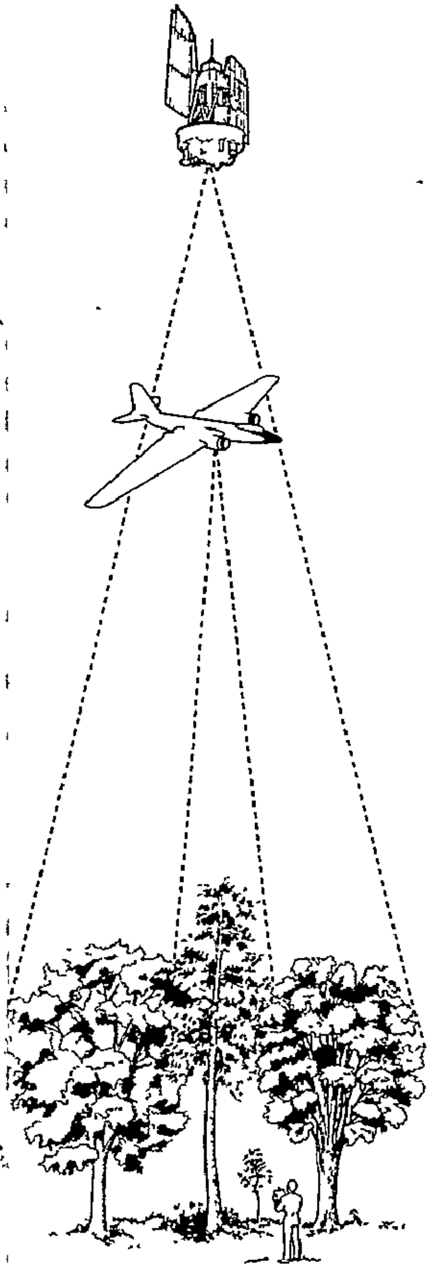
NATIONWIDE FORESTRY APPLICATIONS PROGRAM

(E78-10061) NATIONWIDE FORESTRY APPLICATIONS PROGRAM. TEN-ECOSYSTEM STUDY (TES) SITE 1, GRAND COUNTY, COLORADO Final Report (Lockheed Electronics Co.) 54 p HC A04/MF A01 CSCL 08F G3/43 N78-17434 Unclas 00061



TEN-ECOSYSTEM STUDY (TES) SITE I, GRAND COUNTY, COLORADO, FINAL REPORT

R. D. Dillman
Lockheed Electronics Company, Inc.
Systems and Services Division
Houston, Texas 77058

NAS 9-15200
LEC–10691
September 1977
Final Report (Type II) for
Action Document 63-1347-5325-12

Prepared for
EARTH OBSERVATIONS DIVISION
SPACE AND LIFE SCIENCES DIRECTORATE
LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS 77058

National Aeronautics and Space Administration

Lyndon B Johnson Space Center
Houston Texas 77058

FOREST SERVICE
U.S. Department of Agriculture

NOTE: In 1976, the Nationwide Forestry Applications Program was expanded from a Regional project by co-operative agreement between the Forest Service, U. S. Department of Agriculture, and the National Aeronautics and Space Administration (NASA). The Program is designed to sponsor research and development on the application of remote sensing analysis techniques to problems arising from the need to inventory, monitor and manage forests and rangelands, including the assessment of impacts on forest stands from insect and disease damage.

1 Report No LEC-10691	2 Government Accession No	3 Recipient's Catalog No

4 Title and Subtitle
TEN-ECOSYSTEM STUDY (TES) SITE I, GRAND COUNTY, COLORADO, FINAL REPORT

5 Report Date
August 1977

6 Performing Organization Code

7 Author(s)
R. D. Dillman

8 Performing Organization Report No
LEC-10691

9 Performing Organization Name and Address
Lockheed Electronics Co., Inc.
Systems and Services Division
Houston, Texas 77058

10 Work Unit No

11 Contract or Grant No
NAS 9-15200

12 Sponsoring Agency Name and Address
Earth Observations Division
Space and Life Sciences Directorate, NASA/JSC
Houston, Texas 77058

13 Type of Report and Period Covered
Type II

14 Sponsoring Agency Code

15 Supplementary Notes
Original photography may be purchased from:
EROS Data Center
Sioux Falls, SD

16 Abstract
This report presents the results, conclusions, and recommendations of the analysis which was carried out on Grand County, Colorado, Site I of the Ten-Ecosystem Study. Conclusions on best season, type mapping and area determination accuracies, recommendations on technical approaches, and classification output products are included.

ORIGINAL CONTAINS COLOR ILLUSTRATIONS

17 Key Words (Suggested by Author(s))
Ten-Ecosystem Study
Forestry
Grasslands
ADP Landsat

18 Distribution Statement

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

*For sale by the National Technical Information Service, Springfield, Virginia 22161

SUMMARY

This document presents the conclusions and recommendations resulting from the analysis which was carried out on a portion of Grand County, Colorado. The site (56 by 56 kilometers in size) is the first of 9 areas being studied as part of the Ten-Ecosystem Study being carried out by the Nationwide Forestry Applications Program.

The Ten-Ecosystem Study has been designed to apply the same set of automatic data processing analysis procedures, to test them, and to improve them over each of 10 general forest and rangeland ecosystems in the United States. The specific objectives are (1) to determine the feasibility of a nationwide study, (2) to identify and recommend solutions to specific site/ecosystem problems, (3) to determine type mapping accuracies achievable on Land Satellite multispectral scanner data, (4) to determine the best seasons, (5) to refine procedures for and recommend an automatic data processing system for a nationwide demonstration project, and (6) to transfer the technology to the U.S. Department of Agriculture, Forest Service.

The site (Site I in Grand County) was selected as being representative of the Coastal Range and Rocky Mountain Ecosystem and was analyzed using the IMAGE-100 system to identify softwood, hardwood, grassland, and water.

Three Land Satellite data sets consisting of August (ID 1388-17131), October (ID 1802-17023) Landsat frames, and a temporal data set (bands 2 and 4 for both August and October) were analyzed.

The following conclusions have been reached as a result of the work performed on Site I.

a. All the desired objectives can be accomplished with the Ten-Ecosystem Study technical approach and procedures at a cost of about 18 cents per square hectometer (7 cents per acre).

b. Mapping accuracies for Level II features were about 74 percent in terms of overall probability of correct classification using two methods to develop signatures.

c. Level II area determination is usually an underestimate. But a two-stage sampling scheme using regression estimates could be used to produce acreage estimates with less than 5 percent relative standard error.

d. Level III forest species separation was not possible. Aspect was a more dominant factor than species. Sagebrush and improved pasture were separable features in this site.

e. There was little difference between the August, October, and the temporal data sets for analysis of vegetation. The August data set did provide more consistent hardwood classification.

f. The accuracy, probability of correct classification, and total site classification were almost the same whether the separability or inventory signature were used. The difference was in the error for determining acreage where the inventory procedure produced higher error estimates.

g. The simulated inventory results indicate that signatures derived from only 10 percent of the study site can be used to estimate acreages from the entire site [56 by 56 kilometers (35 by 35 miles)] with a standard error of 3.1 percent for softwood which comprises 51.7 percent of the scene. This is an acceptable error with most current inventory techniques. A further reduction in the standard error could be achieved by using more primary sampling units.

This document is submitted in response to Job Order 75-325, Action Document 63-1347-5325-12. Distribution of this report to the National Aeronautics and Space Administration and the U.S. Forest Service has been approved by the manager of the Earth Observations Exploratory Studies Department and the manager of the Nationwide Forestry Applications Program.

CONTENTS

Section | Page
1. INTRODUCTION 1-1
2. STUDY SITE 2-1
3. TECHNICAL APPROACH 3-1
4. ANALYSIS RESULTS 4-1
4.1 PRELIMINARY SITE ANALYSIS 4-1
4.1.1 TASK I.5 — IMAGE EVALUATION 4-1
4.1.2 TASK II.1 — SITE FAMILIARIZATION. 4-5
4.2 MACHINE PROCESSING. 4-5
4.2.1 PREPROCESSING 4-5
4.2.2 PROCESSING. 4-8
4.2.2.1 Separability Results. 4-8
4.2.2.2 Simulated Inventory Results 4-9
4.2.3 TASK II.4 — POSTPROCESSING. 4-11
4.3 TASK II.5 — EVALUATION. 4-11
4.3.1 CLASSIFICATION MAP EVALUATION 4-11
4.3.2 ACREAGE EVALUATION. 4-13
4.3.3 REGRESSION ESTIMATION OF ACREAGE. 4-15
4.4 OUTPUT PRODUCTS 4-17
4.5 ADDITIONAL ANALYSIS 4-17
5. RESOURCE REQUIREMENTS. 5-1
6. CONCLUSIONS AND RECOMMENDATIONS. 6-1
6.1 CONCLUSIONS 6-1
6.2 ASSESSMENT OF TECHNICAL PROCEDURES. 6-4

