NOTICE

THIS DOCUMENT HAS BEEN REPRODUCED FROM MICROFICHE. ALTHOUGH IT IS RECOGNIZED THAT CERTAIN PORTIONS ARE ILLEGIBLE, IT IS BEING RELEASED IN THE INTEREST OF MAKING AVAILABLE AS MUCH INFORMATION AS POSSIBLE

"Made available under NASA sponsorship in the interest of early and wide dissemination of Earth Resources Survey Program information and without liability for any use made thereof."

E81-10 120

Final Report

April, 1980 to December, 1980

NASA SCIENCE GRANT 2341

S.M. Norman, NASA Technical Officer

NASA - AMES RESEARCH CENTER

(E81-1012C) THE APPLICATION OF LANDSAT

REMOTE SENSING TECHNOLOGY TO NATURAL

RESOURCES MANAGEMENT. SECTION 1:

INTRODUCTION TO VICAR - IMAGE CLASSIFICATION

MODULE. SECTION 2: (Humboldt State Univ., G3/43 00120

The Application of Landsat Remote Sensing Technology to Natural Resources Management

Lawrence Fox III, Principal Investigator School of Natural Resources Forestry Department Humboldt State University Arcata, California 95521

CONTENTS

	Page
Introduction	1
Section One - VICAR Training Module	
Project Overview	3
Image Classification Lecture	4
Image Classification Exercise	8
Appendix I VICAR Command	
Sequence - Stat. Editing	17
Appendix II VICAR Command	
Sequence - Image Classification	19
Project Summary	20
Section Two - CDF Inventory	22
Introduction	23
Methods	25
Results	30
Conclusions	37
References	38
Concluding Remarks	39

INTRODUCTION

The purpose of this project was to develop optimum techniques for training and consulting with users and potential users of Landsat digital computer technology. Specific goals were to:

- Develop technical schedules and consulting techniques for integrating Landsat classifications with geo-based information systems (GIS) through cooperation with the U.S. Forest Service, McCloud Ranger District, data base integration project.
- 2) Develop appropriate methodologies for expanding the information displayed by an unsupervised Landsat classification of the State of California for a limited geographical region. This investigation was to be done in cooperation with the California Department of Forestry (CDF) in a northern California county.
- 3) Produce a Semi-annual and a Final Report documenting significant results and conclusions.

The first goal, integration of a GIS, could not be accomplished since the Forest Service decided against incorporating Landsat data into their GIS at this time. It was therefore decided to alter the grant through mutual agreement between the Technical Officer and the Principal Investigator. The first goal was changed to: developing Landsat training techniques through the support of a VICAR Software Workshop at the NASA Ames Research Center. The material developed for this workshop and a short summary statement of results are presented in Section One of this report.

The second goal, Landsat information extraction, was accomplished as planned and a description of the findings is included as Section Two. The semi-annual report has been submitted and this document is the Final Report, fulfilling goal number three.

SECTION ONE

VICAR Workshop

Project Overview

The Principal Investigator developed and tested a teaching module on image classification procedures using the VICAR computer software package. The module was developed to optimize the training benefits for State users of the VICAR programs. The module was field tested at a NASA sponsored VICAR users workshop in the Fall of 1980. The module is organized into three basic sections: (1) Lecture on Image Classification using VICAR, (2) Exercise on classification with VICAR and (3) Discussion of results.

ANC Workshop: Introduction to VICAR

IMAGE CLASSIFICATION MODULE

-- L. Fox - Humboldt State University

COLTENTS:

Lecture/Discussion on Image Classification

DAY TWO 1330-1530

Exercise on Training Site Selection

DAY TWO 1545-1745

Exercise on Creating Classification Statistics and Batch Classification

DAY TWO 1900-2200

Discussion, Examine Classification Output

DAY THEEE 900-1000

LECTURE OUTLINE:

(Day Two 1330-1530)

- I. Introduction, Computer Classification and Manual Image Classification
 - A. Goal --- Produce map of resource types, common to both approaches
 - B. Technique
 - Computer assigns pixels to resource types according to spectral signature.
 - 2. Photo interpreter (PI) uses color/tone, texture, shadow, sine, shape, pattern, association.
 - As resolution decreases, PI relies less on size, shape, pattern, texture, shadow. Therefore, advantages of PI reduced.
 - 4. For large data sets, computer classifies very rapidly compared to Pi.

II. Spectral Space

- A. Description and comparison to photo products.
 - 1. One dimension, the "histogram."
 - a) Henno
 - b) Variances

- 2. Two dimensions, the "two channel plot."
 - a) Means
 - b) Variances
 - c) Covariances
- 3. Many dimensions, compare to reflectance curve.

III. Information Classes vs. Spectral Classes

- A. Ideal Conditions
 - 1. Every information class represented by one spectral class.
 - 2. Every information class represented by more than one spectral class.
- B. Unacceptable Conditions
 - 1. Information classes not represented by any spectral class.
 - 2. Spectral classes representing more than one information class.

IV. Computer Classification

- A. General problem Subdivide spectral space such that:
 - 1. All spectral classes are separable.
 - 2. All spectral classes represent one information class.
- B. General Procedure
 - 1. Define each spectral class (for every channel) by:
 - a) Econ reflectance value.
 - b) Variance (Std. dev.2).
 - c) Covariance matrix shows attitude of class.
 - 2. Use the classes defined (called statistics) to classify the Landart score by:
 - a) Haximum libelihood.
 - b) Keny other decision rules.

V. Supervised Classification

- A. Approach
 - 1. Select training areas define information classes.
 - 2. Create statistics files
 - 3. Maximum likelihood clr.scification.
 - a) Relate to historym.
 - b) Spectral confusion problems.

- 4. Evaluate results.
- 5. Iterate if accuracy is poor.
- B. When used: resource information lasses represented by one spectral class.

C. Problems:

- 1. Total spectral variation of information classes is un'mown.
- 2. Several spectral classes may occur within one information class or training area. No way of knowing this.
- General concept subdivides spectral space into information classes not spectral classes.

VI. Unsupervised Classification or Clustering

A. Approach

- 1. Select number of spectral classes desired.
- Use clustering routine to establish spectral classes, a final stat file.
- Determine the information class associated with each spectral class defined.
- B. When used: resource information classes represented by several spectral classes.

C. Problems

- 1. Limits on the number of pixels clustered at one time.
- 2. Pixels containing two or more information classes are sampled which may distort statistics.
- 3. Limited control of what classes are created.
- 4. Concept results in a stat file based on spectral information only.
 No concern for what the unalyst may want out of the data.

