SPACE
SCIENCES
LABORATORY

CALIFORNIA LANDSAT-AIDED CROP ESTIMATION

Principal Investigator:

R. N. Colwell

andy
Contributors: R. W. Thomas (Project Scientist)
L. H. Beck
C. E. Brown
S. L. Wall

Annual Report
1 January 1982 - 15 January 1983
USDA Cooperative Agreement No. 58-319T-2-0341X
Remote Sensing Research Program
Space Sciences Laboratory

UNIVERSITY OF CALIFORNIA, BERKELEY

CALIFORNIA LANDSAT-AIDED CROP ESTIMATION

Principal Investigator:

R. N. Colwell

Contributors: R. W. Thomas (Project Scientist)
L. H. Beck
C. E. Brown
S. L. Wall

Annual Report
1 January 1982 - 15 January 1983
USDA Cooperative Agreement No. 58-319T-2-0341X
Remote Sensing Research Program
Space Sciences Laboratory
University of California
Berkeley, California 94720

Acknowledgements

The authors wish to gratefully acknowledge the following individuals for their contributions to this study: Mr. Charles Caudill of U.S. Department of Agriculture's (USDA) Statistical Reporting Service (SRS), Statistical Research Division; Dr. Rich Allen, Mr. William Wigton, Dr. Richard Sigman, and Mr. Martin Holko of the USDA-SRS Remote Sensing Branch; Mr. Robert McGregor, Mr. Ron Bosecker, and Mr. David Kleweno of the California Crop and Livestock Reporting Service (sponsored jointly by the USDA-SRS and the California Department of Food and Agriculture); Mr. Glenn Sawyer and Mr. Charles Ferchaud of the California Department of Water Resources' (DWR) Statewide Planning Branch, Land and Water Use Section; Mr. Jay Baggett of the California DWR and the University of California at Santa Barbara; Dr. Dale Lumb, Mrs. Ethel Bauer, and Mrs. Chris Hlavka of the NASA-Ames Research Center, Technology Applications Branch; and Mr. Ed Sheffner and Mr. Byron Wood of Technicolor Government Services at the NASA-Ames Research Center.

We would like to especially recognize Mr. Marty Holko for his frequent assistance in USDA-SRS coordination, project task setup and Landsat data acquisition; Mr. David Kleweno for his assistance in coordination with the California office and important ground data editing effort; Mr. Charles Ferchaud for his invaluable assistance in ground data collection, interpretation of agricultural conditions, especially relative to project classification goals; Mr. Jay Baggett for a similar ground data effort, an even greater effort in segment digitizing, and for his contributions relative to project scope and task definition; Mrs. Ethel Bauer for her irreplaceable assistance in obtaining NASA support for a variety of Cooperative Study tasks (including U-2 photography of the entire test site); Mr. Ed Sheffner for his effectiveness in the construction of data sets through NASA-Ames and his valuable assistance in the small grains work; and Mrs. Chris Hlavka for her general review, and for contributions to the small grains effort.

A number of other individuals also deserve recognition for their time spent collecting data in support of the 1982 Landsat pass field transect program. These include: Mr. Charles Bell of the U.S. Soil Conservation Service, Colusa County; Mr. Jack Bertholot and Mr. Jim Williams of the California Department of Water Resources for their effort in Fresno County; and Mr. Paul Lavine and Mr. Bill Weir of the U.C. Cooperative Extension Service in Stanislaus and Merced Counties, respectively.

We would also like to express our appreciation to several members of the U.C. Berkeley Remote Sensing Research Program staff who also contributed valuable time and effort to this study. These include: Ms. Julie Odenweller, Mr. Stephen DeGloria, Mr. Paul Ritter, Mr. Agnis Kaugars, Mr. Tony Travlos, Ms. Claire Hay, and Ms. Betsy Ross. Our thanks also to Mr. Stephen Harui of Peninsula Photographics for his effort in producing prints from aerial photographs for the California Crop and Livestock Reporting Service.

TABLE OF CONTENTS

Page

I. INTRODUCTION

In early 1982, the U. S. Department of Agriculture (USDA) joined with the National Aeronautics and Space Administration (NASA) and the State of California in sponsoring a remote sensing technique development study in California. The primary goal of this effort was to determine the extent to which remote sensing data could be used in various state and federal agricultural information programs within California. A complementary goal was to explore the possibility of sharing this technology between state and federal users.

The participants in the resulting cooperative study include the USDA Statistical Reporting Service (SRS), acting through its Remote Sensing Branch (RSB); the USDA-SRS California State Statistical Office (SSO), acting on behalf of the California Department of Food and Agriculture; the California Department of Water Resources (DWR); the NASA Ames Research Center; and the University of California at Berkeley (UCB), Remote Sensing Research Program. This report describes and summarizes work performed at UCB during 1982 in support of this effort under USDA/SRS sponsorship.*

---

*USDA/SRS Cooperative Agreement Number 58-319T-2-0341X

II. UCB PLANNING PHASE ACTIVITIES

A. Identification of Estimation and Mapping Objectives

Estimation and mapping objectives for the Cooperative Study were reviewed and refined during the first part of 1982 through continued communication with personnel from USDA, DWR, and NASA-Ames Research Center. This activity resulted in the formulation of several University of California at Berkeley experiments described in Appendix I. These were later grouped into three major task areas in the subsequent USDA-UCB research agreement. These task areas included:

1) Development of efficient procedures for small grains mapping and estimation, including irrigated acreage breakdown;

2) Development of efficient procedures for multicrop, end-of-season classification and acreage estimation; emphasis to be placed on

   a) new classification strategies

   b) sampling strategies for minor crops

   c) inventory and mapping techniques for small-area estimation of minor crops; and

3) Development of efficient early-season classification procedures and estimation techniques for selected crops.

Emphasis in each task area was to be placed initially on techniques appropriate to statewide estimation. Later, as techniques were developed and tested, the planned focus would shift to acreage estimation at the county or county group levels.

In order to accomplish these technique development goals, a phased series of estimation objectives were proposed for each task area. These are shown in Table 1. Each phase corresponds roughly to a level of technique difficulty required to achieve the associated estimation objectives. Development was proposed to proceed in a stepwise fashion through each succeedingly difficult phase.

B. Formulation of An Approach and Supporting Tasks Necessary to Accomplish Objectives

A preliminary approach was defined for each UCB Experiment during January 1982 (see Appendix I). These were refined in cooperation with the USDA-RSB and stated in general terms in the research agreement. More detailed definition of supporting tasks followed during late winter and spring. Supporting tasks included 1982 data acquisition, summary and analysis of factors affecting crop development in California, and initial crop classification technique development, which together consumed most of UCB's effort during the 1982 calendar year. Work on each of these tasks is described in the following sections.

TABLE 1. Area Estimation Technique Development Objectives for the Cooperative California Landsat Study

ACTIVITY PHASE	SMALL GRAINS	MULTICROP EOS	MULTICROP PRIOR-EOS (JES TO EOS TIME FRAME)
1	1. Total small grains acreage @ state and regional levels 2. Total wheat & total barley acreage at state & regional lvls 3. Total irrigated small grains acreage @ county levels 4. Total wheat and total barley irrigated acreage @ state & regional levels	1. Area estimates for cotton, rice, tomatoes, corn and alfalfa @ the state & regional levels	Develop preliminary techniques for area estimation . Cotton . Rice . Alfalfa . Pasture
2	1. Total wheat and total barley acreage @ county level 2. Total wheat and total barley, irrigated acreage @ state and regional levels 3. Optional: similar estimates for oats	1. Total acreage of above crops @ county level 2. Total acreage of other crops of significant areal extent @ state & regional levels, e.g.: dry beans, pasture, sugar beets, sorghum 3. Irrigated acreage estimates for crops in #1 or #2 if different from total acreage for those crops 4. Area of major cover categories: small grains, field crops, vegetables, vines, orchards, pasture @ count unit & all other levels	1. Acreage estimates for above @ state & regional levels 2. Technique development for other crops . Beans . Melons . Sorghum . Other
3	1. Total wheat and total barley acreage @ subcounty levels 2. Total wheat and total barley irrigated acreage @ county level	1. Area of other crops as above - @ county level 2. Area of selected minor crops @ state & regional levels 3. Area of major & significant other crops @ the DWR Data Analysis Unit level	Acreage of major crops @ county level

Footnote: @ means: "Technique defined for use at"

III. COLLECTION OF DATA NECESSARY TO SUPPORT DEVELOPMENT AND TEST OF LANDSAT-AIDED CROP ACREAGE INVENTORY TECHNIQUES

A. Acquisition of 1982 Landsat Data

A major activity during the spring and summer time period involved:

1) specification of the Landsat scenes for which data would be acquired in support of the Cooperative Study;

2) tracking the availability and quality of this data; and

3) establishing acquisition priorities and responsibilities among cooperators. This effort required a significant amount of communication with other cooperators, especially with the USDA-RSB and NASA-Ames. A list of available Landsat scenes was maintained that included EROS Data Center codes for image quality and cloud cover, and codes showing Landsat products acquired by each cooperator.