Section | Page

6.3 RECOMMENDATIONS FOR FURTHER STUDY 6-4

7. REFERENCES 7-1

PRECEDING PAGE BLANK NOT FILMED

TABLES

Table		Page
1	HIERARCHY OF FEATURES CLASSIFIED FOR SITE I	2-4
2	PCC OF INTERPRETED LANDSAT DATA VERSUS AERIAL PHOTOGRAPHY	4-2
3	THREE-WAY COMPARISON OF VEGETATION CLASSIFICATION RESULTS	4-6
4	TRAINING FIELD CLASSIFICATION ACCURACIES FROM THE SEPARABILITY STUDY	4-10
5	LEVEL III CLASSIFICATION FOR SOFTWOOD AND GRASSLAND FROM THE AUGUST DATA SET	4-10
6	PCC AND CONFIDENCE INTERVAL FOR TWO CLASSIFICATION METHODS FROM THE AUGUST DATA SET	4-12
7	STATISTICAL ANALYSIS OF ERRORS (90-PERCENT CONFIDENCE LEVEL) IN CLASS PROPORTION ESTIMATION FOR SITE I USING THE SEPARABILITY AND SIMULATED INVENTORY CLASSIFICATIONS	4-14
8	REGRESSION ESTIMATE OF SOFTWOOD ACREAGE	4-15
9	SOFTWOOD AREA ESTIMATES AND COMPARISON FOR SITE I	4-16
10	ACCURACY OF SOFTWOOD TRAINING FIELDS CLASSIFIED BY ASPECT	4-22
11	RESOURCES UTILIZED	5-2
12	PROJECTED OPERATIONAL RESOURCES NEEDED TO PROCESS SITE I	5-4

FIGURES

Figure		Page
1	Site I (Grand County, Colorado) location.	2-2
2	Vegetation features of interest	
	(a) Spruce fir on steep slopes with wet meadow and alder bushes in valley bottom . .	2-5
	(b) Lodgepole pine on northern slopes and sagebrush on dry southern slopes	2-5
	(c) Aspen regeneration on southern slope following a burn	2-6
	(d) Irrigated pasture in background with sagebrush appearing on dryer slope in the foreground	2-6
3	A steropair of the aerial photograph used in imagery evaluation	4-3
4	Landsat image of Site I	4-4
5	Classification using separability signatures, segment 2	4-18
6	Classification using inventory signatures, segment 2	4-19
7	Classification using inventory signatures after applying the GETMIX/CLEAN algorithm	4-20
8	Example of map produced from black and white positive transparency	4-21

ACRONYMS AND ABBREVIATIONS

ADP	Automatic data processing
DAS	Data Analysis Station
ERIPS	Earth Resources Interactive Processing System
JSC	Lyndon B. Johnson Space Center
Landsat	Land Satellite
MSS	multispectral scanner
PCC	probability of correct classification; derived from the evaluations procedures and representing the number of correctly classified SSU's divided by the total possible SSU's.
Pixel	picture element
PSU	primary sampling unit; first stage of a two-stage sampling scheme which consists of an area 50 by 50 pixels located at random throughout the site.
RMS	root mean square; root of the arithmatic mean of a set of numbers. Used to evaluate the difference between measured and calculated point locations for image-to-image to ground registration.
SSU	secondary sampling unit; second stage of two-stage sampling scheme which consists of 2 by 2 pixel blocks located at random within each PSU.
TES	Ten-Ecosystem Study
USDA	U.S. Department of Agriculture
USFS	U.S. Forest Service

1. INTRODUCTION

This report documents the work performed on a study site in Grand County, Colorado, as part of the Ten-Ecosystem Study (TES). This site is one of nine being analyzed and was selected to be representative of the Coastal Range and Rocky Mountain Ecosystem (ref. 1).

The following sections present a description of the study site, the technical approach used, analysis results from four tasks (preliminary site evaluation, machine processing, evaluation of results, and output products), resources utilized and conclusions and recommendations.

The TES was designed as an automatic data processing (ADP) study using Land Satellite (Landsat) data, supporting aircraft imagery and ancillary information for performing a forest, grassland, and inland water inventory of chosen sites within ten ecosystems of the United States (ref. 2).

The primary objectives of the TES are to:

a. Investigate the feasibility of state-of-the-art ADP remote sensing technology to classify forest, grassland, and inland water areas.

b. Identify processing problems and recommend solutions that are specific for individual sites or ecosystems.

c. Recommend the definition and requirements of an integrated ADP system to support a nationwide forest and grassland remote sensing test project.

Secondary objectives include:

a. Determining type mapping accuracies at two general levels of hierarchy in the ecosystems.

b. Establishing the season or seasons that offer the greatest potential for type mapping in each ecosystem.

c. Providing the U.S. Department of Agriculture (USDA), Forest Service with project findings and conducting evaluation workshops.

This report presents the results of the study performed on the Grand County, Colorado, site, hereafter referred to as Site I. Individual site tasks have been reported previously with more detailed results in references 3, 4, and 5.

Site I represents an area of about 310 000 square hectometers (768 000 acres) in north central Colorado. The preliminary establishment of the two best seasons for ADP analysis was based on an evaluation of eight Landsat transparencies. The data processing consisted of two phases: the separability study and the simulated inventory study. The former was designed to establish the level of classification detail possible on Landsat multispectral scanner (MSS) data, and the latter was to determine how successfully ADP technology can extend limited ground truth for large area inventories. Classification results from both phases were analyzed statistically to determine map classification and feature proportion accuracy. Man and machine hours were recorded to help establish guidelines for future data processing planning. The site was initially classified into Level II features (refs. 2, 6) which included softwood, hardwood, grassland, water, and other. If accuracies from the separability study were greater than 90 percent for all features, except grassland which was 80 percent, Level III classifications were tried. Level III features include lodgepole pine, spruce fir, wet meadow, and sagebrush.

The TES is being conducted by the Nationwide Forestry Applications Program, a cooperative program of the USDA Forest Service and the Earth Observations Division of the National Aeronautics and Space Administration (NASA), Lyndon B. Johnson Space Center (JSC) in Houston, Texas.

2. STUDY SITE

The 56- by 56-kilometer (35- by 35-mile) study is located 70 kilometers (44 miles) west of Denver, Colorado, in the southeast corner of Grand County (fig. 1). The center coordinates are 39°58' N, 105°59' W. The site includes an intermontane basin known as Middle Park, the Fraser Experimental Forest, and about 80 percent of the Arapaho National Forest.

The weather is mostly influenced by easterly-moving cyclonic systems with the isopleths of precipitation and temperature oriented parallel to the east slope of the Rockies. In Grand County, the average warm season precipitation, which occurs from April through September, ranges from 25.4 centimeters (10 inches) of rain at low elevations to 40.64 centimeters (16 inches) at the upper elevations. The warm season temperatures average between 9° C and 13° C.

The Middle Park Basin contains elevations which range from 2250 to 3660 meters (7382 to 12 000 feet) with all drainage flowing to the west via the Colorado River. The drainage patterns are the result of the geologically recent gradual uplifting of the Rocky Mountains which allowed streams to cut through both the sedimentary and granitic rocks.

In general, the soils were formed during the late Pleistocene Epoch and Recent times and are very shallow. Annual rainfall and temperature patterns combined with parent materials determine present soil associations. In the valleys and lower elevations cryoborolls and cryaquolls are present with middle slopes having cryobaralfs and high elevations having cryumberls and cryaquepts (ref. 7). The soils and vegetation are zoned by altitude.