VII. Guided Clustering

A. Approach

- 1. Select training areas generally define information classes.
- 2. Create histograms of each training area for every channel.
- Determine the number of spectral classes present in the information class by clustering.
- 4. Repeat steps 1-3 for every general information class.
- 5. Combine all classes.

- 6. Pool similar classes or delete classes as necessary to create stat file.
- 7. Faximum likelihood classification.

B. Advantages

- Control over the clustering process as a limited number of spectral classes are requested within a training area.
- Pixels containing many information classes are relegated to a separate class which can be deleted.
- 3. Clustering finds spectral variation within information classes.
- 4. Possible to concentrate on separating information classes which are similar spectrally.

C. Problems

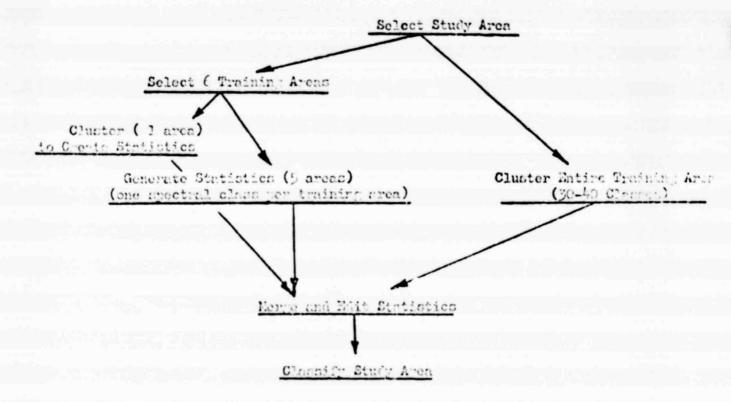
- Spectral classes not occurring in training fields are not considered, causing missclassification.
- Difficult to establish training areas containing enough pixels (several spectral classes per training area).

CIASS EXCERCISES

IMAGE CLASSIFICATION WITH VICAR

- Purpose: 1) To illustrate the supervised, unsupervised, and guided clustering approach to statistics file generation using VICAR.
 - To demonstrate the VICAR hybrid classifier using a parallelpiped/maximum likelihood classification algorithm.

Flow of Works



1. Select Study Area

A. Site

The selection of an area for computer classification is influenced by several factors:

- 1) Management unit or region of responsibility
- 2) Natural boundaries of ecosystems, such as watersheds
- 3) For experimental work, the degree to which lund-use categories of interest are represented on the ground.

The instructors have selected a study area for the exercises today.

It is a coastal region in Northwestern Columbia, South America. The region was selected because it illustrates the international scope of the Landsat system, contains several general categories of land-use, and we happened to have the Landsat CCT sitting on the shelf.

B. Land Cover Information Categories

The selection of information categories is the beginning step in classifying Landsat data. Information category selection can be thought of as goal setting in the planning process. Decisions must be made on what land cover types are needed for the inventory and what landsat spectral signatures are apt to represent those cover types.

Examining air photos of the study area will give an indication of the land cover categories present. We have some space photography from Skylab and some high resolution RBV imagery of the study area for you to use in this exercise. An enhanced Landsat, false color composite of the study area is also available. You should select approximately 3-5 categories of land-use which you believe to be present in the study area.

II. Select 6 Training Areas

Training areas should be selected from grey scale computer print .out of the study area. These areas should be representative samples of the information categories you have selected and contain at least 100 pixels. Avoid border pixels. Compare the grey scale print-out with the photos and false color composite to locate training areas for about 3 of your resource categories that appear very homogeneous. Outline, in pencil, training areas for 3 of your resource categories on the grey scale print-out, in rectangular form. VICAR can accept irregular shapes through the "VERT" format but we will stay with rectangular today. Record the starting line, starting sample, number of lines and number of samples for each training area. We would like you to select 5 training areas. This means you have room for 2 resource categories to be represented by 2 training areas each. The remaining resource categories are to be represented by one training area. In an actual application run, one would want several training areas to be specified for each resource category of interest. This allows for a more representative sample to be selected.

Now, you will probably find that at least one of your resource categories did not fit nicely into a rectangular, homogeneous area. This is a very common problem in Landsat classification work. Some resource categories have very heterogeneous light reflectance patterns (Landsat spectral signatures). Some are interspersed among other categories. Many wildland vegetation categories are very heterogeneous and obviously require special treatment.

We will use "clustering" for this heterogeneous condition. Select one training area which represents a land-use category of interest yet is very dissected and mixed up in appearance on the false color composite. Because one is less sure of actual ground location when specifying these heterogeneous types, a larger training area is preferred; say 200 to 500 pixels. Record the coordinates for this area in the same format that the other five areas were recorded in.

III. Generate Statistics (5 areas)

Landsat spectral statistics are the driving force behind the image classification process. They are use by the maximum likelihood classifier to calculate the probability of a pixel falling into a particular class. It is central to classification accuracy that these statistics be as homogeneous as possible (i.e., low variance in each channel) and yet represent the land cover type as accurately as possible.

We are taking the supervised approach to spectral statistics development for the 5 homogeneous training areas we have selected. We simply tally all pixels occurring within the training area and compute the multi-variate means, variances and covariances for the data set defined by the training area boundaries.

The VICAR program "STATS" is used to generate supervised training statistics. A complete VICAR run is explained in Appendix 1.

IV. Cluster (1 area)

We feel that our heterogeneous land-use category probably contains several spectral classes which will likely color the same color (i.e., place in the same final class) in the final photo product of the finished classification. Yet we are unsure of where these spectral classes are located (i.e., which pixels are which) and how many are there. Statistical clustering will allow the identification of natural grouping or clustering in the data set. We will use the VICAR program "CLUSTER" to automatically select 10 spectral classes from our heterogeneous area. See Figure 1 for details on using "CLUSTER."

V. Cluster Entire Training Area

It is often useful to cluster the entire study area in order to pick up spectral statistics which may not be in the training areas you have defined. We will not take the time to do this additional clustering during the workshop.



VI. Merge and Edit Statistics

An analyst cannot assume that the statistics produced by the supervised approach and by clustering are statistically distinct from each other since the work was done independently. A classification of pasture produced from clustering may have near identical spectral patterns as a pasture class developed from a pasture training area. Therefore, it is not enough to simply combine all of the statistics produced from various approaches; they must be edited.