The important first step in this process was to specify a test site in California that would provide an effective test-bed for candidate classification and sampling techniques, while at the same time limiting processing requirements to control costs. Three contiguous Landsat scenes covering a significant portion of the Sacramento Valley and the upper San Joaquin Valley were selected for this purpose. Subsequent to this recommendation a line-start problem developed on Landsat 3, resulting in the elimination of data for the western one-third of each Landsat scene. In order to compensate for this data loss, the number of Landsat scenes was expanded to five. Sixteen contiguous counties were selected within this five scene area (see Figure 1) to represent the formal test site. These counties were chosen on the basis of the amount of agricultural land present, the relative importance of the crops present to USDA and DWR, and the degree to which they contributed to a coherent test site.

To assist in establishing Landsat acquisition priorities, a significant effort was undertaken to predict which Landsat passes would contain the most valuable information regarding crop separation. This work (summarized in Appendix II) was considered critical to obtaining the most valuable Landsat data before 1 October, the date on which EROS instituted a substantial price increase.

The bulk of the Landsat data was acquired by USDA-RSB and NASA-Ames. NASA-Ames assistance was especially important in supplementing USDA's acquisitions with Fall '81, Spring '82, and Summer '82 Landsat CCT's and transparencies required for new technique development.

During the fall of 1982, EROS was contacted several times regarding the nature and impact of its preprocessing techniques on Landsat digital data. UCB was concerned that geometric correction by cubic convolution resampling (the EROS standard) might alter "raw" spectral response in each of the four MSS bands in such a way as to significantly affect greenness-brightness measures and classification performance. It was determined

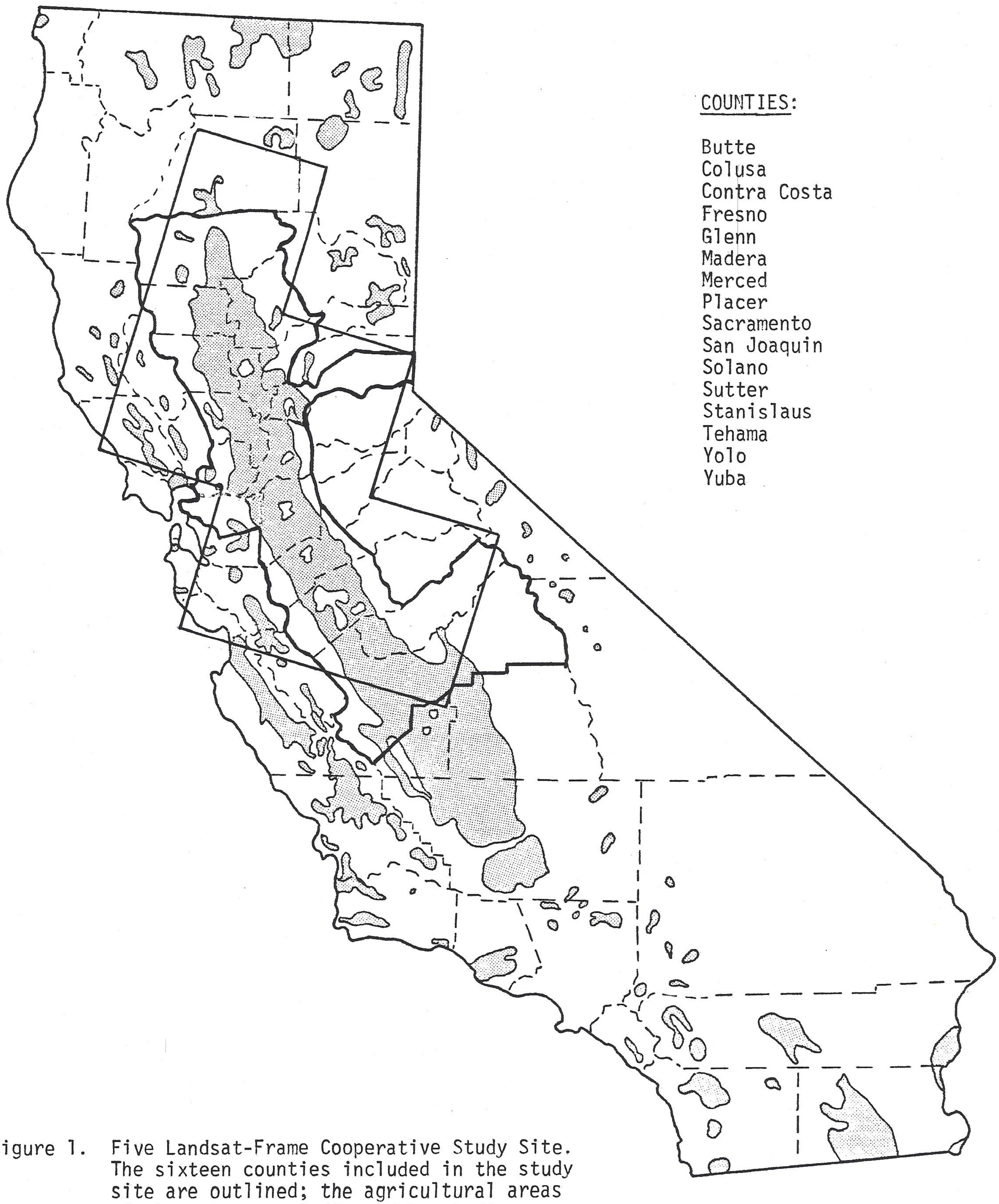


Figure 1. Five Landsat-Frame Cooperative Study Site. The sixteen counties included in the study site are outlined; the agricultural areas of the state are indicated by shading.

that the impact of this resampling procedure was probably most pronounced at ground class boundaries. In EROS's experience, the actual reflectance shift was minimal (1 to 2 counts) even in these situations. Since the current USDA-RSB procedure employs cubic convolution resampled data, it was decided to maintain this pattern - but at the same time to explicitly recognize this resampling step as a source of variation during technique development.

B. Acquisition of 1982 Ground and Aircraft Data

1. June Enumerative Survey (JES) Data

UCB assisted JES data collection at a number of points. These included coordination with the USDA-RSB on the form of the JES questionnaire for segments to be included within the study site. Later, UCB participated with the other cooperators in specifying counties in which the specialized recording forms would be used. Representatives from UCB also attended the first day of enumerator training sessions and made brief presentations regarding the nature of the Cooperative Study at each.

UCB shared with the SSO the responsibility for preparing the 1982 JES segments for digitization. UCB's primary activity was locating the JES segments on current (Spring '82) U-2 color infrared aerial photography provided for the Cooperative Study by NASA-Ames. Approximately 215 segments were annotated on the U-2 photography using older black-and-white photos and maps provided by the SSO; for another 140 segments where black-and-white photos were not available, lists were prepared giving the U-2 frame number where the segment should be found. In addition to the annotated photos and lists, photo center flight lines for the entire U-2 mission were provided to the SSO.

For a small number of segments, current enlarged aerial photography was needed for the actual digitization process being done by the California Department of Water Resources. UCB provided the SSO with examples of U-2 photography enlarged to 1 mile = 8 inches, 1 mile = 6 inches, and 1 mile = 4 inches. Following the SSO's evaluation, UCB coordinated the enlargement of 18 segments to a scale of 1 mile = 6 inches.

UCB also shared in the actual editing of a limited number of segments. The participation in the editing was particularly valuable for UCB in understanding the detailed nature of the editing procedure and its place in the processing flow of the USDA classification methodology.

Because California's agriculture is comparatively complex relative to much of the U.S., UCB organized two one-day field trips for SSO personnel. The trips were designed to help answer questions arising during the editing, give some insight to potential classification problems, and provide an opportunity to interact with other Cooperative Study participants. Charles Ferchaud of the California Department of Water Resources provided the vehicles and led the field trips. The combination of routes and dates (August 25 and October 21) on each trip provided the opportunity to see much of the agricultural diversity found in the 16-county test site.