NASA-S-77-10119

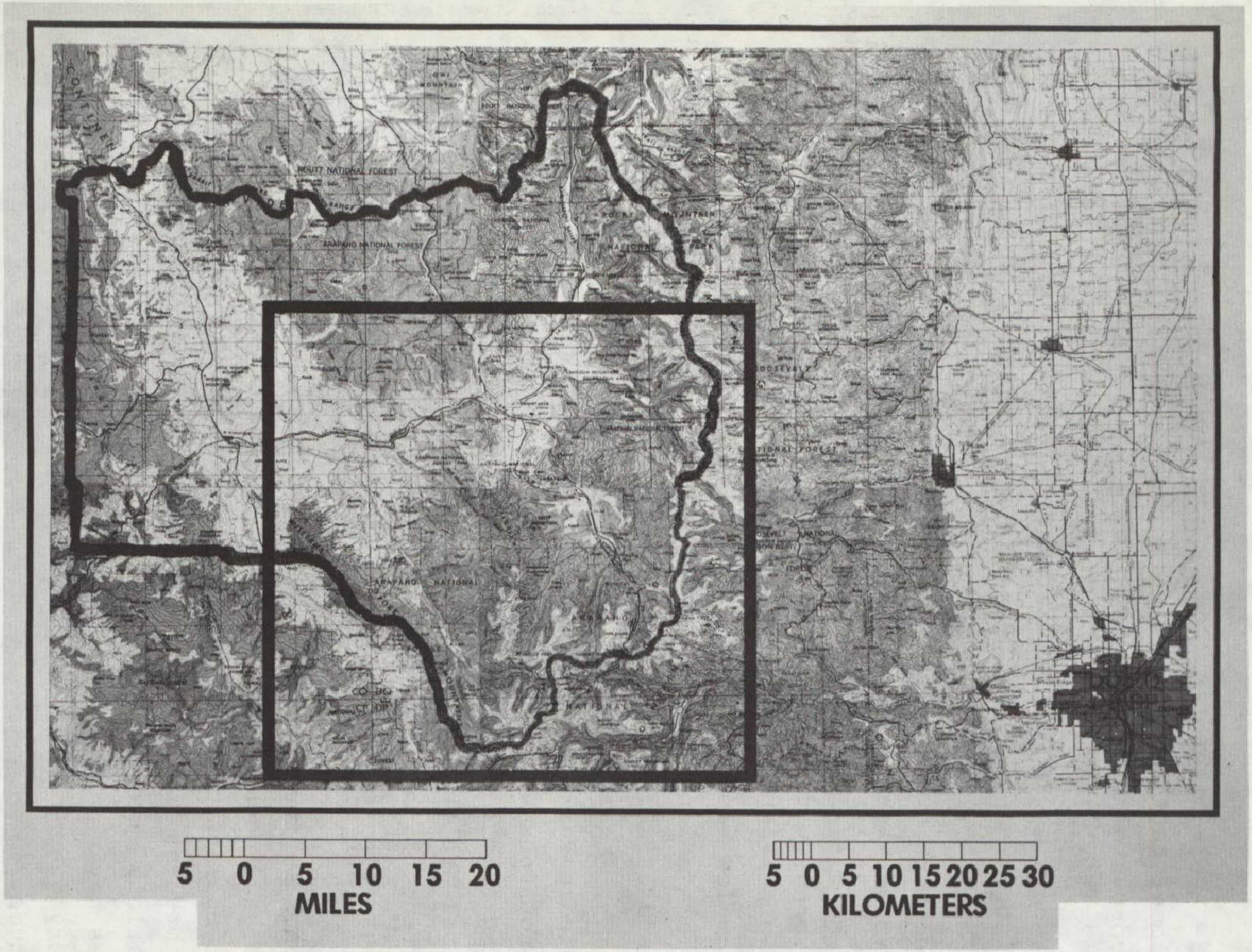


Figure 1.— Site I (Grand County, Colorado) location (outlined area).

ORIGINAL PAGE IS OF POOR QUALITY

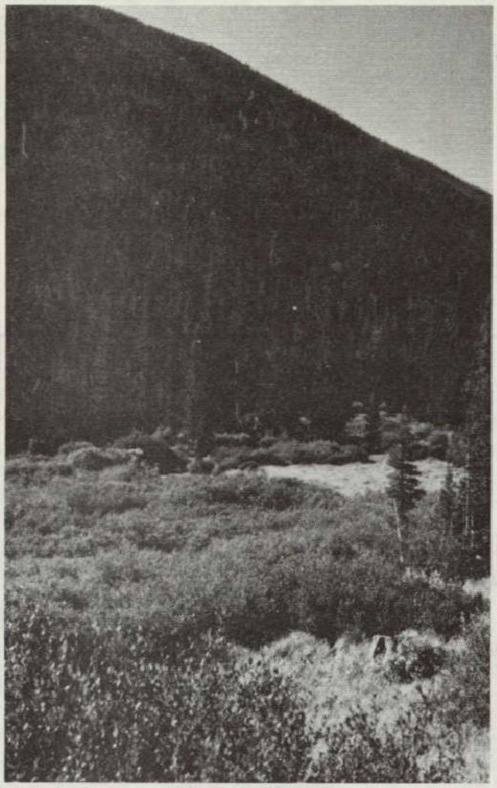
Vegetation in the Rocky Mountains is generally zoned by altitude in response to the temperature and moisture. The Middle Park Basin area on the western slope receives more moisture in the form of rain and snow than the eastern slope. The increased moisture has caused greater weathering of rocks with corresponding deeper soil developed. Consequently, vegetation growth is more continuous, with less obvious zonation.

The tree species found in the study site include: lodgepole pine (*Pinus contorta*), spruce (*Picea engelmannii, Picea pungens*), fir (*Abies Lasiocarpa*), and aspen (*Populus tremuloides*). These three forest types grow in the 2440 to 3200 meters (8000 to 10 500 feet) altitude zone with the actual distribution controlled by temperature and moisture requirements. Lodgepole pine grows on the dryer and hotter southern slopes and is a pioneer species on burns and cuts. Spruce and fir develop at higher elevations and on northern slopes where temperatures are lower and soil moisture is more favorable. Aspen occurs in only small areas throughout the site and is usually a pioneer species coming in after burns, landslides, or other disturbances.

The other two vegetative features of the ecosystem include improved pasture and sagebrush, which occur respectively in valleys at elevations of 2440 to 2740 meters (8000 to 9000 feet) and on dry sites. Generally, sagebrush occurs in all areas where there is not enough moisture for growth of pasture or forests. Census water is defined as being at least 16 square hectometers (40 acres) in size or 183 meters (200 yards) in width. This distinction for water can be made as one of the post-processing steps (ref. 8). The hierarchy of features investigated for this site is presented in table 1. Figure 2 presents ground photographs of the vegetative features of interest.

TABLE 1.- HIERARCHY OF FEATURES
CLASSIFIED FOR SITE I

Level I	Level II	Level III
Forest	Softwood	Lodgepole pine Spruce and fir
	Hardwood	Aspen
Nonforest	Grassland	Pasture sagebrush
	Other	
Water	Water	Census
		Noncensus

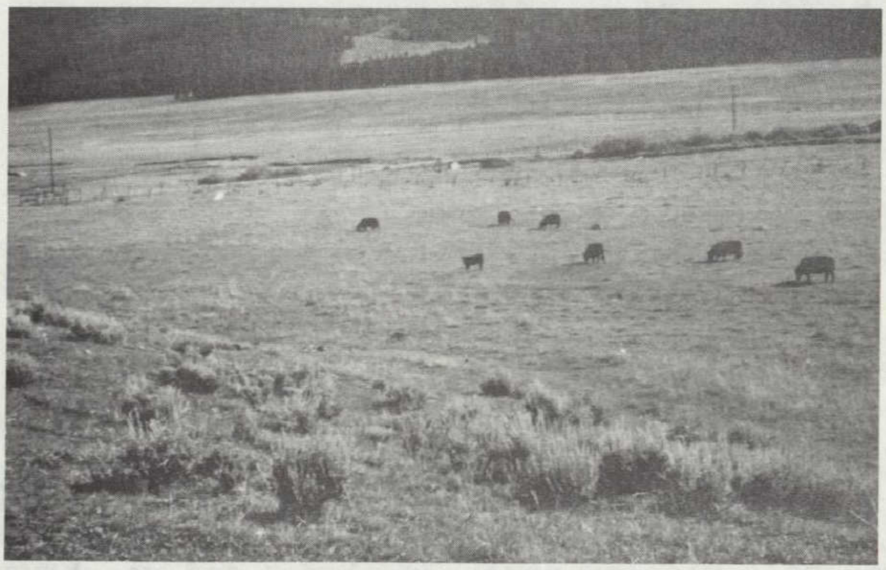


(a) Spruce fir on steep slopes with wet meadow and alder bushes in valley bottom.

(b) Lodgepole pine on northern slopes and sagebrush on dry southern slopes. (North is to the left in picture.)

Figure 2.— Vegetation features of interest.

(c) Aspen regeneration on southern slope following a burn. (Lodgepole pine is in foreground and background.)



(d) Irrigated pasture in background with sagebrush appearing on dryer slope in the foreground.

Figure 2.— Concluded.

3. TECHNICAL APPROACH

This section will summarize the approach followed in processing, analyzing, and evaluating aircraft and Landsat data collected over the study site.

The TES is composed of three phases: Phase I is the planning and data acquisition phase of the investigation; Phase II is the data reduction; and Phase III is the final analysis of and reporting on all nine sites.

Phase I has been documented previously in references 1 through 4 in which the objectives, rationale, sites selected, and procedures to be used for processing the Landsat data are discussed. This report summarizes the results presented in references 3, 4, and 5 and presents the conclusions and recommendations which are most significant as a result of this study.

Task I.5 was a photointerpretation analysis of Landsat color composite transparencies viewed at 1:200 000 scale versus high altitude color infrared photography viewed at 1:27 000 scale. This task had three objectives: (1) selecting the two Landsat data sets which provided the potential for the highest classification accuracy for two distinct phenologic dates; (2) providing the site scientists with a familiarization of the area; and (3) selecting potential training fields for use in ground checks during task II.1 (ref. 3, section 4.1.1).

Task II.1 consisted of a site visit and an analysis of the training sites selected (ref. 4, section 4.1.2). Task II.2 (preprocessing) consisted of selecting the study site, registering the two Landsat dates to each other, and registering the Landsat data to a map (ref. 4, section 4.2.1).

Task II.3 (processing) consisted of two investigations. The first was a separability study which used an analysis of training fields chosen from throughout the study area, to determine the level of vegetation detail which could be classified within the site. The second study was a simulated inventory study which relied on signatures from only a 10-percent area of the site to classify the entire study site (ref. 4, section 4.2.2).

Task II.4 (postprocessing) was designed to produce different examples of the possible classification output products at different scales. Also, primary sampling units (PSU's) were produced for the evaluation phase (ref. 5, section 4.4).