It is quite possible that two classes may be similar enough spectrally to cause mis-classification. When two spectral classes overlap excessively in spectral space, the computer experiences difficulty in deciding which class to place pixels into. One tool which assesses the spectral similarity or uniqueness of two classes is the separability statistic. Separability statistics are computed for each pair of spectral classes and combined in a separability matrix. In this example, two classes were very similar with a separability of 0.13. These classes, 4 and 5, were developed from the two cypress training areas which seemed to appear slightly different on the red (channel 2) line printer map. This low separability statistic indicates that these two classes could not be accurately distinguished by the classification algorithm. Since neither of these classes conflicted with other classes, they should be pooled together to form one spectral class for cypress swamp. This was done with the following core VICAR commands:

E,STATEDIT, IN,GRP, ,PS1

P.PS1

SELECT, 1, 2, 3, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, POOL, 4, 5 SEPAR A two channel map is also useful for visually comprehending the positions of each class in spectral space. A two channel plot is drawn in Figure 1. Classes 1-5 were developed by a supervised approach and show wide dispersion except for 4 and 5. This is to be expected since we trained on vastly different resource categories to form classes 1-3. Classes 6-15 were developed from clustering the heterogeneous mountainous area near the urban center. Note how the clustering has produced a "cloud" of similar yet spectrally distinct classes. This is typical for classes developed from clustering. Also note that the euclidian distance between many of the unsupervised clusters is as small as the distance between class 4 and 5. The variances are smaller for the clustered data, however, allowing for less spectral overlap. In general, clustering allows the identification of more spectral classes than the supervised approach.

We did not experience any spectral overlap between classes developed from clustering and supervised classes. We were lucky. In a more thorough development of a classification, one would have trained on more upland, mountainous areas and experienced some spectral overlap and confusion. When the editing process becomes more complex, two good general rules to follow are these: 1) When two classes are very similar yet don't overlap with other classes, they should be pooled. 2) When a class is similar to two or more other classes the class causing the problem should be deleted. Of course many other more subjective considerations affect the decision to pool or delete, such as the number of pixels in the classes, the importance of a particular class to the user and the variances of the classes involved.

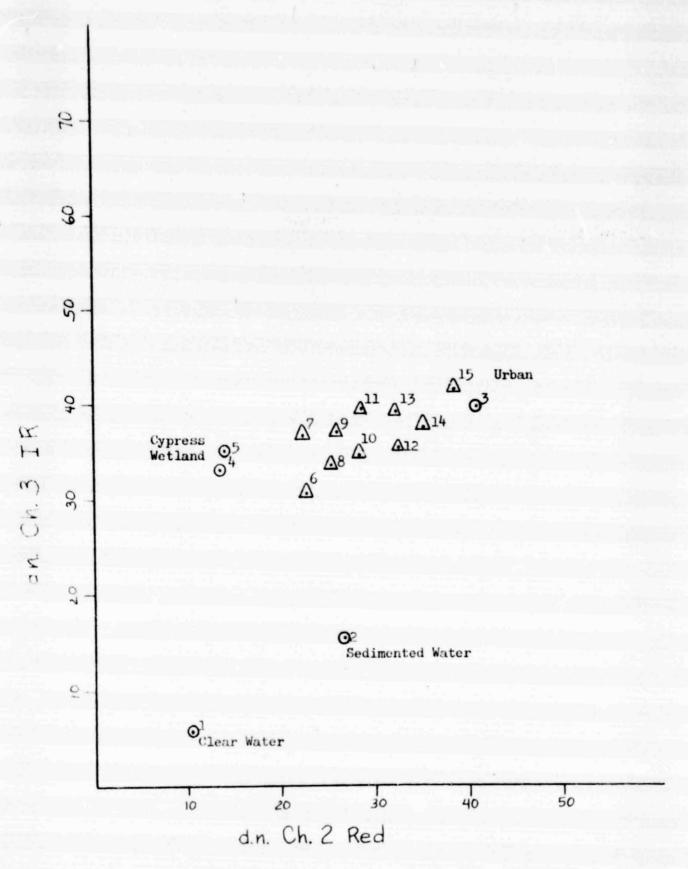


Figure 1. A two channel plot of the spectral classes developed from supervised training (\odot) and guided clustering (Δ) techniques.

VII. Final Classification

The VICAR software has two classification algorithms available, BAYES and FASTCIAS. BAYES is a multispectral classifier which uses the Bayesian maximum likelihood algorithm. FASTCIAS is a multispectral classifier using an algorithm which combines the parallepiped and Bayesian techniques. The classification is done in a two-step process. First, parallepiped decision boundaries are set up at plus-or-minus a number of standard deviations from the mean in each channel. The user specifies the number of standard deviations to use. In order for a pixel to be placed into a specific class, the mean value must fall within the decision boundaries for that class. If a pixel's spectral signature falls outside the decision boundaries for all classes, the pixel is assigned to the unknown class. A pixel whose spectral signature falls within the decision boundary for more than one class is considered ambiguous. The user can resolve the ambiguity by classifying this pixel with a maximum likelihood algorithm. The following basic VICAR commands were used for FASTCIAS:

E,FASTCIAS,(A,GRP),CAT,,PS2 P,PS2

MSS,4 SIGMA,9.0

These statements will execute FASTCIAS using an input data set,
"A," and a statistics file, "GRP." The output file is called "CAT."

A complete listing of the VICAR command sequence to edit the statistics and run BAYES and FASTCLAS is presented as Appendix II.

APPENDIX 1

Annotated VICAR Command Sequence to; (1) Generate Statistics, (2) Cluster to Create Statistics and (3) Edit Statistics.

Each section of the following VICAR run is lettered and the explanation for each section is located after the letter as a footnote.

Line

```
SETUP NEED TAPE WJTOO8 9 TR800 BPI WITHOUT RING
 2
           TIME, 500
 3456789
           READ, , WJT008, N, 9
           SAVE, 2, 101, 101, VICAR1, COLST, COLCLUSM
           RESERVE, 4, 514, 514, *, (COL1, COL2, COL3, COL4)
           RESERVE, 1, 2055, 514, . , A
           E,SAR, NO1, COL1, (211, 328, 512, 512)
           E,SAR, NO2, COL2, (211, 328, 512, 512)
           E.SAR, NO3, COL3, (211, 328, 512, 512)
10
           E.SAR, NO4, COLA, (211, 328, 512, 512)
11
           E,STATS, (COL1, COL2, COL3, COL4), COLST, PS1
12
           P, PS1
                HIST SPEC TRAIN
13
14
                CLASS, 1,240,430,20,20
15
16
17
18
                CLASS, 5, 466, 853, 20, 20
19
           E,MSS,(COL1,COL2,COL3,COL4),A
20
           E,CLUSTER, A,COLCLUSM, , PS2
21
           P.PS2
22
                MSS, 4 SPEC NCATS, 10 NCMIN, 5
23
                HIST, 2,4 DELTA, 0.45 CONV, 98.0
24
                SAMPLE, 4 TRAIN
25
                CLASS, 1, 328, 500, 100, 100
26
           E ,STATEDIT ,COLST , * , ,PS3
27
           E,STATEDIT, COLCLUSM, *, , PS3
28
           P.PS3
29
                COPY ALL CHECK SEPAR
           END
30
```

EXPLANATION OF LETTERS

- A. These statements are standard set up cards for a VICAR run.
 - Line 1, tells the operator what tape to use
 - Line 2, sets an upper time limit of 500 seconds
 - Line 3, reads the tape and labels the tape "N"
 - Line 4, reserves 2 permanent data sets called "COLST" and "COLCLUSM" on a disk pack called VICAR1.
 - 101 samples per line by 101 lines
 - Line 5, reserves 4 temporary data sets in the public memory (*) called COL1, COL2, COL3, COL4. Each data set is 514 samples x 514 lines.