UCB also consulted with NASA-Ames regarding desired aircraft coverage requirements for small grains and other crops during the spring and summer of 1982. Two NASA-Ames highflight missions resulted. The imagery from the first of these was used to supplement the JES editing process. Finally, UCB provided a limited amount of programmer consulting to DWR and NASA-Ames regarding JES segment digitizing software and equipment at DWR.

2. Landsat Pass-Specific Ground Observation Data

a. Uses of Ground Data

Ground data collection throughout the 16-county area was intended for the following uses:

(1) To generate location-specific crop calendars for each of the five Landsat scenes, as crop development varies throughout the Central Valley

(2) To develop basic awareness of cropping and cultural practices in order to resolve labelling anomalies

(3) To identify physical factors affecting spectral response, such as canopy density, biostage, surface moisture, and row direction

(4) To identify relationships between percent ground cover and biostage by crop type

(5) To develop relationships between physical factors and spectral measures, such as Tasselled Cap greenness/brightness or 7/5 vegetation indicators

(6) To provide a basis for optimum selection of Landsat acquisitions, through awareness of opportunities for separation of confusors based on crop development sequences

(7) For defining classification procedures, and for modifying procedures following testing and analysis.

b. Frequency of Observations

Given the diversity and variability within the California agricultural environment, it was anticipated that efficient classifier development and evaluation would depend to a significant extent upon correlation of field conditions with coincident Landsat spectral data. Unfortunately, uncertainty regarding weather conditions, satellite/receiving station problems, and data quality made it impossible to predict which overpasses would require supporting ground data. Therefore it was decided that ground data would be collected for every Landsat 3 overpass. This frequency of observation was also consistent with that required for monitoring two of the most important factors affecting crop spectral response - crop biostage and crop canopy development.

c. Participants

As the 16-county area was too extensive to survey every 18 days by the UCB staff, personnel from other agencies were asked for assistance. The following individuals were enlisted into the 1982 ground data crew:

- Department of Water Resources
    - Charles Ferchaud: Tehama, Butte, Glenn, Colusa & Yolo Counties
    - Jay Baggett: Yolo County
    - Jack Bertholot and Jim Williams: Fresno County
- U.C. Cooperative Extension Service
    - Paul Lavine: Stanislaus County
    - Bill Weir: Merced County
- U.S. Soil Conservation Service
    - Charles Bell: Colusa County
- U.C. Berkeley
    - Louisa Beck and Cathy Brown: San Joaquin County

Most began their observations in April, 1982, and continued through December, 1982. In all, 404 fields had 18-day ground data coverage. (See Table 2 for a breakdown of fields by crop type, and Figure 2 for the approximate locations of the seven transects.)

d. Types of Data Collected

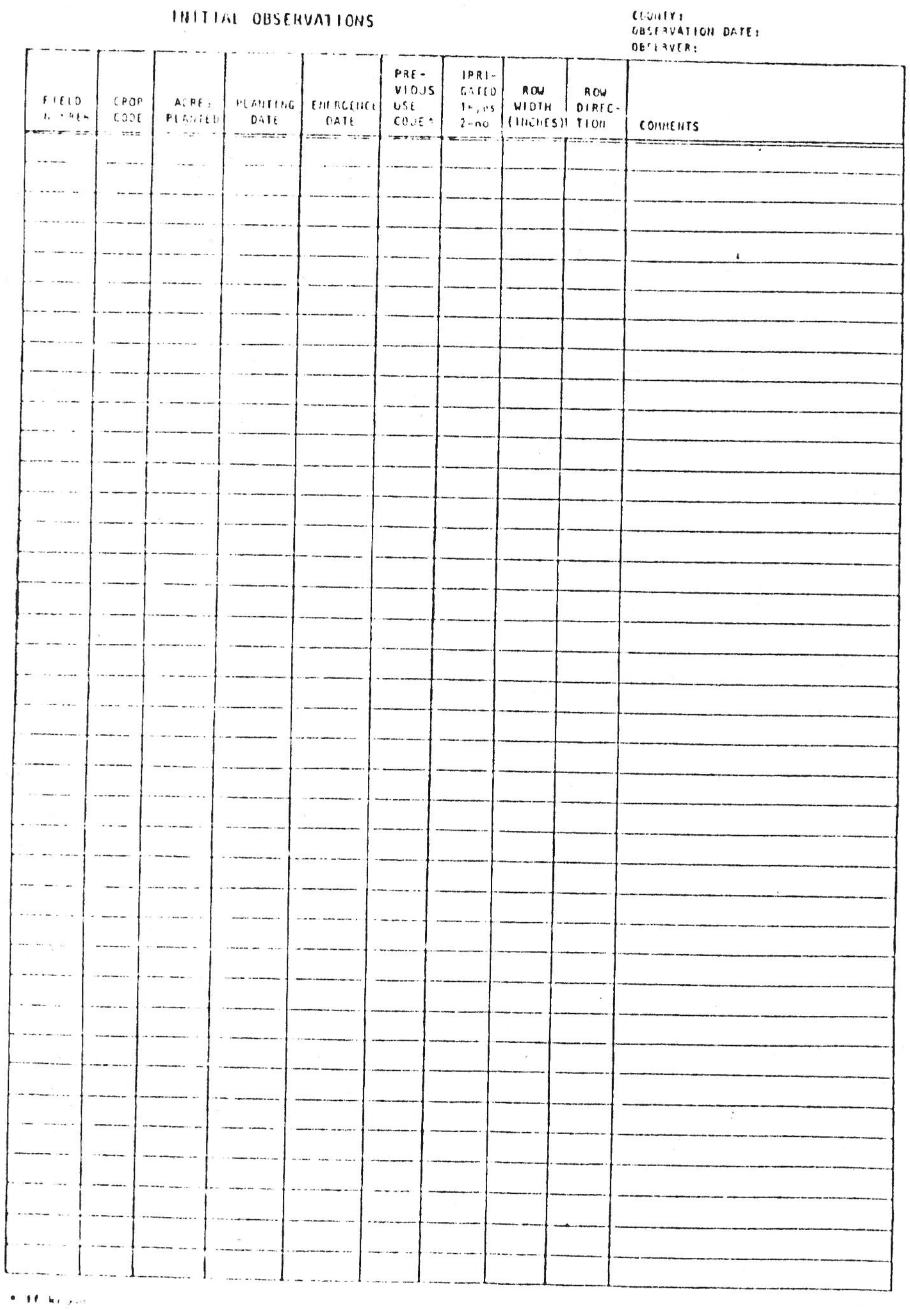
In order to accomplish the goals of the ground data collection effort, some basic types of data were required. These types of data were designed to take a minimum of training to collect, and recording sheets were designed such that field recordings could be expedited. Examples of the recording sheets are shown in Figure 3 and Figure 4. The types of data collected were as follows:

- Initial observations
    - crop code
    - planting date (if known)

Table 2: 1982 field observations for each county transect, showing a breakdown of fields by crop type.

	GR	DB	SR	CR	SB	TO	SU	TR	SF	OR/V	RI	CT	A/P	MISC.
BUTTE	3	6	3	5	5		2				1		1	
COLUSA														
CB	4	5	1	4	4	8			4		2		4	
CF	7	3	1	1	1	4							1	
FRESNO	6	1			3	6		1	1			12	7	
GLENN	8	4	1	3	3	2	2				12		5	
MERCED												3		
SAN JOAQUIN	9	14		9	4	12	1	6	2	4	4		5	
STANISLAUS		7		4		7		5					5	
TEMAMA	6	4	1	4	4								1	2
YOLO														
JB	33			17	7	17	1	2	2	4	6		13	6
CF	4			1		7					1		3	
TOTAL	80	44	7	48	31	63	6	14	9	8	26	15	45	8

Figure 2: Approximate locations of the ground data collection transects within the five Landsat frame study site.