Task II.5 (evaluation) was a detailed comparison of ADP classification versus photointerpreted classification to arrive at overall map probability of correct classification (PCC) and to estimate the major class proportions/acreages for the site.

Task II.6 is report preparation which includes three previous reports (refs. 3 through 5) and this final site report.

Procedures were followed as described in reference 6, with the exception that a good comparison of class acreages was not possible because the site size was smaller than one county. The county is normally the smallest unit for which forest inventory statistics are available. The present evaluation method is a rigorous statistical evaluation which produces standard error and confidence qualifications. Conclusions drawn from such evaluation are as valid as conclusions drawn directly from acreage comparison even when the latter is possible.

4. ANALYSIS RESULTS

This section will summarize the results from Tasks I.5 and II.1 through II.5 which are documented in references 3, 4, and 5.

4.1 PRELIMINARY SITE ANALYSIS

The preliminary site analysis was covered by Tasks I.5 and II.1. The objectives were to select the two best seasons for ADP analysis and to provide for familiarization with the study site. This familiarization allowed the site scientist to better interpret the aerial photography used for ground truth during the evaluating process and to analyze training fields used in the separability study.

4.1.1 TASK I.5 — IMAGE EVALUATION

Photointerpretation of one frame of color-infrared photography (approximately 1:120 000 scale; fig. 3) was compared with the interpretation from the same area on eight different dates of Landsat color composites (1:1 000 000 scale; fig. 4).

Overall Landsat interpretation accuracy (PCC) compared to the high-altitude photography was calculated using 121 randomly selected points. The Level II classes of softwood, hardwood, grassland, water, and other were interpreted.

Table 2 presents the results of the photointerpretation analysis. The August 1973 acquisition date provided the highest accuracy in separating the classes of interest; therefore, it was selected for later ADP analysis. The second acquisition date to be used was in October 1974 because it had a high PCC and provided a second phenologic date for comparison in the temporal analysis (i.e., summer versus fall).

TABLE 2.- PCC OF INTERPRETED LANDSAT DATA VERSUS AERIAL PHOTOGRAPHY

Date	Image ID	PCC (a)	PCC_{HR} (b)	Remarks
8/15/73	1388 - 17131	73.39	87.1	
7/10/73	1352 - 17134	70.97	85.5	
10/3/74	1802 - 17023	67.74	84.7	
5/30/74	1676 - 17070	58.87	85.5	20% snow
1/22/73	1334 - 17135	55.65	83.9	20% snow
4/11/73	1262 - 17143	-		100% snow
2/16/73	1208 - 17142	-		80% snow
1/11/73	1172 - 17135	-		80% snow

[a]PCC for all classes.

[b]PCC for hardwood and grassland only.

LEGEND:

S = Softwood	Sg = Sagebrush	C = Cut
H = Hardwood	Rk = Rock	ALp = Alpine
GL = Grassland	Ur = Urban	Sn = Snow
W = Water		

Figure 3.— A stereopair of the aerial photograph used in imagery evaluation.

NASA-S-77-10102

Figure 4.– Landsat image of Site I. (Black and white reproduction of color composite, August 1973, 1388-17131)

TABLE 3.– THREE-WAY COMPARISON OF VEGETATION CLASSIFICATION RESULTS

	Correct Classification, %			
	Overall	Softwood	Hardwood	Other
Field check versus photointerpretation	88.2	100	80	86.6
Field check versus compartment maps	79.4	92.3	20	93.3
Photointerpretation versus compartment maps	76.4	92.8	14.3	92.3

had a root-mean-square (RMS) error of 1.4 pixels (about 95 meters). This RMS error is very good when one considers the large number of control points and the size of the area.

The eight-channel data tape was then registered to the ground by means of a least-squares-fit program on the PDP 11/45 computer. This ground registration used 12 control points which had an RMS error of 79 meters and 63 meters (259 feet and 207 feet; 1.3 and 0.8 pixels) along and across the scan lines, respectively. The rotation factor was then 0.080456, which represented a shift to the west of about 1 pixel for every 13 lines displayed.

Landsat pixels from the original data tapes normally represent a ground area of 57 by 79 meters (187 by 259 feet). The 57-meter (187-foot) dimension being along the scan line. In order to produce an output display on the Gould printer with known features in their correct relative position, the original data must be expanded north and south in order to produce square pixels, 57- by 57-meters (187- by 187-foot), within a given ground area.

Each segment is 485 pixels across the scan line and a least-squares-fit program indicated that the lines should be stretched 1.3 times. This means that 352 original Landsat lines are read into 485 lines by duplicating every fourth line displayed. This produces a digital data set which corresponds to a square ground area of 27 kilometers (17 miles) on each side.

Each pixel represents 0.330 square hectometers (0.816 acres) on the ground. Also, in order to fit the entire site onto the IMAGE 100 screen, a final data set of four 485- by 485-pixel segments was produced, representing 310 000 square hectometers (768 000 acres) of land.

Three data sets were created from the eight-channel registered tape for the separability study: (1) the four-channel August data, (2) the October data, and (3) the temporal data, consisting of bands 2 and 4 from both August and October. A dynamic range or gray-level count of 32, 32, 32, and 16 (for MSS bands 1, 2, 3, and 4, respectively) was selected for use in signature extraction, based on the observed signature variance and the appearance of the one-dimensional thresholded classification.

4.2.2 PROCESSING

Each data set was processed to determine the degree of separability of the classes of softwood, hardwood, grassland, water, and other. For Level II classes with accuracies above the established limits (see section 1), a more detailed classification level was investigated.

The data set which provided the greatest degree of separability was used for the simulated inventory study.

A total of 85 training fields were selected from throughout the study site, based on the ground-checked points, additional photo-interpretation, and other collateral data. Signatures for the same features were combined and any overlap was eliminated by a maximum likelihood decision program.

4.2.2.1 Separability Results

The accuracies produced by classifying the training fields with the composite signatures is presented in table 4. The difference in overall accuracies between dates was a maximum of 2 percent and the differences for vegetation features was on the order of 8 percent except for hardwood. Hardwood showed a high accuracy of 86.1 percent for August and a low of 36.6 percent for the temporal data. The August data set was selected for further analysis because of the higher classification accuracy of hardwood.

The Level II accuracies for softwood and grassland exceeded the established limits; therefore, a Level III classification was performed. The results of the Level III classification are presented in table 5. Improved pastures and sagebrush were separable from each other with accuracies of 84 to 98 percent, respectively. Level III softwood species had poor separability. When lodgepole pine and spruce fir signatures were used on training fields, most of the pixels were classified as spruce fir. This is because the training fields of spruce fir cover a larger range in signature variability than lodgepole pine.

4.2.2.2 Simulated Inventory Results

The August data set was used to determine how successfully ADP technology can extend limited ground truth for large area inventories. It was selected because of the high overall classification accuracy and the high accuracy for classifying hardwood (table 4). Training fields for classification were selected from an area representing about 10 percent of the study site outside of the field-checked areas. These were used to classify the entire area. Aerial photography (1:120 000 scale) was used to identify the training fields. An evaluation of the classification produced by these signatures is presented in section 4.3.1.

From a qualitative visual inspection of the inventory classification, it was anticipated that the PCC would be very low. The low PCC was expected because there was a large area of the site not classified and it was felt this would lead to high errors of omission and low PCC. Therefore, the study site was classified again using the composite signature from the separability study. This classification was also evaluated and the results are presented in section 4.3.1.

TABLE 4.— TRAINING FIELD CLASSIFICATION ACCURACIES FROM THE SEPARABILITY STUDY

Feature	Separability study, %		
	August	October	Temporal
Softwood	99.2	99.8	99.7
Hardwood	86.1	65.3	36.6
Grassland	91.9	96.9	93.5
Water	100	100	100
Overall Accuracy	97.0	98.0	96.1

TABLE 5.— LEVEL III CLASSIFICATION ACCURACY FOR SOFTWOOD AND GRASSLAND FROM THE AUGUST DATA SET

[Values may not add to 100 percent due to thresholding.]

	Lodgepole pine	Spruce fir	Improved pasture	Sagebrush
Lodgepole pine	9.0	82.7	(a)	(a)
Spruce fir	0	100.0	(a)	(a)
Improved pasture	(a)	(a)	84.1	1.6
Sagebrush	(a)	(a)	1.7	98.3

[a]Comparison not made because of previously demonstrated high separability.

4.2.3 TASK II.4 – POSTPROCESSING

Postprocessing involved two basic functions. First, classification tapes from the processing task were used to make examples of several output products. The products included color transparencies and Gould printer theme prints for production of lithograph copies of site maps at different scales (see section 4.4). Second, PSU's randomly selected from within the study area were printed on the Gould printer for use in the evaluation task (section 4.3).

4.3 TASK II.5 – EVALUATION

Reference 3 outlines a detailed procedure to be used to assess the overall accuracy of the classification and the capability to accurately estimate the acreage of different features from the map. One PSU[1] (50 by 50 pixels) and 10 secondary sampling units (SSU)[2] (2 by 2 pixels) were selected at random within the PSU. A total of 20 PSU's or 200 SSU's were analyzed for this site. The PSU's were registered to the aerial photography, and the area covered by the SSU was interpreted for comparison with the ADP classification. The registration process is a digital process where control points are selected and mathematically fitted by a least-squares analysis.