- Line 6, reserves one temporary data set, 2055 samples by 514 lines in the public memory. The file is to be called "A."
- Lines 7-10, request that four bands of Landsat data be copied from tape "N" to disk files called COL1-4. (COL for Columbia) Only a subwindow of data is to be copied however, starting at line 211, sample 328 and continuing 512 samples by 512 lines.
- B. These statements generate spectral statistics for 5 training areas.

 Lines 11-18, generate spectral statistics using 4 bands of Landsat data

 (COL1-4) and call the output file COLST (for Columbia statistics).

 The size field is defaulted (,,) and the statistics are to be generated according to parameter statement 1. Parameter statement 1 (lines 12-18) calls for histograms of all input bands of each spectral class and spectral plots of all bands. The word "TRAIN" tells the computer that training fields are to follow. Lines 14-18 give rectangular window coordinates (starting line, starting sample, number of lines, number of samples) for 5 training areas.
- C. This one statement (line 19) reformats the 4 bands of data into an MSS format window called "A."
- D. These statements request a clustering algorithm to be passed through data set "A."
 - Line 20, requests the "CIUSTER" routine to be used on data set A. The output file is to be called COICIUSM (for Columbia clustering on a mountainous area) and the size field is defaulted (,,). The parameter statements (lines 22-25) indicate that; (1) the input data is 4 channel, MSS format, (2) spectral plots of all channels are requested, (3) number of classes to be generated is 10, (4) minimum number of classes after merging is 5, (5) histograms are requested of channels 2 and 4, (6) the delta or minimum separability statistic is 0.45, (7) the percent convergence required is 98, (8) every 4th pixel is to be sampled and (9) one training area is to be submitted, 100 lines by 100 samples.
- E. This section requests that separability matrices be printed out for both statistics files created, COLST and COLCIUSM.

APPENDIX 2

VICAR command sequence to edit the statistics and run BAYES and FASTCIAS.

```
CHARGE , T4352
TIME.1900
SETUP, NEED TAPE WJT008 9TR 800 BPI WITHOUT RING
SETUP, NEED TAPE WJT018 9TR 800 BPI WITH RING
READ, . WJT008, N.9
WRITE, , WJT018, T,9
RESERVE, 4,514,514, *, (COL1, COL2, COL3, COL4)
RESERVE, 1, 2055, 514, ., IN
FIND, (GRP, VICAR1)
E.SAR, *NO1, COL1, (211, 328, 512, 512)
E,SAR, NO2, COL2, (211, 328, 512, 512)
E,SAR, NO3, COL3, (211, 328, 512, 512)
E,SAR, NO4, COLA, (211, 328, 512, 512)
E,MSS,(COL1,COL2,COL3,COL4),IN
E,FASTCIAS,(IN,GRP),COL1,,PS2
P.PS2
   MSS,4 SIGMA,9.0
E,STRETCH,COL1,COL2,,(LINEAR,0,63)
E,DISPLAY,COL2, ., PS3
P.PS3
   .-=+/N°ON=)(4S
             +M/ = /
E, BAYES, (IN, GRP), COL3, (MSS, 4)
E,STRETCH,COL3,COL4,,(LINEAR,0,63)
E.DISPLAY, COLA, ... PS4
P.PS4
     .-=+/WOW=)(4S
            +M/ = /
E,SAR,COL1,T
E,SAR,COL3,T
```

END

Project Summary

In general, the module was well received by the VICAR users and potential VICAR users at the NASA Workshop. There was enough time allotted for the lecture and exercises. The material seemed to be addressing the appropriate level of sophistication for efficient technology transfer.

The schedule called for the students to actually run VICAR soft-ware, "live" at the workshop. This was quite beneficial for establishing user confidence, yet required a tremendous amount of backup effort. Mike MacDonald of Technicolor gave his energy above and beyond the call of duty. If Mike had not been willing to put in three 18 hour days, the "live" portion would have failed miserably.

The lesson learned was not to underestimate the amount of support required for students in a workshop to actually process Landsat data. I would suggest that 4 or 5 consultants be available to review student card decks before submission to the computer and to help de-bug VICAR runs after they have failed to operate.

The discussion session would have been improved substantially if more examples of VICAR classifications had been available. The Principal Investigator came up short in this area. In the future, it would be best to have several comparative VICAR classifications (e.g., FASTCLAS and BAYES with supervised statistics, with guided clustering and with unsupervised statistics) available for student review and discussion. I was counting on the student output which didn't materialize to the level required for an in-depth discussion.

SECTION TWO

FOREST RESOURCE ASSESSMENT

OF

HUMBOLDT COUNTY

By:

Lawrence Fox III and Kenneth E. Mayer
SCHOOL OF NATURAL RESOURCES, DEPARTMENT OF FORESTRY
HUMBOLDT STATE UNIVERSITY

A Report of Work Performed Under NASA Grant No. NSG-2341

NASA TO MICAL MONITOR:

NASA TECHNICAL CONSULTANTS: James Brass and David Peterson

In Cooperation With

FOREST RESOURCE ASSESSMENT PROGRAM (FRAP)

and

CALIFORNIA STATE DEPARTMENT OF FORESTRY (CDF)

FRAP COORDINATOR: Nancy Tosta-Miller

INTRODUCTION

On January 8, 1980, Judge Lawrence K. Karlton of the U.S. District Court of Eastern California ruled that the National Environmental Impact Statement (NEIS) supporting the second Roadless Area Review (RARE II) did not satisfy the requirements of the National Environmental Policy Act (NEPA). The court then enjoined the U.S. Forest Service (USFS) from developing any of the 47 disputed non-wilderness areas in California pending proper consideration of wilderness values in compliance with NEPA. Primary deficiencies cited by the judge included the "lack of site-specific data on RARE II areas" (J. of Forestry, 1980).