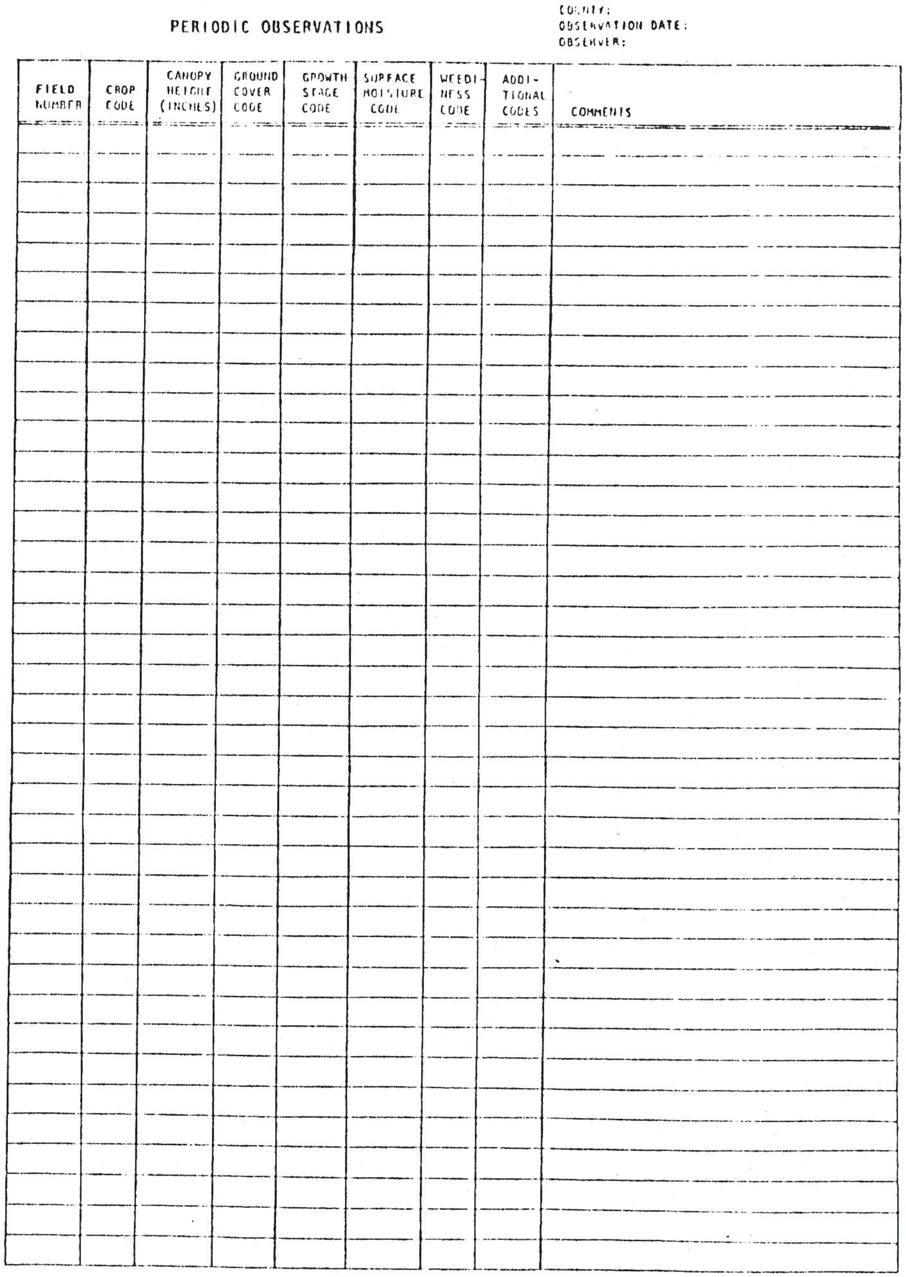
INITIAL OBSERVATIONS

COUNTY:
OBSERVATION DATE:
OBSERVER:

FIELD NUMBER	CROP CODE	ACRES PLANTED	PLANTING DATE	EMERGENCE DATE	PREVIOUS USE CODE*	IRRIGATED 1=yes 2=no	ROW WIDTH (INCHES)	ROW DIRECTION	COMMENTS

* If [illegible]

Figure 3: An example of the form used by the field crew for recording initial field observations.



PERIODIC OBSERVATIONS

COUNTY:
OBSERVATION DATE:
OBSERVER:

FIELD NUMBER	CROP CODE	CANOPY HEIGHT (INCHES)	GROUND COVER CODE	GROWTH STAGE CODE	SURFACE MOISTURE CODE	WEEDINESS CODE	ADDITIONAL CODES	COMMENTS

Figure 4: An example of the recording form used for 18-day observations.

- emergence date (if known)
- previous use (if known)
- irrigation practice
- row width
- row direction

- Periodic observations - 18-day
  - canopy height
  - percent ground cover
  - biostage
  - surface moisture
  - weediness
  - comments pertaining to field condition

Field observers were instructed to set up their own transects for 18-day coverage which they could cover in four to six hours, in most cases. The transects were to encompass the variation in crop type, as well as variation within crop type. As each observer was working his/her area of familiarity, he or she would be the best judge of this variability. The observations were to be made from field edges, as permission was not sought to enter fields.

e. Products to Date

The field observations have been used by personnel at UCB to (1) assist in the formulation of recommendations for purchase of Landsat scenes prior to October 1, 1982 (Beck, 1982) and to (2) begin the development of spectral crop calendars for 1982. The preliminary calendars, which are shown in Tables 3 through 10 are based on the presence or absence of Landsat-detectable green canopy cover. In order to be considered "detectable", the recorded ground cover percent had to equal or exceed 20 percent*. For each date (which corresponded to a Landsat 3 overpass), all fields were checked for the presence of detectable canopy cover. If the majority of fields of a given crop type was above the threshold, a "1" was recorded for that particular crop on that date. All crops below the threshold were given a "0".

In many cases, there was variation within crop type that precluded the use of a single binary code. Tomatoes, for example, exhibited a high degree of between-field biostage and canopy cover variability as a result of planting and harvesting date ranges that were spread over several months. Therefore, both an "0" and a "1" were assigned to the crop when a mix occurred.

---

*A threshold of Landsat crop canopy detection, based on work reported by Rice et al., 1979.

These calendars will be used later for correlation of canopy cover with Tasselled Cap data. It should be noted that the (0,1) sequences were based, in many cases, on a relatively small number of fields per crop type, and may not necessarily encompass the complete range within (or between) crop types.

The data collected during 1982 will also be used for other spectral analyses scheduled for 1983. These analyses will focus on establishment of relationships between field/crop condition and spectral response, and upon the development of classification methodology taking advantage of these relationships.

TABLE 3

TEHAMA

	4/28	5/17	6/3	6/21	7/10	7/29	8/16	9/9	9/18	10/8	10/27	11/18	12/13
grain	1	1	1-0	0	0	0	0	0	0	0	0	0-1	0-1
sorghum	0	0	0	1	1	1	1	0	0	0	0	0	0
corn	0	0	0	1	1	1	1	1-0	1-0	0	0	0	0
s beet	0	0	0	0	1	1	1	1	1	1	1	1	1
dry bean	0	0	0	0	1	1	1	0	0	0	0	0	0

Note: In Tables 3 through 10, a "1" indicates the presence of Landsat-detectable green vegetation ($\geq$20 percent ground cover), and a "0" signifies the absence of detectable vegetative cover. These tables were generated using the 1982 ground data that were collected coincident with Landsat 3 overpasses.

TABLE 4

BUTTE

	4/28	5/17	6/3	6/23	7/10	7/29	8/20	9/9	9/17	10/9	10/27	11/17
grain	1	1-0	0	0	0	0	0	0	0	0	0	0
corn	0	0	0	1	1	1	1	1-0	0	0	0	0
sunflwr	0	0	1	1	1	1	1-0	1-0	0	0	0	0
rice	0	0	1	1	1	1	1	1	1	1-0	0	0
s beet	0	0	0	1	1	1	1	1	1	0	0	0
dry bean	0	0	0	0	1	1	1	1	1-0	1-0	0	0
sorghum	0	0	0	1	1	1	1	1	1	0	0	0

TABLE 5

GLENN

	4/28	5/19	6/3	6/21	7/9	7/28	8/17	9/8	9/17	10/8	10/25	11/18
grain	1	1	1-0	0	0	0	0	0	0	0	0	0
s beet	0	1	1	1	1	1	1	1-0	1-0	0	0	0
rice	0	0	0-1	1	1	1	1	1	1-0	0	0	0
tomatoes	0	0	0	1	1	1	1	1-0	1-0	1-0	0	0
corn	0	0	1	1	1	1	1	0	0	0	0	0
dry bean	0	0	0	0	1	1	1-0	1-0	0	0	0	0
sunflwr	0	0	0	0-1	1	1	1	1	0	0	0	0
sorghum	0	0	0	0	1	1	1	1	1	0	0	0