4.3.1 CLASSIFICATION MAP EVALUATION

The PCC of the map is a more detailed and more appropriate accuracy criterion than the training-field and test-field accuracy

---

[1] PSU cell size is 50 by 50 pixels, located at random from throughout the site. They represent the first stage of a two-stage sampling strategy. The two-stage sampling has the advantage of providing estimates of a given precision with less effort than a completely random sample.

[2] SSU cell size is 2 by 2 pixels located at random within each PSU. This is the second stage of two-stage sampling.

criterion traditionally used to evaluate ADP classification maps. The TES procedures are elaborate and represent an effort to evaluate the classification maps versus ground data on a 2- by 2-pixel basis selected randomly on the map. The evaluation method also eliminates any possible residual error resulting from misregistration of the PSU's. Table 6 presents the PCC for the classification using the separability study, the simulated inventory study, and the corresponding confidence interval.

Contrary to expectation, the PCC's of the classification maps derived from the two signature schemes were similar. A t-test[3] was performed and indicated that the 2-percent difference in PCC could be the result of chance and that the true difference could be zero.

TABLE 6.— PCC AND CONFIDENCE INTERVAL FOR TWO CLASSIFICATION METHODS FROM THE AUGUST DATA

Classification method	PCC, %	Confidence interval
Separability signatures	73	±4.8
Inventory signatures	75	±5.4

The confidence interval of ±5.4 percent was 0.4 percent wider than the ±5 percent criterion established by the procedures. The additional effort needed to reduce it to less than 5 percent was projected to be excessive and not worthy of evaluation.

[3]The t-test for paired plots is a standard statistical method to test the hypothesis that there is no difference in the means from two populations. In this instance, the calculated t-value (-0.438) was much smaller than the tabulated t-value (2.093) at the 95-percent level of significance for 19 degrees of freedom indicating that the hypothesis could be accepted (ref. 9).

4.3.2 ACREAGE EVALUATION

Another evaluation of map accuracy was the precision of acreage (proportions) estimation of a given feature, with a corresponding confidence interval. This accuracy measure has not received much attention in past remote sensing applications, but it is nonetheless a very important inventory parameter. In fact, this parameter would be more meaningful than the map PCC for many inventory projects.

Table 7 presents a tabulation of the errors in estimating class proportions for the two studies, based on an analysis of 20 PSU's. The table presents the true class proportion, p (considered as ground truth but actually estimated from the photography); the pixel-count class proportion, $\hat{p}$, from the ADP classification; the error of the estimated proportion, B; and the confidence interval of the error, 0.9, at the 90-percent confidence level.

In order to evaluate the errors, B, of the proportions, the confidence interval should be studied. If the range of the confidence interval includes zero, the error is inferred to be insignificantly different from zero. This evaluation indicates that the softwood from the separability signature classification constitutes over 51.7 percent of the scene, and its acreage can be estimated with almost no error. This is because the error is -3.8 (confidence interval of -0.085 to 0.009) which includes zero. However, the softwood proportion from the inventory classification is underestimated by 10.4 ± 4.3 percent.

4.3.3 REGRESSION ESTIMATION OF ACREAGE

To improve the estimate of classification acreage and to reduce the variance of the estimate, a linear regression analysis was performed using the estimated versus true class proportions for the softwood class. The F-test[4] was used to evaluate the

TABLE 7.— STATISTICAL ANALYSIS OF ERRORS (90-PERCENT CONFIDENCE LEVEL) IN CLASS PROPORTION ESTIMATION FOR SITE I USING THE SEPARABILITY AND SIMULATED INVENTORY CLASSIFICATIONS[a]

Classes	True class proportion,	Separability		Confidence interval, 0.9 level	True class proportion, p	Simulated inventory		Confidence interval, 0.9 level
		Estimated class proportion, $\hat{p}$	Error, B			Estimated class proportion, $\hat{p}$	Error, B	
Hardwood	0.001	0.02	-0.019	-0.026, -0.012	(b)	(b)	(b)	(b)
Softwood	.517	.555	.038	-.085, .009	0.531	0.4275	0.1035	0.0596, 0.1474
Grassland	.0055	.045	.0105	-.016, .037	.056	.055	.001	-.0227, .0247
Water	.0065	.0075	-.001	-.003, .001	.0065	.005	.0015	-.0009, .0039
Other	.42	.3725	.0475	.004, .091	.4063	.5125	-.106	-.143, -.069

[a]True class proportion (p) comes from photointerpretation, and estimated class proportion ($\hat{p}$) comes from pixel-counting of ADP classifications. The true class proportions are slightly different between the two classification methods. This results from the use of one set of SSU locations, for each PSU, in the inventory study and the use of a different set of random SSU's for each PSU in the separability study. The later procedure will be used in all subsequent evaluation in TES.

[b]Extensive hardwood sites did not occur in the area from which signatures were extracted and thus this class was not considered for this portion of the evaluation.

significance of the regression. The equation and coefficients are given in table 8.

Since the features of hardwood, grassland, and water constitute only small proportions of the scene, such an evaluation procedure and an acreage estimation formulation produce excessive variance; conclusive statements were not drawn. A different sampling strategy would have to be applied in order to adequately evaluate these features. Therefore, only the softwood acreage estimates will be analyzed.

TABLE 8.— REGRESSION ESTIMATE OF SOFTWOOD ACREAGE

	0.99-Level of significance	Coefficient of Determination, (R^2)	Regression equation	Standard error, %
Separability signature classification	Yes	0.87	$p = 0.75\hat{p} + 0.10$	2.3
Simulated inventory	Yes	.77	$p = 0.98\hat{p} + 0.10$	3.1

Table 9 presents a comparison between softwood acreages derived from Fraser Township, the softwood acreages of the whole site, and the adjusted acreages and confidence interval; the adjustment used the regression equation developed in the separability signature classification (i.e., the first equation in table 8). The basis for comparison in Fraser Township is a Forest Service

[4]The F-test is a standard statistical method to analyze the variance of several measurements to determine if there is a significant difference between the measurements. In the case of the regression, the F-test indicates what part of the variation in Y, which is explained by the fitted line, is significantly greater than the part that the line left unexplained.

TABLE 9.— SOFTWOOD AREA ESTIMATES AND COMPARISON FOR SITE I.

Softwood area estimates	Fraser Township, square hectometers (acres)	Study site, square hectometers (acres)
USFS stand map[a]	6748 (16 678)	--
Photointerpretation[b]	7004 (17 310)	--
ADP pixel counts[c]	6777 (16 751)	136 473 (337 304)
Corrected (regression) estimate	6050 (14 952)	133 806 (330 712)
Adjusted standard error of regression estimate[d]	233 (576)	7678 (18 977)
Confidence interval of regression estimate (0.9 level)[e]	5647, 6453 (13 957, 15 949)	120 531, 147 081 (291 902, 363 523)
Total area inside unit	9552 (23 607)	314 099 (776 318)

[a]USFS stand map over township was planimetered; stand map was more than five years old.

[b]Area planimeterdd off interpreted overlays of 1:120 000-scale color infrared photographs (not corrected geometrically); photographs were taken in September 1972.

[c]ADP in August 1973 Landsat data, using signatures developed in the separability study.

[d]Precise standard error depends on how different the area to be corrected is from the average area used in the regression analysis. Adjusted standard error (ASE) relates to standard error (SE) by:

$$(ASE)^2 = (SE)^2 \left(1 + n(P_o - \bar{P})^2 / \Sigma (P - \bar{P})^2\right)$$

where n = number of data points in regression
P_o = the area to be corrected
P = area (data value) of regression data
P = average of $\bar{P}$.

[e]Confidence interval = t(ASE), where t = t - value.

stand map and photointerpretation which were planimetered to determine the acreage of different features. Both of these sources have built-in errors associated with the system and method used to derive acreage. The standard errors of the estimates are very small, 233 square hectometers [(576 acres) i.e., 3.5 percent of the ADP count of 6777 square hectometers (16 751 acres)]. There was no good basis for comparison of the total softwood acreages for the entire study site.

4.4 OUTPUT PRODUCTS

The classification tapes from the study site were run through the DAS in order to produce color transparencies of the maps from the separability study (fig. 5), the inventory study (fig. 6), and the inventory classification after the program GETMIX/CLEAN was run (fig. 7). GETMIX/CLEAN was used to eliminate all groups of pixels less than 4 square hectometers (10 acres) for softwood, hardwood, and grassland, 0.4 square hectometers (1 acre) for noncensus water, and 16 square hectometers (40 acres) for census water. A Gould printer output of the separability classification was used to produce black and white positive transparencies for each class which were used to produce color lithograph maps of the study site (fig. 8). The vertical striping and lighter tone of the lower left hand quarter is a combined result of the variable output quality from the Gould printer, the high contrast transparencies, and the photographic process used to produce transparencies for the lithographic process.

4.5 ADDITIONAL ANALYSIS

In an effort to better understand the complexity of the site, the study site was classified and evaluated using the combined signature from the separability study. This effort was not required by the TES investigation plan but was performed and reported in section 4.3.