This dramatic court decision with its far-reaching national implications highlights a pressing need for inventory data on a site specific basis. Landsat, a NASA-sponsored earth imaging satellite, is providing the potential for site specific forest inventories of very large areas. One Landsat scene covers approximately 3.4 million ha (8.46 million ac) or 34,255 km² (13,214 square miles) with a grid cell information base, each cell having a ground area of approximately 0.46 ha (1.12 ac). The satellite measures reflected "light" in four bands of the electromagnetic spectrum, two visible and two non-visible infrared wavelength bands. Reflectance patterns called "spectral signatures" are measured by the satellite as it passes in sun synchronous, near polar orbit. The spacecraft returns to the same location on the earth every 18 days (Sabins, 1978).

The legislators of California have responded to forest inventory needs by passing the California Forest Resources Assessment and Policy Act of 1977 (FRAPA). This bill (AB 452) established a project to improve the information base, upon which forest policy decisions are formulated. Under policy guidance from the California Board of Forestry and with the help of the Secretary of Resources, the Director of Forestry was required to prepare a forest resources assessment and analysis by July 1, 1979. The bill also provides for an updated assessment every five years.

The staff of the Forest Resource Assessment Program (FRAP), a branch of the California Department of Forestry (CDF), cooperated with the NASA Ames Research Center (NASA-ARC) Moffett Field, California, to use Landsat data in the preliminary assessments. Forest cover was classified by computer, using Landsat data of the entire state. Land condition and type of vegetative cover was placed into one of 17 categories which included conifer forest, conifer/hardwood forest, hardwood forest and brush. Area summaries were provided by category for each county of the state (Peterson and Tosta, 1979).

To further develop the detail of the assessment program, a cooperative project was undertaken in 1979-80 between CDF, NASA, Humboldt State University (HSU), California Polytechnic State University at San Luis Obispo, and the Tahoe National Forest. The objective was to intensively inventory three areas in California with Landsat digital data to determine

ORIGI AL PAGE IS OF POOR QUALITY if more classification detail could be achieved on a local basis than had been possible in the state-wide survey. The results would assist CDF in fulfilling future forest assessment responsibilities.

This report is a description of the Landsat inventory of Humboldt County, completed by HSU. It therefore represents one phase of the cooperative project. The goals of the Humboldt project were to:

- Develop an intensive forest land inventory strategy for Humboldt County that could evaluate the full discriminatory power of Landsat spectral signatures for vegetation classification.
- Map and inventory the forest land of Humboldt County using Landsat digital data via computer classification techniques.

PROJECT AREA

Humboldt is California's 14th largest county comprising 9254 km² (3573 square miles) (Figure 1). The county is approximately 170 km (105 miles) long, north and south, and 56 km (35 miles) wide, east and west. Humboldt is primarily mountainous with the exception of the alluvial lowlands in the Eel River Valley and Humboldt Bay Area. Elevation ranges from sea level to over 2133 m (7000 feet). Precipitation occurs during the winter months, with an average rainfall of 101 cm (40 inches) with excesses of 229 cm (90 inches) occurring in isolated coastal mountain areas. Temperatures are highly variable in the county with lows for the winter of -8.8°C. (16°F.) in the mountainous areas and 12.2°C. (54°F.) in the Eureka area. Summer high temperatures range from 18°C. (60°F.) on the coast to 40°C. (100°F.) in the inland areas.

The 1980 census showed 107,500 county residents with approximately 50 percent living within a 25-mile radius of Eureka. Humboldt has a higher population than any of its neighboring counties. Del Norte to the north, Mendocino to the south, and Trinity and Siskiyou to the east.

Twenty-seven percent of the county is in public (federal, state, county, and municipal) ownership and 73 percent is private (Humboldt Atlas, 1975). The Forest Products Industry provides approximately one job in four. Humboldt is the top lumber-producing county in California and ranks among the top lumber-producing counties in the United States. Redwood (Sequoia sempervirens) is the major commercial species with Douglas-fir (Pseudotsuga menziesii) and hardwoods providing the remainder of the important timber resources. Old growth timber is limited, creating a major shift towards second growth harvesting and management.

Important conifer species are: redwood, Douglas-fir, grand fir (Abies grandis), sitka spruce (Picea sitchensis), Western hemlock (Tsuga heterophylla), Port Orford-cedar (Chamaecyparis lawsoniana) and Western red cedar (Thuja plicata). Important hardwood species are: tanoak (Lithocarpus densiflorus), Pacific madrone (Arbutus menziesii), giant chinquapin (Chrysolepis chrysophylla), red alder (Alnus rubra), Oregon white oak (Quercus garryana), and California black oak (Quercus kelloggii).

INVENTORY DESIGN

Landsat provided a comprehensive data base with a grid cell coordinate system. The inventory was designed to classify every picture element (grid cell) and therefore was a 100 percent inventory with no sampling. While it was certain that all of the area would be inventoried to a .47 ha cell size, the exact classification was not prepared beforehand. The spectral signatures developed from the Landsat pixels may have represented various levels of classification detail depending on environmental conditions, forest types involved, the discriminatory power of the computer programs, and the skill and experience of the analyst.

We began the project with the goal of identifying two major conifer species types, redwood and Douglas-fir, and various mixes, as well as several forest density categories and size classes. A literature review of work completed prior to 1975 indicated that we would not reach this level of detail (deStinger, 1978). Our experience had been that the initial classification goals were achievable (Fox and Mayer, 1980) and recent work by others indicated a high probability of success (Walsh, 1980).

ANALYSIS TECHNIQUES

The classification of Landsat data can be divided into three dictinct tasks:

- Developing a set of descriptive statistics for the spectral classes required.
- 2) Classifying the study area into these classes.
- 3) Identifying the resource label for those spectral classes.

To complete this project, we used unsupervised and guided clustering techniques (Flemming, 1975) to develop the needed spectral classes and multivariate discriminant analysis for classifying every pixel in the study area (Sabins, 1978).

Scene Selection and Image Mosaic

After a review of available Landsat images of Humboldt County that were cloud-free and date sequential, two April 12, 1977 scenes were selected. Each image covered approximately 50 percent of the county. A computer compatible tape was acquired from Earth Resources Observation Systems (EROS) Data Center. Radiometric anomalies including bad data lines and points, radiometric striping and atmospheric scattering were not corrected or normalized before processing. These errors were not significant for the data tape.

The two images were overlapping north to south by approximately 200 lines. To assure an accurate line-by-line match for joining the two images, the overlapping lines were removed, which was completed on the IBM-360 computer at NASA-ARC. This assured continuous, non-overlapping data north to south. Each scene was copied to tape as well as to computer disk for easy access during processing.

