for the system.

E. Transferability

- addresses the ease which the technology can be transferred, and implemented at St. Regis.

F. Languages

- identifies software programming languages.

G. Interface

- describes how the user would utilize the system.

As complete a definition as possible would be provided for the above items prior to the next committee meeting. Through a review of these proposals we hope that the committee will form a consensus regarding the preliminary design for FRIS.

2.4 Cost Unit

Based on a partially completed review of the literature about the economics of information and information systems, the following concepts are considered pertinent in the evaluation of the FRIS project. First in decision and information theory a clear distinction is made between information and data. Information is the attributes of data which make an impact on or influence decision making. Data is a collection of facts and figures which have not been analyzed and/or arranged in an useful order. This distinction is important because the value of FRIS is not in the data collection phase, but in the development of information used by managers at all levels of the firm in decision making.

2.4.1 Value of Information

The value of information is, therefore, the usefulness of analyzed and sorted data in improving the decision making of managers. Three components of value can be identified and assessed to determine the value of the information system. They are:

1) the relevance of the information provided to the decisions to be made,
2) the timeliness of the information, and
3) the accuracy of the information.

Relevance is the degree to which appropriate information is made available for decision making. While seemingly obvious that only information which is relevant would be provided to the decision makers, all information from a data base should be reviewed in the light of this criterion. Since any information created for its sake only is a misallocation of manpower and equipment. For the current project the relevance question has been addressed in the Southern Timberlands Division's "Forest Resource Information System - The Rational and Approach, Who Needs a FRIS."

Timeliness is an important yet difficult component of value to evaluate. The value of timely information is equal to cost savings in reducing the time a manager spends when making a decision, to competitive advantage which may be gained over other firms because of more rapid decision making, and to the fact that information decreases in value with respect to time.

The accuracy of information involves the degree, if any, of biasiness and the amount of variance or uncertainty surrounding the information. If bias is known it can be corrected and the information derived is not affected. If bias is unknown it is assumed to not exist and the information derived is not affected. Bias may arise in the statistical manipulation of the data during the collection and analyzing phases of the information system. The existence of bias is usually determined from statistical theory and should be identified by the data analyst and corrected for during data processing.

Variance or uncertainty about the information arises from a variety of sources. One may be the sampling procedure used in the forest inventory. Another may be the uncertainty about future events or currently held data (timeliness) and projection methods. In all cases the results of variance-uncertainty is to reduce the decision maker's confidence in the information received and his use of that information. If information is not used it has no value. Therefore, reduction of variance/uncertainty can increase the value of information by making the decision maker more confident in the information and hence making the information useful.

2.4.2 Measuring the Value of Information

Relevance is assumed to be satisfied by the report referenced earlier. Therefore the value of information is assumed to be at a maximum and constant with respect to relevancy.

Timeliness can be measured by calculating the difference in profit earned by a firm when using a more timely information system as compared to a slower information system. The calculation of this difference is replete with many problems due to the various cost savings and added value which might occur. Some of these items are the reduction of time spent on routine decisions, the added value from time saved but expended on more difficult (less certain) decisions, the reduction of time in updating "old" information, the increased productivity stemming from a better understanding of real world situations due to the timeliness of the information. Thus, there are a great many cost savings and added values which may occur and a method of measuring and quantifying them has not yet been devised.

Accuracy which includes bias and variance may be measured and quantified to some degree. For the purpose of this demonstration we are concerned with measuring the effect that Landsat data has on the information flow. Precisely we are concerned with how to measure this effect within FRIS. Furthermore, we hope to attach a cost-benefit figure to this assumed value increment.

3.0 Summary

Accomplishments during this quarter have significantly advanced the overall FRIS Project Goal. Noteworthy among the many project achievements are:

- Implementation of a benchmark classification evaluation framework.
- Refinement and definition of FRIS preprocessing activities.
- A framework for developing a preliminary system.
- Identification of potential geo-based referencing systems as components of FRIS.
- Definition of a context in which to evaluate FRIS costs and potential benefits.

References

Barker, G. R., 1976.

"Project Plan - Applications System Verification and Transfer." Forest Resource Information System. 71 pp.

_______________, 1978.

FRIS information needs. Internal project memoranda.

_______________, 1978.

"Forest Resource Information System - The Rational and Approach, Who Needs a FRIS." Internal project communications. 8 pp.

Davis, B., 1978.

Benchmark Evaluation Criteria. Project review draft. 14 pp.

Hoffer, R. M. and Staff, 1975.

"Computer-Aided Analysis of Skylab Multispectral Scanner Data." LARS/Purdue University Information Note 121275.

Knapp, E., 1978.

"Geographic Information System Survey Interim Report." CSC/SD-78/6066, Computer Sciences Corporation. 37 pp.

LARS, 1977.

"Crop Identification and Area Estimation Over Large Geographic Areas Using Landsat MSS Data." LARS/Purdue University, Technical Report 022477.