ORIGINAL PAGE IS
OF POOR QUALITY

NASA-S-77-24233

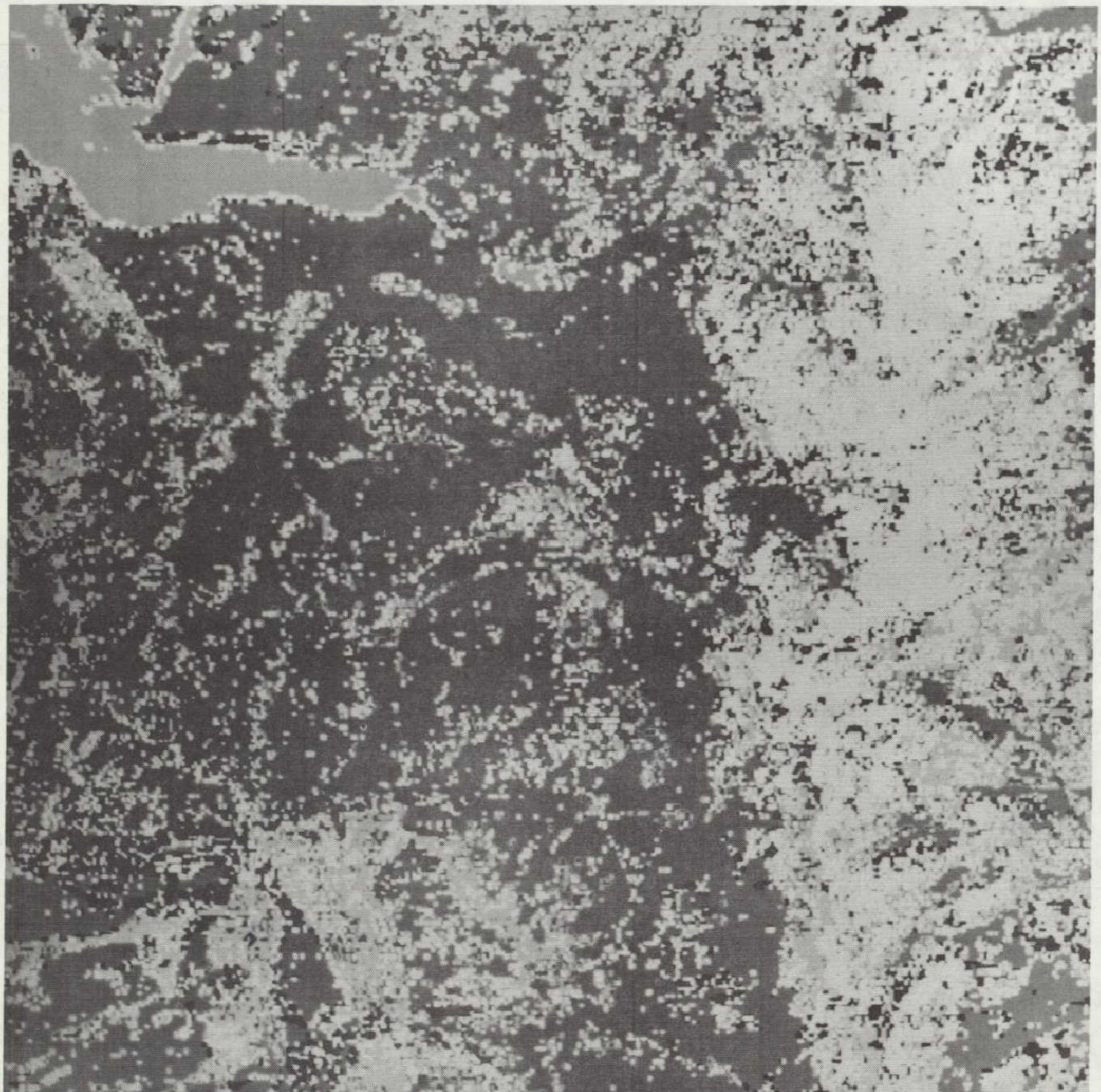


Figure 5.— Classification using separability signatures, segment 2.

NASA-S-77-24230

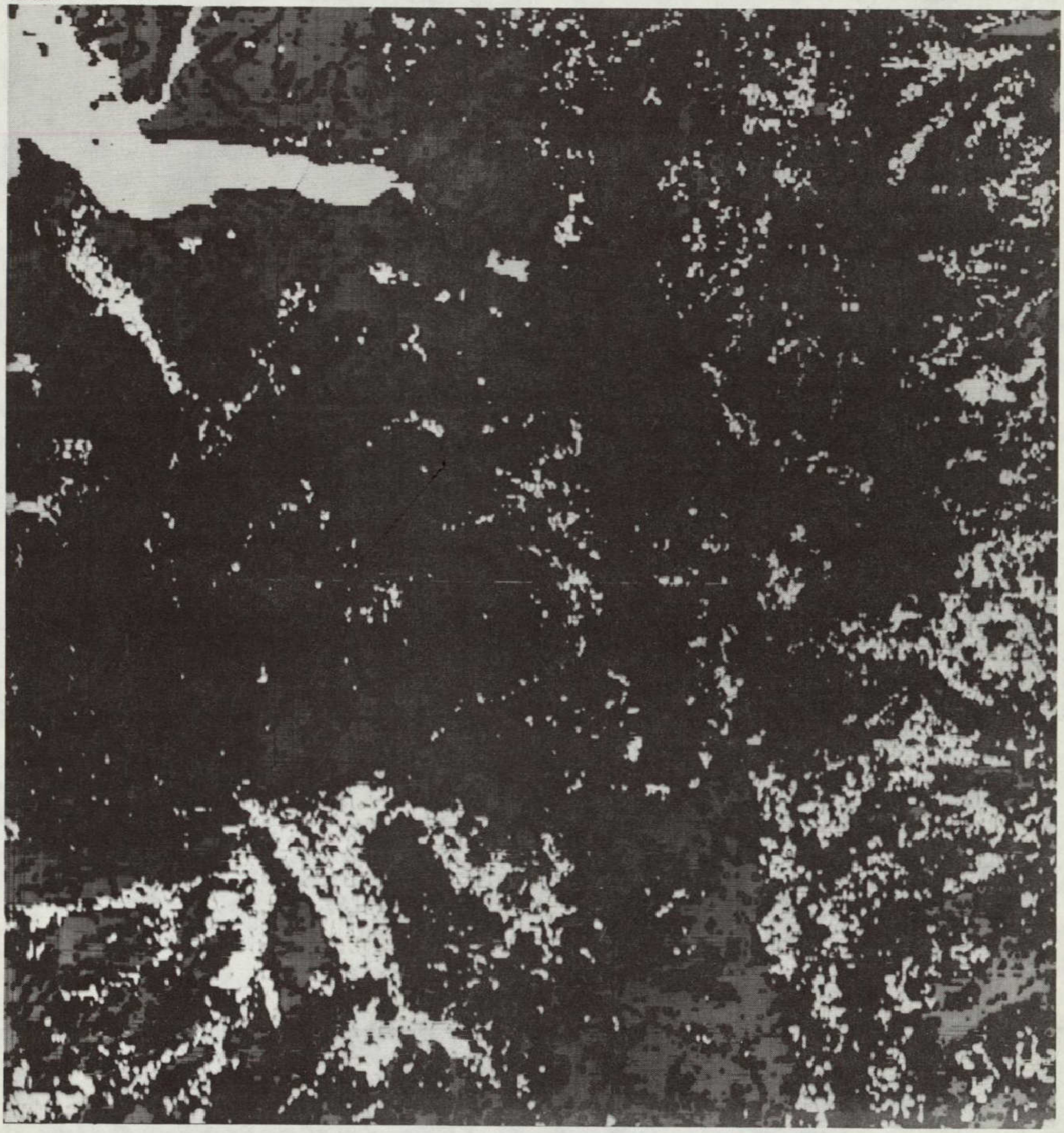


Figure 6.— Classification using inventory signatures, segment 2.