Resource Category Selection

In order to determine the maximum detail that could be obtained from Landsat digital data classifications, resouce categories were selected for targeted forest land variables. The approach focused on conifer species differentiation, tree size class, and crown closure category. Furthermore, forest categories such as hardwood, brush, and regeneration areas were also being selected for classification.

Before a feasible classification scheme was selected, an examination of the vegetation communities was completed. As the different forest communities were identified, a determination was made relative to the following criteria:

- Is the resource type an important part of the vegetation mosaic of Humboldt County?
- Is there enough acreage of that vegetative community to facilitate computer processing?
- 3) Would the resource type be spectrally separable from the other types?

If it was determined that a vegetative community met these criteria, it was included into the preliminary classification scheme. The CALVEG classification scheme was used as reference to develop the vegetation types (Parker and Matyas, 1980). It was also decided that resource categories and classification goals would change throughout the classification process as the limitations of spectral discrimination and computer software were encountered.

Training Area Selection and Groundtruthing Effort

Approximately 10 homogenous training fields were selected for each resource category. The training areas were at least 10 ha (25 ac) in size, with the majority of training fields being > 20 ha (50 ac). The training areas were located on U-2 1:32,500 color infrared aerial photography and outlined on U.S. Geological Survey (USGS) 7½' quadrungle maps to facilitate groundtruthing.

A field visit was made to each training area to determine plant species, vegetative crown closure, age, size class, aspect and slope. This information was recorded and used as an aid in the labeling of spectral classes generated.

Landsat Classification Procedures

In order to extract the "raw" (unaltered) Landsat spectral information from within each training field, the Landsat images were displayed on the Interactive Digital Image Manipulation System (IDIMS) cathode ray tube. The training areas were located by visual inspection of the Landsat data. These areas were removed from the data set and copied on disk to facilitate guided clustering.

Guided clustering was completed on each set of training areas. EDITOR software was used to analyze the Landsat data. Procedures used were similar to those reported by Fox and Mayer, 1979; Forbes et al. 1980; and Mayer et al. 1980.

Unsupervised clustering was performed on each Landsat image to develop spectral statistics unique to the resources present in the scenes. The two statistics files were compared for spectral separability (Fox and Mayer, 1979). The statistics files were edited to eliminate classes with spectral similarity. The final statistics were merged with the statistics developed from guided clustering.

The final, edited statistics file was used as input to a maximum likelihood decision function. This step completed a multivariate discriminant analysis of the Landsat data and classified every pixel into one of the spectral classes developed.

Many more spectral classes (unique light reflectance pattern) are usually generated than the number of resource categories defined (Mayer and Fox, 1979). In order to comprehend the data, several spectral classes must be grouped together into resource categories. Spectral classes were grouped according to geographical proximity and similar resource classification as observed on color infrared photography. Spectral curves of Landsat generated spectral classes were also used to determine the proper resource identity (Mayer and Fox, 1980).

Accuracy Evaluation

The Landsat classified scenes (April 12, 1977) were evaluated using U-2 color infrared photography (May 8, 1978). Classified data windows were selected from both scenes that could be easily located on the ground and on the photographs. Single pixels were randomly selected from these windows. A two man team was used to complete the evaluation. One member of the team located the pixel to be evaluated on the photographs. The second member completed the photointerpretation to determine the "truth" (actual resource identity).

The overall mean probability of correct classification was calculated by:

 $P = -\frac{\gamma}{X}$

WHERE: Y is equal to the number of correct Landsat pixels and X is the total number of pixels identified by photointerpretation.

For a complete technical description of the evaluation process, refer to Mayer and Fox, 1980 and Fox et al. 1981.

Area Summary

Area summaries for the Landsat resource categories were produced for Humboldt County. The boundary of the county was digitized into the computer data base in order to create a specific geographical segment of the categorized data. Area summaries within that segment were accessed through the EDITOR software system.

RESULTS

Resource Category Description

The Landsat spectral data was classified into 64 classes, each one representing a statistically unique light reflectance pattern (Table 1). The spectral classes were assigned to 15 resource categories based on similarities in composition (Table 2). The forested land was classified into nine categories, based on species composition (redwood through hardwood). All of the forest categories represented closed canopy forest stands with various proportions of hardwoods and conifers. The total vegetative crown closure was consistently greater than 65% for all nine forest categories.

The "Redwood" category is dominated by redwood (Sequoia sempervirens) which represents 80% of the conifer trees present. Up to 20% of the conifers present may be one or more of the following species: Douglas-fir (Pseudotsuga menziesii), grand fir (Abies grandis), sitka spruce (Picea sitchensis), and western hemlock (Tsuga heterophylla).

The "Redwood" class may contain small amounts of hardwood, usually red alder (Alnus rubra). The stands are definetly conifer dominated; however, with hardwood seldom occupying more than ten percent of the forest canopy.

The "Douglas-fir >80%" category also represents a closed canopy conifer stand. Douglas-fir dominates and comprises over 80% of the total conifer cover. Associated conifers may occupy up to 20% of the stand and be any or all of the following species: ponderosa pine (Pinus ponderosa) sugar pine (Pinus lambertianna), incense-cedar (Libocedrus decurrens), white fir (Abies concolor), and noble fir (Abies procera). Small amounts of hardwood (less than 10% of total trees) may also be present in the stand.

The "Douglas-fir/Hardwood" categories (61-80% and 26-60%) are fully stocked stands containing a mixture of Douglas-fir and hardwood. Douglas-fir dominates, representing 61 to 80% of the trees present in the 61-80% category. Hardwoods are principally tanoak (Lithocarpus densiflorus) and Pacific madrone (Arbutus menziesii). Hardwoods dominate the 26-60% category. Since Douglas-fir accounts for 26-60% of the trees present, hardwoods and/or brush comprise the remaining proportion of the stand. Up to 20% of the conifers in these stands may be comprised of species other than Douglas-fir, the same associates listed for the previous Douglas-fir class.

lable 1. Means, Variances, and the Number of Pixels for Each Spectral Class, Humboldt County, 1977.