TABLE 6a

COLUSA (C. Ferchaud)

	4/28	5/19	6/4	6/21	7/9	7/27	8/17	9/8	9/20	10/8	10/26	11/17
grain	1	1	1-0	0	0	0	0	0	0	0	0	0
s beet	1	0	0	0	0	1	1	1	1	1	1	1
tomatoes	0	0-1	0-1	1	1	1-0	1-0	1-0	0	0	0	0
dry bean	0	0	0	0	0	1	1	1	1-0	0	0	0
corn	0	0	0	0	0	1	1	1	1	0	0	0
sorghum	0	0	0	0	0	1	1	1	1	0	0	0

TABLE 6b

COLUSA (C. Bell)

	4/12	4/29	5/17	6/4	6/22	7/9	7/27	8/17	9/1
grain	1	1	0	0	0	0	0	NA	0
tomatoes	0	0	1	1	1	1	1	NA	0
s beets ow/spr	1/0	1/0	0	1	1	1	1	NA	1
rice	0	0	0	1	1	1	1	NA	1-0
barley	1	1	1	0	0	0	0	NA	0
safflwr	0	0-1	1	1	1	1	1-0	NA	0
corn	0	0	0-1	1	1	1	1-0	NA	0
dry bean	0	0	1	1	1	1	1-0	NA	0

TABLE 7a

YOLO (C. Ferchaud)

	4/28	5/19	6/4	6/23	7/9	7/27	8/20	9/8	9/20	10/8	11/17
grain	1	1	0	0	0	0	0	0	0	0	0
tomatoes	0	0	0	1	1	1	1-0	1-0	0	0	0
o.w.beet	1	0	0								
corn	0	0	0	1	1	1	1	1-0	0	0	0
rice	0	0	0	1	1	1	1	1	1-0	0	0

TABLE 7b

YOLO (J. Baggett)

	5/3	5/17	6/4	6/22	7/9	7/28	8/18	8/25	9/20	10/7	10/27	11/17
grain	1	1-0	0	0	0	0	0	0	0	0	0	0
tomatoes	0	0	1	1	1	1	1-0	1-0	0	0	0	0
s beets ow/spr	1/0	1-0/0	1-0/0	1-0/1	0/1	1	1	1	1	1	1-0	1-0
corn	0	0-1	1	1	1	1	1-0	1-0	1-0	0	0	0
rice	0	0	0-1	1	1	1	1	1	1-0	0	0	0
sunflwr	0	0	1	1	1	1	1-0	0	0	0	0	0
safflwr	0	0	1	1	1	1	1-0	0	0	0	0	0
dry bean	0	0	0	0	0	1	1	1	1-0	0	0	0

TABLE 8

SAN JOAQUIN

	3/25	4/12	4/29	5/17	6/3	6/22	7/9	7/28	8/17	9/1	9/30	10/8
grain	1	1	1	1-0	0	0	0	0	0	0	0	0
tomatoes	0	0	0	0	0-1	0-1	0-1	1	1-0	1-0	1-0	0
s beets ow/spr	1/0	1/0	1/0	1/0	1/0	1-0/0-1	1-0/1	0/1	1	1	1	1
corn	0	0	0	0-1	1	1	1	1	1	1	1-0	0
rice	NA	NA	NA	NA	NA	NA	1	1	1	1	1	1
dry bean	0	0	0	0	0	0	0-1	0-1	1	1	1-0	1-0
peppers	0	0	0	0	0	0	0	1	1	1	1	1-0

TABLE 9

STANISLAUS

	5/17	6/4	6/22	7/9	7/28	8/16	9/2	9/20	10/7	10/26	11/12	12/1
tomatoes	0	0	1	1	1	1-0	0	0	NA	0	NA	0
corn	0	1	1	1	1	1	1-0	0	NA	0	NA	0
dry bean	0	0	0	0-1	1	1	1-0	0	NA	0	NA	0
melons	0	1	1	1	1-0	1-0	0	0	NA	0	NA	0

TABLE 10

FRESNO

	4/28	5/17	6/3	6/21	7/9	7/27	8/16	9/2	9/20	10/7	10/26	11/12
grain	1	1	0	0	0	0	0	0	0	0	NA	0
barley	1-0	0	0	0	0	0	0	0	0	0	NA	0
tomatoes	0	1	1	1	1	0	0	0	0	0	NA	0
s beet	1	1	1	1	1	1	1-0	0	0	0	NA	0
cotton	0	0	0	0-1	1	1	1	1	1-0	1-0	NA	0
dry bean	0	0	0	0	0	0	1	1	1	1	NA	0
safflwr	1	1	1	1	1	0	0	0	0	0	NA	0

IV. DEVELOPMENT OF PROCEDURES FOR LANDSAT-AIDED CROP GROUP AND CROP TYPE CLASSIFICATION

A. Development of Procedures for Classification of Small Grains

1. Background, Objectives, General Approach

This effort represented the initiation of development and test activity under UCB Experiment #1. The initial objective was to develop a digital classification procedure that would mimic the manual technique developed by Jay Baggett and Charles Ferchaud at DWR. There were two reasons for selection of this approach to small grains classification. First, the manual procedure had worked extremely well - superior to any digital classification to date. Second, there was a need to gain another perspective on the classification process, one that would focus greater attention on the spectral change information utilized by the manual technique, and one that would provide, more generally, information concerning the manner in which spectral data should be partitioned in order to obtain the best performance.

An approach to digital classification based on Boolean combination of spectral reflectance data from several dates was selected as the initial vehicle for this work. The procedure was called polygon vector classification (or simply, polygon classification) since it involved the partitioning of two-dimensional distributions by date into class polygons, and Boolean combination of these polygons over dates to form polygon vectors. This procedure had its roots in earlier UCB work by Hay in LACIE (there called Delta Function Stratification) and by Hay and Odenweller in the corn/soybeans portion of FCPF AgRISTARS. Polygon classification represented an attempt to generalize these earlier techniques and thereby enable mimicking (within limits) manual small grains classification.

The polygon approach also enabled the determination of classification performance associated with non-parametric definition of clusters. At a more basic level, polygon vector classification was viewed as a tool for enabling the documentation of characteristics of spectral distributions by date and their probable physical causes (in conjunction with periodic field observation data). From these data, it was anticipated that information would be forthcoming relative to features useful in distinguishing crop types and to appropriate methods for cluster partitioning.

Development of the polygon classification technique proceeded on two previously acquired data sets. These were the 1980-81 Yolo County small grains data set obtained through NASA-Ames, and a 1976 multidate data set for a portion of Sacramento Valley prepared during the California Irrigated Lands Project. A decision was made to use a Tasselled Cap (also known as "Kauth") representation of the Landsat spectral data exclusively in Yolo County

(1981) based on its theoretical advantages (Kauth and Thomas, 1976, Kauth et al 1979) and its demonstrated utility during the past several years (Hall 1982, Cicone, Hay, et al 1981). Within the 1976 Sacramento data set, however, simple band combination (e.g. the MSS band 7 to band 5 ratio) was retained as an alternative spectral representation (see, for example, Wall et al, 1981) along with the Tasselled Cap transformation. This provided an opportunity for later assessment of performance differences resulting from the use of the two types of spectral bands.

2. Small Grains Classification Technique Development Using the 1981 Yolo County Data Set

a. The 1981 Yolo County Data Set

The Yolo County Data Set included registered Landsat MSS data for the following dates*:

- 17 November (1980)
- 7 April
- 13 May
- 31 May
- 6 July

In addition to the spectral data, the data set contained twenty-five USDA-JES segments with 1981 DWR land use survey ground data, and the USDA/SRS Yolo County land use stratification. Both the ground data and stratification were registered to the MSS data.

b. Data Preparation

Prior to polygon vector analysis, all MSS pixels were normalized (using a sun angle correction algorithm) and processed using the Tasselled Cap Transformation. The Tasselled Cap greenness and brightness bands for each date were subsequently used for all analysis, as they accounted for nearly all the spectral variation found in the four MSS bands, and required half the processing time and storage space. In order to reduce the number of pixels for intensive processing, a subset of the JES segments were selected for training and analysis. Figure 5 shows the approximate location of the seventeen sample segments used, and the USDA strata in which the segments fell.