ORIGINAL PAGE IS
OF POOR QUALITY

NASA-S-77-24232

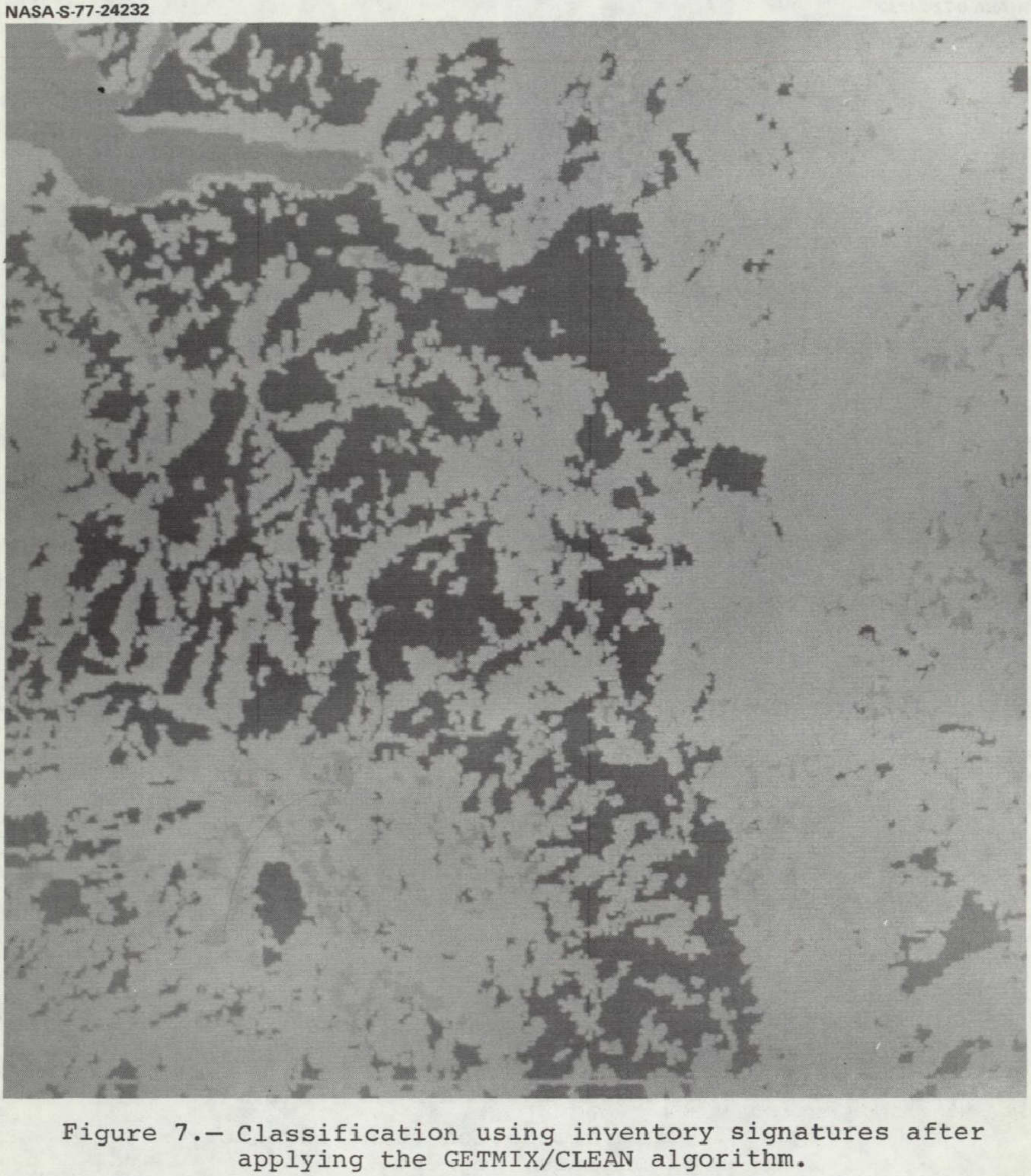
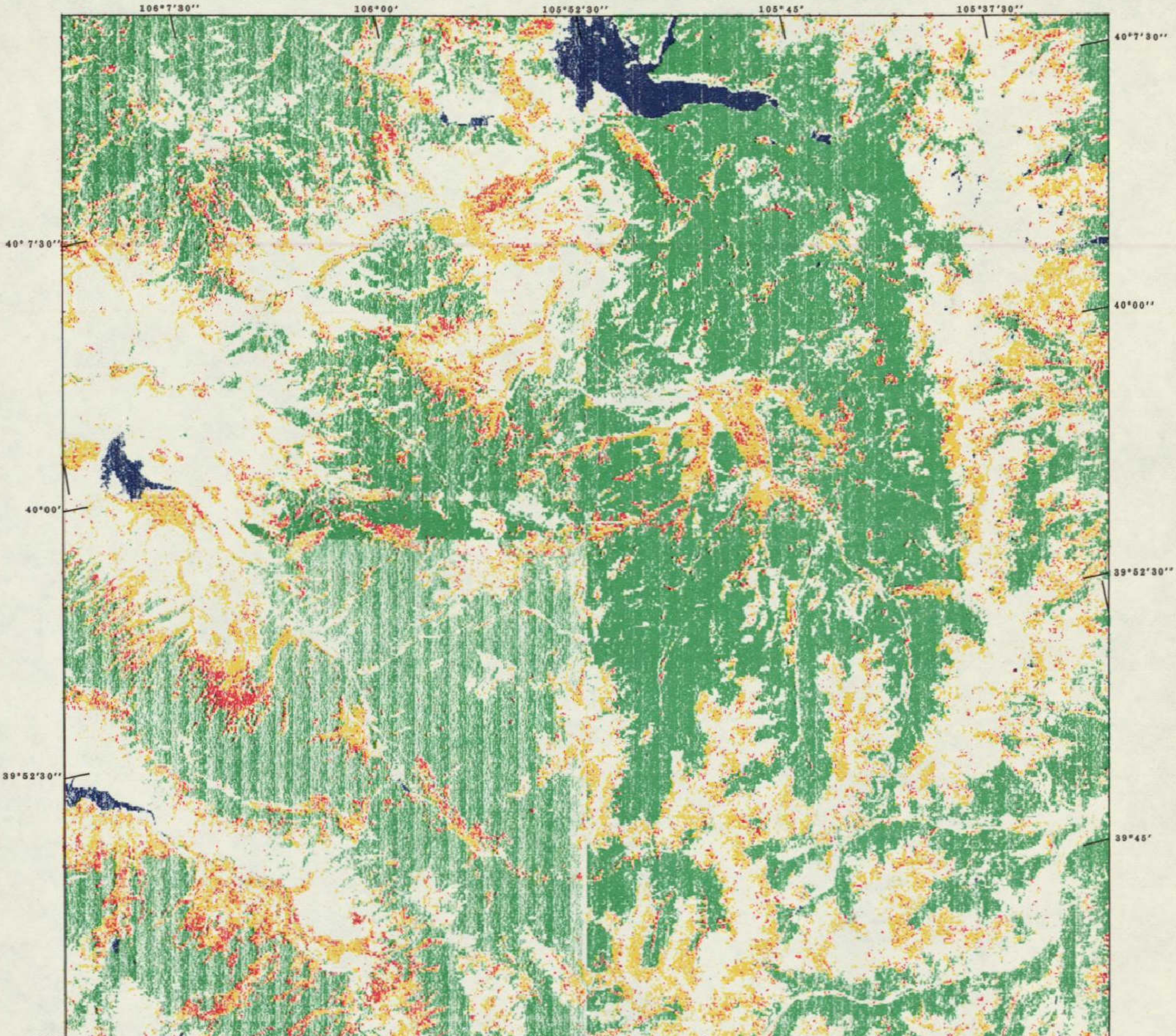


Figure 7.— Classification using inventory signatures after applying the GETMIX/CLEAN algorithm.

NATIONWIDE FORESTRY APPLICATIONS PROGRAM

TEN ECOSYSTEM STUDY
GRAND COUNTY, COLORADO



CLASSIFICATION FROM LANDSAT I
1388-17131 AUGUST 1973

LEGEND

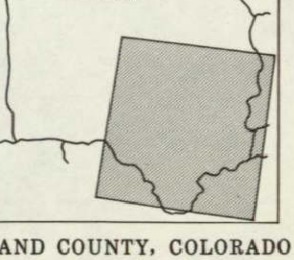
- SOFTWOOD
- HARDWOOD
- GRASSLAND
- WATER
- OTHER

N

1 0 1 2 3 4 5
MILES

1 0 1 2 3 4 5
KILOMETERS

LOCATION PLOT



GRAND COUNTY, COLORADO

PREPARED BY:
SFB. CARTOGRAPHIC TECHNOLOGY LABORATORY,
EARTH OBSERVATIONS DIVISION,
DS & LS JSC-NASA.
HOUSTON, TEXAS APRIL 1977

Also, a brief study of the effects of aspect on the separability of signatures was investigated. When the same softwood training fields, which were used for Level III classification, were grouped by aspect irrespective of species, the overall accuracy was 73.1 percent (table 10) compared to 53.2 percent for species classification. This indicates that aspect is a more controlling factor in areas of high relief than is species.

TABLE 10.— ACCURACY OF SOFTWOOD TRAINING FIELDS CLASSIFIED BY ASPECT

Aspect	Northwest, west	Northeast, east
Northwest, west, southwest	70.2	29.8
Northeast, east	71.2	88.8
	Overall Accuracy	73.1

5. RESOURCE REQUIREMENTS

The resources required for this site fell into three areas: site data (i.e., photographic imagery and Landsat imagery), manpower, and machine and equipment times. The per-area cost can be calculated for a given task if costs are assigned for data acquisition, man-hours, and machine time.

Site data used included eight color composite Landsat frames, two Landsat scenes in the form of computer compatible tapes, and aerial color infrared photographs from Mission 76-096, June 1976. The aerial photographs were used throughout the study for selecting the best Landsat date, training field verification, and evaluating the classification results. Aerial photographs, which covered only sections of Site I, from WB57F Mission 211, September 1972 and Mission 248, August 1973 were also used for comparison. These aerial photographs were at a scale of 1:120 000.