			Contractor and the second	Band 7		Number of Pixels Per
Means Va	riance	Means	Variance	Means	Variance	Spectral Clas
8.70	0.70	27.18	6.09	15.36	2.40	74
10.50	5.29	33.37	27.83	18.53	6.12	30
21.47	9.55	36.07	10.78	15.67	6.24	15
19.29	1.60	89.76	6.57	37.41	3.01	17
9.41	1.70	41.25	1.41	23.50	1.27	70,914
9.06	0.72	16.99	0.41	8.02	2.86	171
9.27	0.51	21.39	0,29	10.34	0.49	85
10.35	2.62	23.78	1.04	12.41	1.16	83
14.15	1.91	29.91	1.98	15.30	1.16	117
7.37	2.22	12.69	2.95	5.83	1.66	41,938
7.64	1.00	15.53	2.95 0.79	7.87	1.04	53
7.86	1.14	18.81	0.89	9.57	0.45	12
8.18	0.81	21.06	0.47	11.12	0.67	98
8.47	0.83	23.40	0.27	12.06	0.81	70
13.25	1.93	26.62	4.27	13.87	2.98	. 8
10.97	1.40	32.90	1.30	16.94	1.25	184,395
19.10	7.43	37.63	1.23 8.71	21.48	0.98	108,340
12.44	0.97	27.61	1.78	19.90	1.88	10
15.26	1.76	36.87	1.71	19.26	1.61	161,401
20.16	2.35	37.39	2.17	18.49	1.71	127,988
16.50	1.87	28.87	2.25	15.12	1.32	16
24.33	5.53	33.44	3.79	17.28	2.45	18
26.21	2.69	37.79	4.07	19.40	1.58	43
14.51	1.04	40.38	1.02	20.85	1.13	47
16.24	2.57	41.49	1.56	21.89	1.33	62,899
14.36	0.82	45.03	0.47	23.60	1.09	60
14.77	3.15 0.88	45.70	2.84	23.98	1.66	83,742
16.37	2.62	48.82	1.22 3.57	25.74	0.86	61
19.49	5.83	56.22	3.32	29.46	1.77 2.36	38
9.27	0.62	33.89	0.89	19.48	0.67	36,830
22.16	3.85	43.85	3.37	21.82	1.80	77,840
15.87	2.52	24.66	2.68	11.57	1.87	78,889
20.64	3.27	28.62	3.28	12.85	2.49	70,533
19.60	2.43	33.06	1.55	16.07	1.61	130,924
24.95	4.44	35.50	5.17	15.96	3.20	68,199
12.10	2.13	44.97	2.05	25.54	1.99	85,344
13.04	4.85	50.58	3.67 5.11	25.93	1.95	56,091
19.21	8.16	62.94	5.56	28.67	3.52	38,078
16.96	4.18	67.90	3.06	33.69	2.71	25,722
12.65	0.64	74.90	8.04	45.22	2.23	40
15.07	1.70	64.64	2.16	36.93	3.03	28
14.54	0.91	69.80	1.76	40.31	1.38	95
14.10	1.20	73.24	1.61	42.07	2.12	71
13.90	0.91	77.50	1.13	43.90	1.37	40
36.52 7.10	22.33	42.34	41.02	17.52	9.33	29
42.98	0.64	39.47	0.78	0.01	0.01	341.319
30.33	11.04	28.51	27.98 14.84	9.47	6.17	12,603
12.75	2.61	19.40	4.42	9.09	4.50 2.59	49,989
37.55	14.69	51.25	22.37	21.44	7.39	24,433
77.76	12.41	72.70	15.56	26.35	3.36	14,346
10.03	0.62	5.75	0.80	0.48		122,841
108.69	13.31	102.59	14.01	37.47		17,387
16.53	8.54	79.22	15.93	45.21	6.22	8,568
59.35			39.00			3,005
39.12 121.13						1,002
127.00						3,494
15.85						4,542
19.05						4,110
59 39 121 127 15	. 12 . 13 . 00 . 85	35 28.74 12 67.97 13 26.30 00 0.00 85 2.23	35 28.74 71.29 12 67.97 84.98 13 26.30 121.47 .00 0.00 126.93 .85 2.23 10.30	35 28.74 71.29 39.00 112 67.97 84.98 95.89 113 26.30 121.47 24.93 .00 0.00 126.93 0.36 .85 2.23 10.30 2.58	35 28.74 71.29 39.00 27.97 112 67.97 84.98 95.89 42.85 113 26.30 121.47 24.93 39.66 0.00 0.00 126.93 0.36 46.82 85 2.23 10.30 2.58 1.38	35 28.74 71.29 39.00 27.97 17.66 112 67.97 84.98 95.89 42.85 32.66 13 26.30 121.47 24.93 39.66 8.80 000 0.00 126.93 0.36 46.82 23.39 .85 2.23 10.30 2.58 1.38 0.93

Table 2. Humboldt County Landsat Resource Categories, Stand Description, And Spectral Classes.

RESOURCE CATEGORY	SIZE CLASS	SPECIES COMPOSITION STOCKING	SPECTRAL CLASS
Redwood (RW)	>15, CD9	>80°. Rw ^b	6,7,8,9
Douglas-fir (DF)	>15,€0	>80%DF	1,2,11,12
Douglas-fir/Hardwood	≥15,00	61-80%DF	13,14
Douglas-fir/Hardwood	>15,00	26-60°DF	16
Mixed Conifer (MC)	>15,CD	> 80°.₩C	18,5,33
Mixed Conifer/Hardwood	>15,00	61-80\$MC	10,15,39
Mixed Conifer/Hardwood	O⊃, 51<	26-60%MC	17,20
Hardwood	>15'HTC	> 90%00م	26,28,30,41,64
Hardwood	>15'HT	26-60%CC	23,37
Soil		< 25%VEG	3,24,36,38,49,54,4,31
Agriculture Grasslands			32,40,42,43,44,45,46 47,48,58,60
Brush	<15'HT	>25%VEG	19,22,27,35,53
Brush/Regeneration	<15'HT	>25%VEG	21,25,29,34
Other (urban, sand, gravel and snow)			55,57,59,61,62,63
Water			50,51,52,56

a = Average crown diameter of dominants and co-dominants.

Total vegetative crown closure is b = Percent of total conifer cover. >65% for all conifer categories.

c = Average stand height of dominants and co-dominants.

d = Total Hardwood crown closure. Other vegetation always present to make total vegetative crown closures consistently above 65%.

The "Mixed Conifer" category represents a closed crown stand in which no one conifer species dominates. The class is made up of redwood and Douglas-fir near the coast and pine, fir, Douglas-fir, and occasionally white fir further inland.

The "Mixed Conifer/Hardwood" classes are identical to the Douglas-fir/ Hardwood classes discussed previously except that the conifers in these categories are mixed with no single species previously predominating. The coastal forest community described as "redwood" was placed in this category (Parker and Matyas, 1980). These forests often contain 30 to 50% hardwood and the conifers rarely are redwood only. Often 30% of the conifers present are other than redwood. The species listed in the "redwood" category are the major ones present.