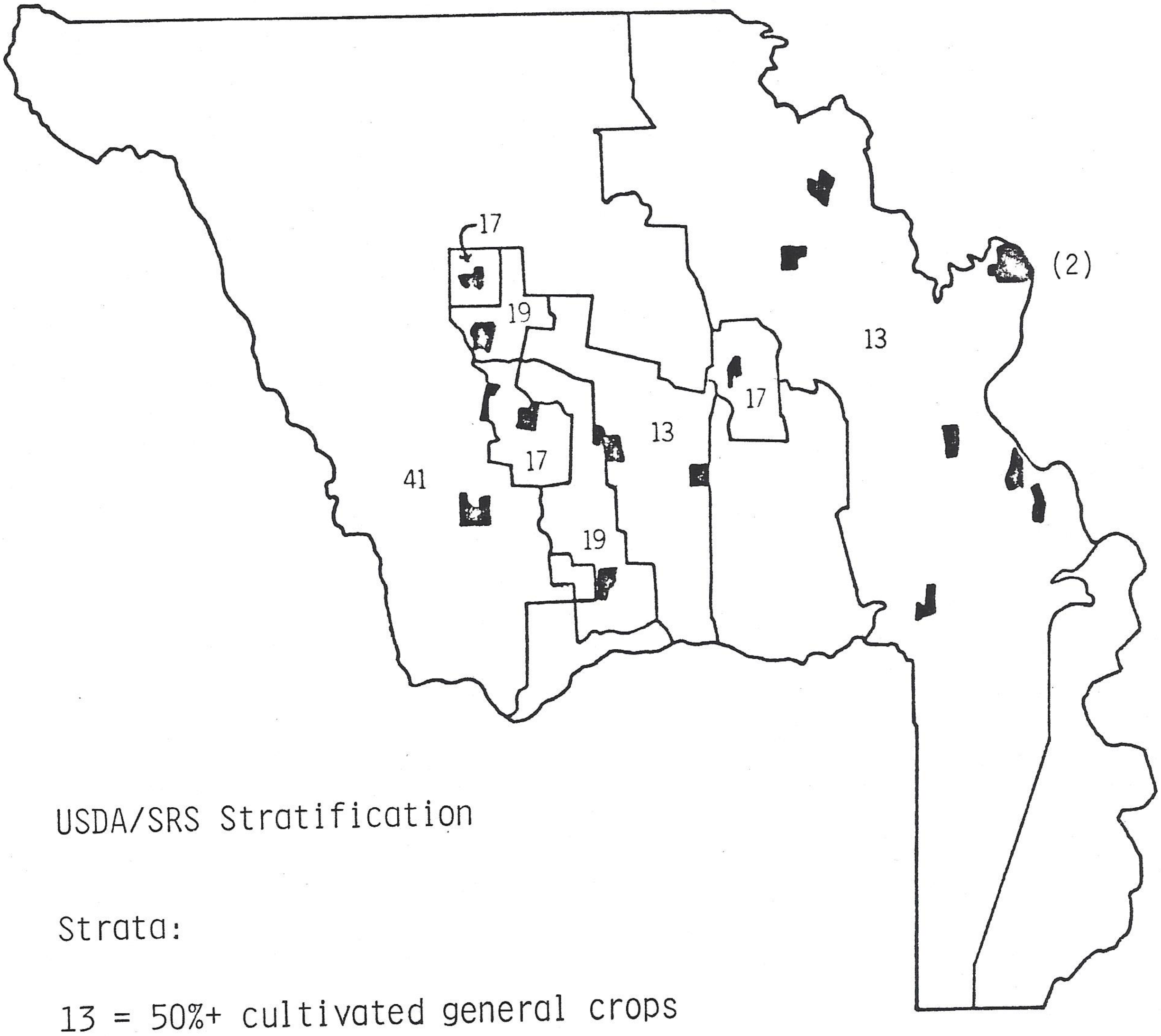
Using the segment ground data to isolate grain fields, scatterplot displays were generated for irrigated grain, non-irrigated grain, and non-grain pixels. Two types of scatterplots were created: greenness Date i versus greenness Date i + 1 for fixed polygon analysis, and greenness Date i versus brightness Date i for a variable polygon procedure.

c. Scatterplot Partitioning

The primary goal of the polygon partitioning procedure was to differentiate grain from its confusors. These confusors,

---

*Registration was performed by personnel at NASA-Ames Research Center



USDA/SRS Stratification

Strata:

13 = 50%+ cultivated general crops

17 = 50%+ cultivated fruits, nuts, grapes

19 = 50%+ cultivated general crops/vegetables

41 = privately owned range, less than 15% cultivated

Figure 5. Approximate location of the JES segments used in this study, along with their respective stratum locations.

which include overwintered sugar beets, native vegetation, and pasture, have periods of green-up and maximum canopy cover which overlap those of grain. Therefore, the polygon work was concerned with separating grain from other crop types by utilizing (a) a temporal difference in canopy development and/or harvest, (b) a consistent difference in maximum greenness values, or (c) a combination of temporal and spectral shifts.

In order to characterize the spectral behavior of grain in Tasselled Cap space, two types of scatterplot partitioning techniques were employed. The first type, called Fixed Polygon Partitioning, used two-date greenness scatterplots upon which a predetermined set of polygons was superimposed. The second type of partitioning required the analyst to define custom polygons for grain and non-grain distributions on greenness-brightness scatterplots for each date. This last technique was called Variable Polygon Partitioning.

(1) Fixed Polygon Partitioning

(a) Method

The first partitioning technique tested utilized two-date greenness scatterplots in order to track grain's spectral and temporal movement through greenness space. The tracking was accomplished using a set of fixed greenness zones, or polygons, in which the presence or absence of grain pixels was noted.

Boundaries of the greenness zones were established by (1) setting a green vegetation detection threshold for each date and by (2) locating a line of equal greenness (no-movement between dates) for each date pair. Tasselled Cap green vegetation thresholds were examined in earlier AgRISTARS work (Hay et al 1979, Rice et al 1980, and Odenweller et al 1983). Using these AgRISTARS studies as a guide, a Tasselled Cap greenness value of six was selected to represent an initial threshold of vegetation detection. An interactive cursor procedure operating on a Data General NOVA system was used to define the greenness-equals-six boundary. A set of four polygons resulted for each two-date greenness scatterplot (see Figure 6). In order to differentiate between pixels that increased in greenness between Date i and Date i + 1 from those pixels that decreased, a line representing equal greenness for both dates was added. The result was a set of five polygons, as described in Figure 7.

Once these five polygons were defined, they were used to partition two-date greenness scatterplots for irrigated grain, non-irrigated grain, and non-grain