Table 11 presents a breakdown of the man-hours and machine hours used for each task in the study.

The actual times will probably be reduced for later sites as procedures are refined and FAP personnel become thoroughly familiar with the analysis. Over 23 percent of the man-hours for this site were spent reporting on progress and procedures. In a more operational mode, these times would be reduced and overall times would probably be lowered.

The precise cost per acre for the entire analysis is difficult to assess because it is difficult to account for all costs. In this case, an effort has been made to include all major technical, scientific, and clerical tasks from project initiation to completion. Added to this could be a proportional surcharge for program management time and overhead for facilities and employee

TABLE 11.— RESOURCES UTILIZED

Task	Total man-hours[a]	Machine hours					Cost/ sq hm, cent
		ERIPS	IMAGE-100	DAS	Dell Foster	Univac 1110	
Preliminary image analysis	200						
Site analysis	250						2.8
Preprocessing	243	16	36	10	10		6.2
Processing	348		54	6	6		6.5
Postprocessing	98		4	7	15	2.1	1.2
Evaluation	144						.5
Reporting (includes reports 1, 2, 3, and 4)	610						
Total	1893	16	94	23	31	2.1	17.8

[a]One to five personnel were utilized for each task.

benefits. Nevertheless, based on the following man-hour rates and rates estimated for government-owned research equipment, an average of 18 cents per square hectometer (7 cents per acre) was estimated.

Systems	Cost/hour	Hours	Total cost
IMAGE 100	$300	94	$28,200
ERIPS	300	16	4,800
Univac	300	2.1	630
PMIS DAS	100	23	2,300
Dell Foster	15	31	465
Total computer cost			$36,395
Man-hours	$10	1893	$18,930

Area costs could be reduced in an operational inventory system which was directed at producing acreage figures for a given class by administrative boundaries. Site analysis and familiarization would be eliminated because of previous area familiarization or the use of local Forest Service personnel. Preprocessing and processing man-hours would be reduced by at least 15 percent by having men and machines devoted totally to one project. Machine time would be reduced by at least 10 percent with operators performing standardized procedures. Reports which summarize the results could be reduced to one and man-hour costs could be reduced by using technicians with salaries on the order of $7 per hour. Table 12 presents the projected costs for processing Site I in an operational situation, with the final cost being 11.8 cents per square hectometer.

TABLE 12.– PROJECTED OPERATIONAL RESOURCES NEEDED TO PROCESS SITE I

Computer	Cost/hour	Hours	Total cost
IMAGE 100	$300	82	$24,600
ERIPS	300	10	3,000
Univac	300	2.1	630
PMIS DAS	100	21	2,100
Dell Foster	15	27	405
Total Computer Costs			$30,735
Man-hours	$ 10	787	$ 7,870

6. CONCLUSIONS AND RECOMMENDATIONS

This section presents the conclusions derived from the results of the processing of Site I, an assessment of how well the technical procedures performed and recommended changes for future processing systems, an assessment of the analysis problems related to site characteristics, and recommendations for further study.

6.1 CONCLUSIONS

The primary objectives of TES and the conclusions derived from Site I are as follows:

Objective A: Investigate the feasibility of state-of-the-art ADP remote sensing technology to classify forest, grassland, and inland water areas.

Conclusion: The IMAGE 100 system can classify Level II features with an overall mapping accuracy of about 74 percent ± 4.8 percent at a cost of 18 cents per square hectometer (7 cents per acre).

The simulated inventory results indicate that signatures derived from only 10 percent of the study site can be used to estimate acreages from the entire site [56 by 56 kilometers (35 by 35 miles)] with a standard error of 3.1 percent for softwood which comprises 51.7 percent of the scene. This is an accepable error with most current inventory techniques. A reduction in the standard error could possibly be achieved by evaluating more PSU's.

Objective B: Identify processing problems and recommend solutions that are specific for individual sites and ecosystems.

Conclusion: The major problems associated with analysis of data in Site I were related to its location in the Rocky Mountains and in an area of high relief.

The Rockies receive measurable amounts of snow beginning in November and ending in May. When snow is on the ground, extensive vegetation classification and evaluation is not possible. The seasons for data analysis are from May to October, about 6 months in which to get cloud-free data.

The high relief causes slope and aspect to be a major factor in controlling growth of some vegetation; more important, it causes variations in illumination. Slope and aspect affect accuracies for Level III classification more than species variation. However, at Level II, slope and aspect did not measurably affect class accuracies.

For more detailed analysis of forest features, some correction must be made for slope and aspect effects.

Objective C: Recommend the definition and requirements of an integrated ADP system to support a nationwide forest and grassland remote sensing test project.

Conclusion: By using the procedures defined for TES, classification accuracies of 74 percent ± 4.8 percent were obtained; acreage estimates with less than 5 percent relative standard error using a two-stage sampling scheme and regression estimates were produced when only 10 percent of the site area was used to extract feature signatures.

The final ADP system recommended for a nationwide test project must await complete processing on all

TES sites and a definition of the objectives of a nationwide test project by the Forest Service.

The secondary objectives of TES and the conclusions derived from Site I are as follows:

Objective A: Determine type mapping accuracies at two general levels of hierarchy in the ecosystems.

Conclusion: Overall Level II training field classification accuracies were 97 percent for an August data set. Softwood had a 99.2-percent accuracy; hardwood had a 86.1-percent accuracy. The evaluation procedure, using a random sampling technique, indicated the overall correct classification was 74 percent ± 4.8 percent.

Level III training field classification accuracies were 84.1 and 98.3 percent for improved pasture and sagebrush, respectively. The Level III classification of lodgepole pine and spruce fir was 9 percent and 100 percent, respectively, which indicated that Level III species separation was not possible.

Objective B: Establish the season or seasons that offer the greatest potential for type mapping in each ecosystem.

Conclusion: Late summer (August), early fall (October), and temporal data sets showed little difference in overall training field classification accuracies. The August data set did provide more consistent hardwood classification.

Objective C: Provide the USDA Forest Service with project findings and conduct evaluation workshops.

Conclusion: This final report to the Forest Service and NASA will document the project findings. In September 1977 the findings will also be presented to a Forest Service workshop on the TES.

6.2 ASSESSMENT OF TECHNICAL PROCEDURES

The procedures used for processing Site I produced overall good results in terms of ability to correctly classify the scene at Level II and to determine acreage of major vegetation features.

The simulated inventory, using signatures from 10 percent of the site, produced an overall PCC differing little from the PCC calculated for the classification from the separability signature. Any improvement in the overall PCC value will probably require the use of some type of ancillary data in conjunction with the Landsat bands.

The acreage estimates for softwood, using the random sampling, were very good and were improved by use of a regression equation. The acreages for hardwood, grassland, and water were not adequately evaluated using this random design because they occupy only about 2 to 6 percent of the scene. Some type of stratified sampling may have to be developed to adequately evaluate features occupying small proportions of the total scene.

6.3 RECOMMENDATIONS FOR FURTHER STUDY

This study has established precise levels of classification and acreage estimation accuracy using digital Landsat data for a portion of the Rocky Mountain ecosystem. Further development will require defining the precise utility of the study for regional and nationwide forest resources planning.

There are three areas of interest which should be investigated in order to fully apply the results of this study. First, the application of the classification results to area or regional resource data bases. These data bases usually involve several layers of information (i.e., land use, vegetation type, and soil type) registered to some coordinate system and used for planning, modeling, and management of resources. It remains to be defined how Landsat classification may be incorporated into the data bases, whether as a standard land cover base or as a means of updating existing data or both.

Second, the usefulness of these procedures for forest type acreage determination must be analyzed further, in light of the accuracy levels available, to determine their usefulness for regional inventory updates to existing detailed inventory systems conducted every 10 years.

The ultimate question to be answered is "What part can Landsat data play in the inventory and analysis of vegetation given the established accuracies levels, classification detail, and information requirements of the Forest Service?"

7. REFERENCES

1. Dillman, R. D.; and Ward, J. F.: Report on Site Selection for the Ten-Ecosystem Study. LEC-8952, October 1976.

2. Kan, E. P.: Ten-Ecosystem Study (TES) Investigation Plan. LEC-8667, July 1976.

3. Dillman, R. D.: Ten-Ecosystem Study (TES) Site I Report, Grand County, Colorado, Report 1 of 4. LEC-9379, October 1976.

4. Dillman, R. D.: Ten-Ecosystem Study (TES) Site I Report, Grand County, Colorado, Report 2 of 4. LEC-10115, February 1976.

5. Dillman, R. D.; and Almond, R. H.: Ten-Ecosystem Study (TES) Site I Report, Grand County, Colorado, Report 3 of 4. LEC-10524, April 1977.

6. Anderson, S. R.; et al.: A Land Use and Land Cover Classification System for Use with Remote Sensor Data. Geological Survey Professional Paper Number 964, 1976.

7. General Soil Map of Grand County, Colorado. Soil Conservation Service, April 1972.

8. Kan, E. P. (ed.): Technical Analysis Procedures for the Ten-Ecosystem Study. LEC-9379, December 1976.

9. Freese, F.: Elementary Statistical Methods for Foresters. USDA, Agriculture Handbook 317. March 1974.