The two "Hardwood" classes are comprised of red alder, tanoak and madrone in various densities. These classes contain woody vegetation greater than 15 feet in height. This is an important distinction since tanoak grows in brush form as well as tree form. Often the only difference between brush and hardwood is a difference in canopy height. The Hardwood, >60% crown closure class contains greater than 60% hardwood forest canopy closure. The understory is brush or grass/forest. The Hardwood, 26-60% crown closure class contains between 26 and 60% crown closure of hardwoods. Brush or grass/forest make up the remainder of the total vegetative cover.

The "Soil" category is primarily bare soil yet may contain up to 25% vegetation, with scattered trees, brush or grass/forest.

"Agriculture/Grasslands" is a combination range/agriculture category which contains both irrigated and non-irrigated pasture.

"Brush" is a category describing woody vegetation less than 15 feet in height. Tanoak, blue blossom (Ceanothus thyrsiflorus), snow brush (Ceanothus velutinus), deer brush (Ceanothus integerrimus), and coyote brush (Baccharis pilutaris) are major species (refer to Parker and Matyas, 1980 for a complete listing). The brush category may contain conifer seedlings that are beneath the brush canopy level. This class does not contain significant amounts of conifer seedlings or saplings protruding above the brush canopy.

The "Brush/Regeneration" category is identical to the brush canopy except that it contains significant amounts of conifer seedlings and/or saplings protruding above the brush canopy. The conifers represent over 40% of the canopy in this class. Conifer regeneration may be any of the conifer species described previously.

The "Other" category describes lands of little importance to forest management, urban, gravel and snow.

The "Water" category describes rivers and marine estuaries in the county.

Accuracy Evaluation

The overall mean probability for correct classification was 82.2 percent. The accuracy estimate was determined from evaluating 320 single pixels scattered randomly throughout the county. Individual category accuracy statements were not generated as photo coverage was not complete, creating problems in finding and locating adequate sampling areas. Further evaluation of detailed resource categories will take place in the future (Prime Timberland Contract 3-080-078), as photographs become available.

Area Summary

Landsat identified 962,291 ha (2,309,512 ac) in Humboldt County (Table 3). It was found that Douglas-fir was the most predominant conifer category, comprising 21 percent (485,481 ac) of the total land base. This was expected, as a preliminary survey of the county suggested that there was more Douglas-fir present than redwood (Plank, 1974).

Redwood>80% comprises four percent of the total land base; approximately 37,928 ha. It is important to note that there is more redwood in the county than is revealed by this figure. Many of the "second growth" redwood stands were placed into the mixed conifer category; primarily due to the heavy intermingling (>20%) of associated conifer species. In the past, many timber inventories referred to any stand containing redwood as a redwood stand. Since Landsat spectral signatures respond to total vegetative cover, it was not practical or feasible to misinterpret these stands as redwood. Consequently, mixed conifer, >80% cc, was the second most abundant category, comprising 13 percent (125,300 ha) of the land base.

Conifer forest as a whole makes up more than 60 percent (580,669 ha) of the land base in the county. This is a graphic example of the wealth of timber resources in Humboldt. Furthermore, pure hardwood comprises just over 6 percent (62,934 ha). This total may seem to be low, but actually it is very realistic. Much of the hardwood in the county does not exist in pure stands, devoid of conifers. Hardwoods are found in much of the Douglas-fir/Hardwood 61-80 percent, 26-61 percent, Mixed Conifer/Hardwood 61-80 percent, and 26-60 percent. These categories comprise approximately 22 percent (215,158 ha) of the total land base, a significant proportion of the forest resources in the county.

Landsat identified 89,947 ha of brush and 96,198 ha of brush-regeneration, approximately nine and 10 percent respectively. Over half of the brush lands identified were indicated as being overtopped with conifers. Landsat data analysis at this level provided no indication of the age structure of the regeneration zones.

Table 3. Area Estimates For Humboldt County, California Derived From Classified Landsat Data - April 12, 1977.

RESOURCE LABEL	ACRES	HECTARES	PERCENT
Redwood >80%CCa	91,029	37,928	4.0
Douglas-fir>80%CC	485,481	202,283	21.0
Douglas-fir/Hardwood 61-80%CC	55,116	22,965	2.4
Douglas-fir/Hardwood 26-61%CC	50,629	21,095	2.2
Mixed Conifer >80%CC	300,722	125,300	13.0
Mixed Conifer/Hardwood 61-80%CC	214,109	89,212	9.3
Mixed Conifer/Hardwood 26-61%CC	196,528	81,886	8.5
tardwood >60%CC	74,648	31,103	3.2
ardwood 26-60%CC	76,396	31,831	3.3
Brush	215,873	89,947	9.4
Brush-Regeneration	230,876	96,198	10.0
Agriculture-Grasslands	152,608	63,587	6.6
Soil < 20%CC-VEG	93,451	38,938	4.0
Other	55,146	22,977	2.4
later	16,900	7,041	0.7
TOTAL	2,309,512 -	962,291	100.0

Approximately 63,587 ha (152,608 ac) of the Agriculture/Grasslands category were found in the county. This total is somewhat lower than the information reported from the Landsat statewide inventory (approximately 400,000 ac). Much of the rangeland included as grassland previously is possibly being classed as brush in this effort.

REFERENCES

- Fox, L. III, K. E. Mayer and A. R. Forbes. 1981. Classifying forest resources with Landsat data. Unpublished manuscript.
- Mayer, K. E. and L. Fox III. 1980. Mapping wildlife habitat on the McCloud Ranger District with Landsat digital data. Unpublished final report.
- Parker, I. and W. Matyas. 1980. A classification of California vegetation. Second edition, U.S. Forest Service in-house report.
- Peterson, D. and N. Tosta-Miller. 1980. 14th International Symposium on Remote Sensing of Environment.
- Plank, R. D. (Ed). 1974. Humboldt County, California Atlas. Humboldt County. 103 pp.
- Sabins, Floyd F. 1978. Remote Sensing Principles and Interpretation.
 W. H. Freeman and Co. San Francisco. 426 pp.

CONCLUDING REMAKRS

The project goals have been met, as evidenced by the documentation included here and the fact that CDF has continued the work begun here through an Inter-Agency Agreement with Humboldt State University. The CDF is moving forward to incorporate Landsat data into their operational forest inventory procedure.

It is unfortunate that the U.S. Forest Service chose not to incorporate Landsat at this time. We continue to work with the McCloud Ranger District on an informal basis and look forward to future Landsat projects.

Please receive my hearty thank you for the opportunity to work in this exciting aspect of modern day remote sensing.