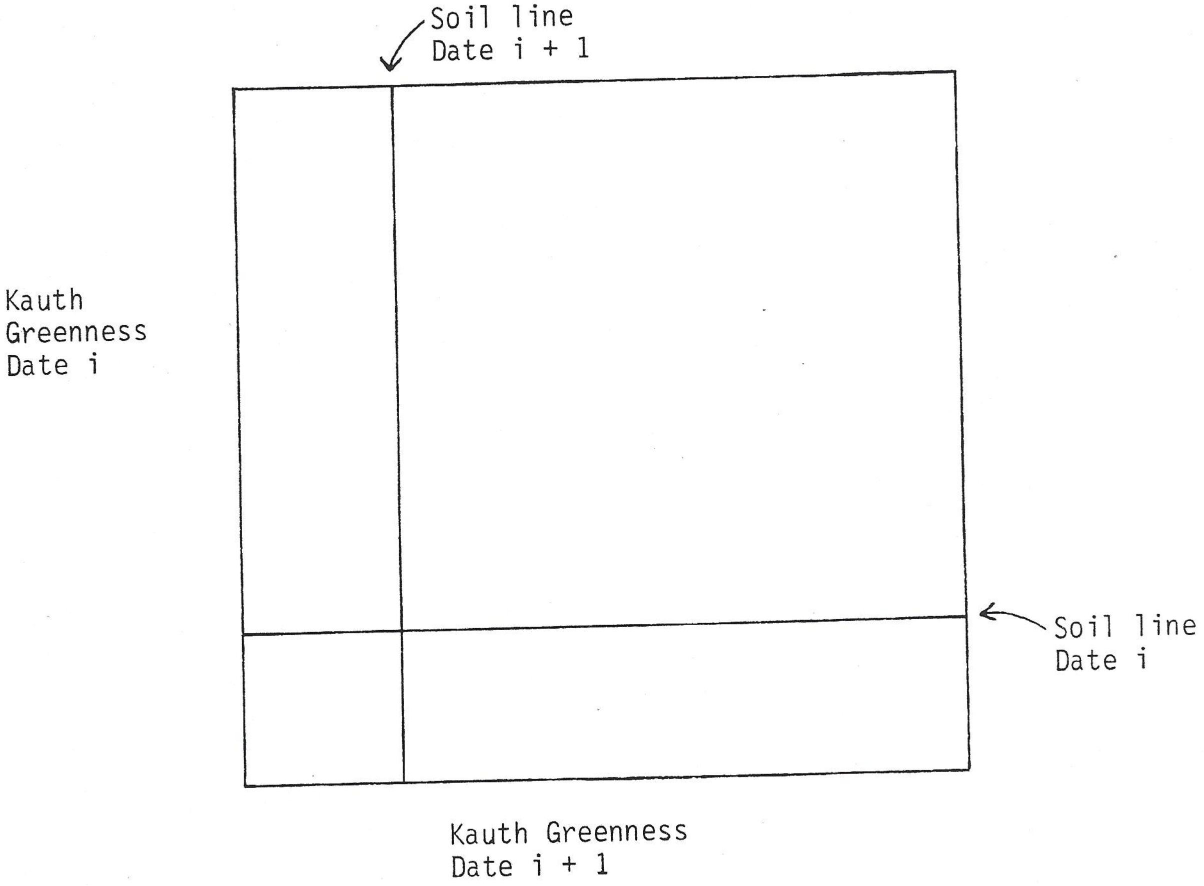
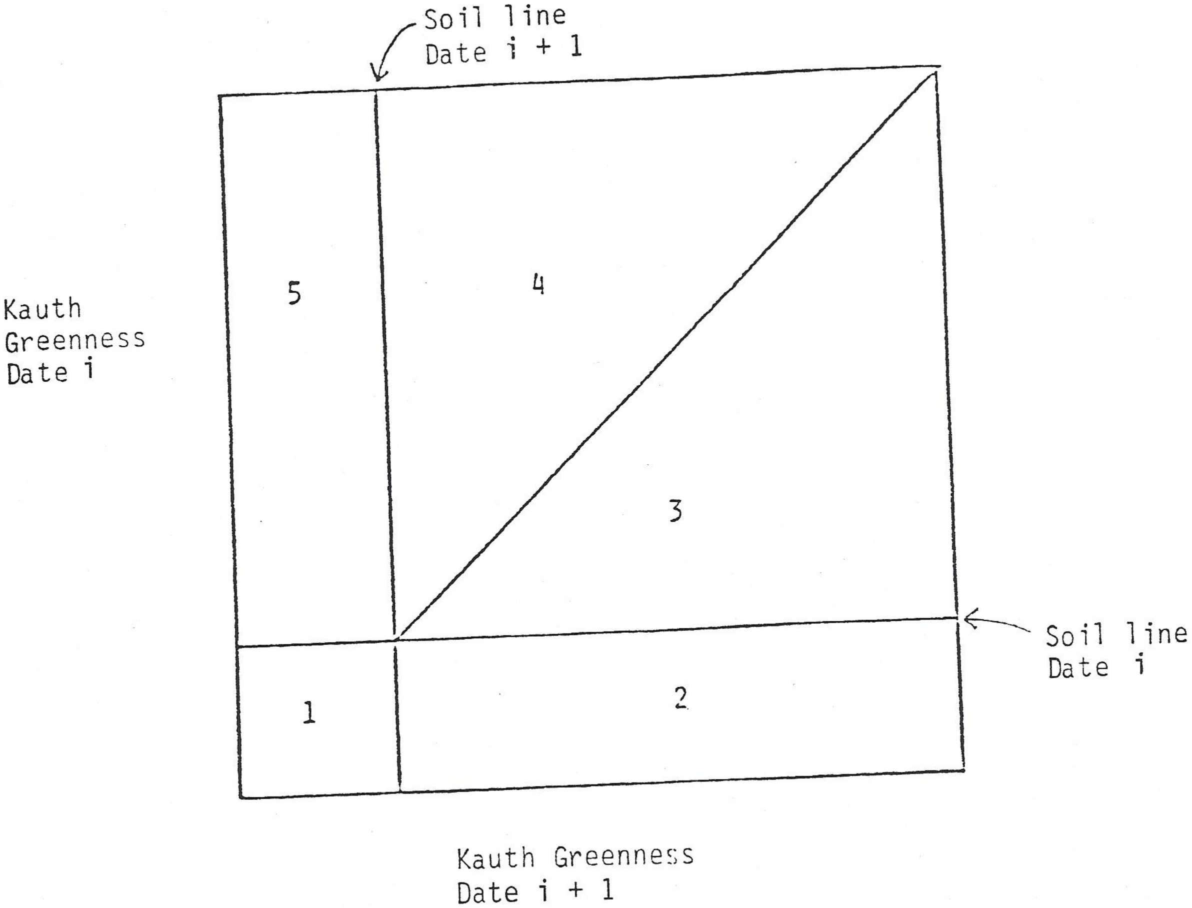


Figure 6. Set of four polygons resulting from specification of vegetation detection thresholds for two Landsat pass dates.



Polygon #	Contents:
1	Pixels that were below the threshold of detection on Dates i and i + 1
2	Pixels that were below the threshold of detection on Date i, above on Date i + 1
3	Pixels that increased in greenness between Date i and Date i + 1
4	Pixels that decreased in greenness between Date i and Date i + 1
5	Pixels that were green on Date i but dropped below the threshold of green detection on Date i + 1

Figure 7. Description of five polygons employed in fixed polygon partitioning on the Yolo County 1981 data set.

for the following date combinations:

- 14 November vs. 7 April
- 7 April vs. 13 May
- 13 May vs. 31 May
- 31 May vs. 6 July

For each successive date pair, the analyst superimposed the same five fixed polygons over the two-date irrigated grain scatterplot. Polygons containing the bulk of the irrigated grain pixels were labelled "iG". Next, the analyst displayed the scatterplot representing all non-irrigated grain for the same date pair. Again, polygons that contained the bulk of non-irrigated grain pixels were noted, and received a label of "nG". Where irrigated and non-irrigated grain coincided, a label of "Gr" was assigned. Finally, the analyst displayed the scatterplot representing all non-grain pixels for the same date pair. Polygons that contained both grain and non-grain pixels were labelled "mixed"; the remaining polygons were assigned a label of "other". The labels, which were assigned numeric equivalents, were then used by the computer as a digital mask in order to sort and label all pixels on the two dates corresponding to the scatterplot. The result of the machine labelling step was a two-date class map, called a polygon assignment band.

The process described above was then repeated until a polygon assignment band was created for all date pairs (Figure 8). Once all four assignment bands were created, the computer used them to track each pixel in the data set through the entire five-date sequence. As a result, four class labels were assigned to each pixel, one label for each date pair. This string of labels, called a vector, was then stored in a vector file in which each unique vector was assigned its own identification number. The frequency of identical vector occurrences was automatically accumulated using a multi-dimensional histogramming program.

The vectors were then automatically grouped according to the dominant cover label in that vector. For example, all vectors that contained three mixed labels were assigned to the mixed class; the order in which these labels occurred in the vector was incidental. The vector file, along with the associated class assignment, was then used by the computer to automatically classify all pixels in the data set.

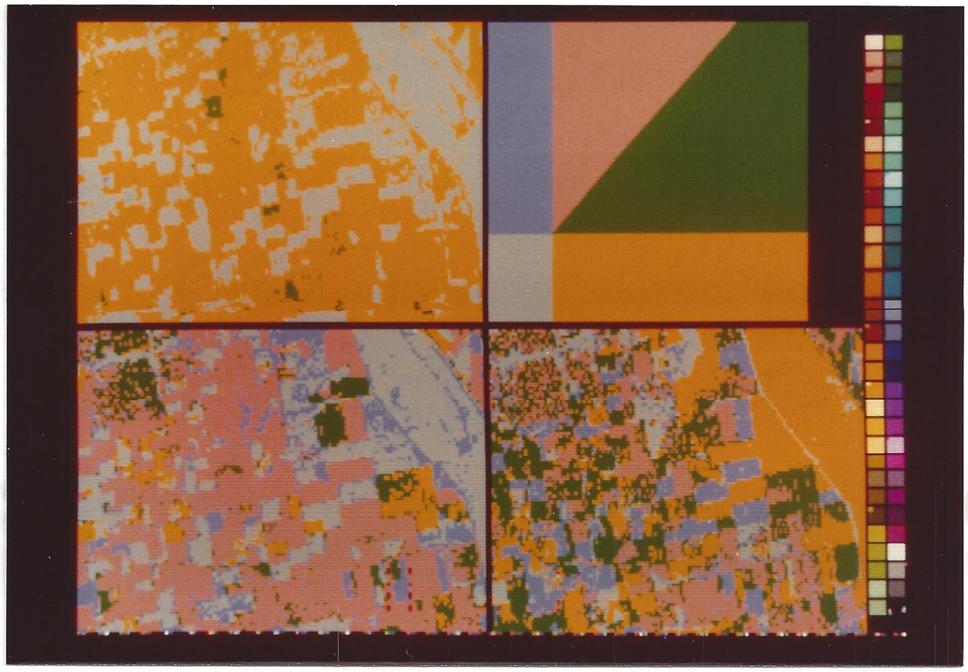


Figure 8. An example of three polygon assignment bands generated using the fixed polygon partitions shown in the upper right. The upper left assignment band represents the November/April greenness bands; the lower left represents the April/13 May assignment band; and the lower right shows the assignment band for 13 May/31 May. Refer to Figure 7 for a description of each polygon.

(b) Results

The results of the Fixed Polygon approach were poor. The final classified map was visually confusing, with a lack of field patterns. A preponderance of the pixels in the data set fell into the mixed class, with few pure grain or non-grain classes. The error was probably the result of the limitations posed by the fixed polygons themselves. The polygons were too general and were not sensitive to the actual greenness distributions of grain pixels throughout the date sequence. For example , both the irrigated grain and non-grain distributions occurred at the intersection of all five fixed polygons on several dates (Figure 9a & b). Rather than label all polygons "mixed" for two of the four date-pairs, the analyst had been forced to make judgements regarding where the cut-off should be to distinguish between dense and sparser pixel distributions.

(2) Variable Polygon Partitioning

(a) Method (see Figure 10 for processing flow)

This type of partitioning differed from the Fixed Polygon procedure in two important respects. First, the Variable Polygon technique employed a series of uni-temporal greenness-versus-brightness scatterplots rather than two-date greenness scatterplots. Second, instead of partitioning the scatterplots using a predetermined set of fixed zones, the analyst had to interactively draw a custom polygon boundary around the distribution of grain pixels on each date. In this way, a set of polygons could be tailored that would mimic the movement of grain through greenness-brightness space over time.

The most expedient way to create the custom polygons was to display one image in which the irrigated grain, the non-irrigated grain, and the non-grain greenness-brightness distributions for one date were represented in three different colors (red, green, and blue). The relative densities for each crop type resulted in various color combinations, allowing the analyst to locate areas in greenness-brightness space that were predominantly grain, non-grain, or mixed (Figure 11). Using an interactive cursor, these distributions were outlined and each polygon given the appropriate digital mask value (1 = other, 2 = mixed, 3 = grain). The analyst repeated this procedure for the first four dates, thereby creating four digital polygon masks (Figure 12).

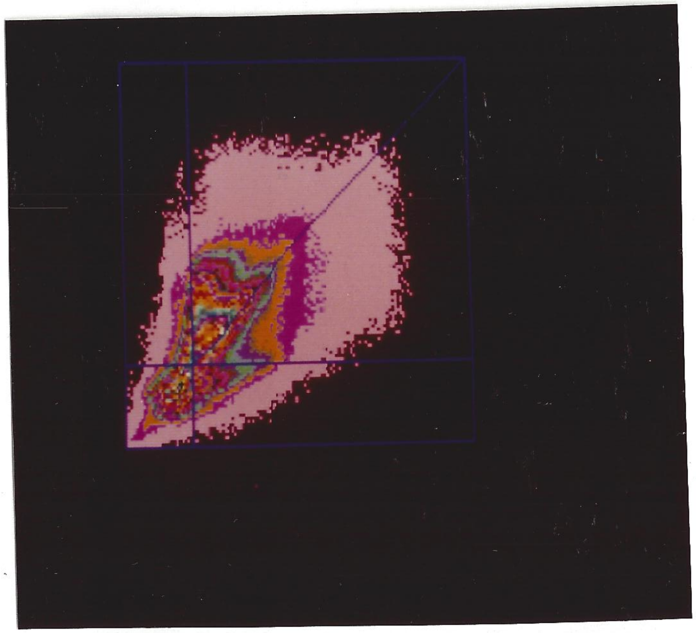


Figure 9a. Location of all crop pixels on the 13 May/31 May greenness scatterplot, with polygon overlay.

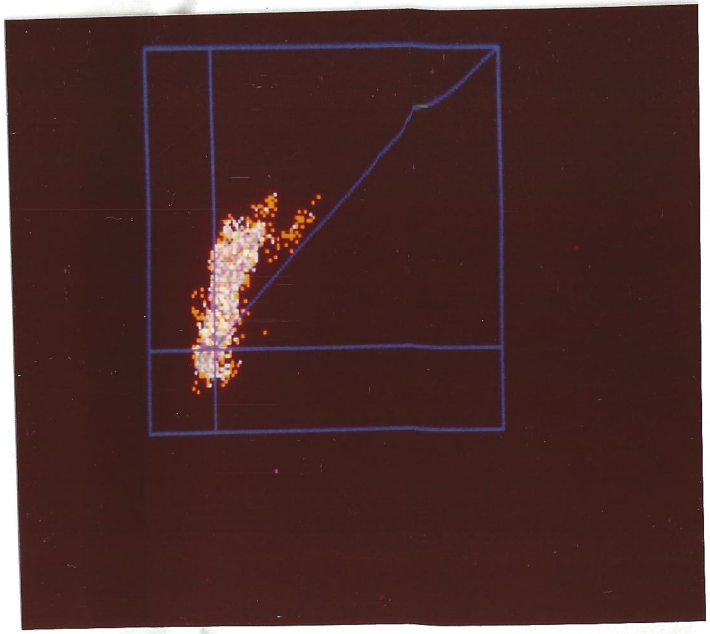
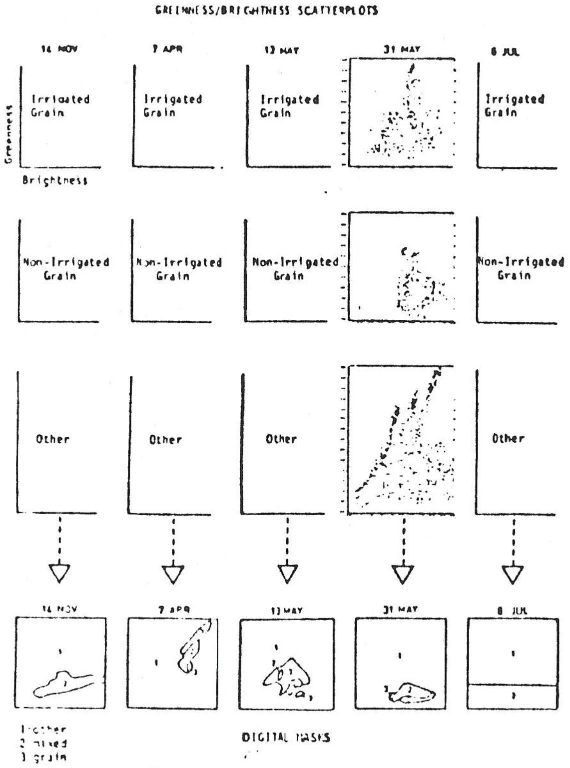


Figure 9b. Location of grain pixels on the 13 May/31 May greenness scatterplot with respect to the polygon overlay.

▶1. Define Polygon Masks



▶2. Generate Vector File

VECTOR NUMBER	COUNT	1 14 NOV	2 7 APR	3 13 MAY	4 31 MAY	5 6 JUL
1026	2122	2 GR	1 OT	1 OT	1 OT	1 OT
4766	1510	2 GR	2 GR	2 GR	3 GR	2 GR
1799	703	2 GR	3 GR	3 GR	3 GR	2 GR
1797	661	2 GR	2 GR	3 GR	3 GR	2 GR
4763	529	2 GR	3 GR	2 GR	3 GR	2 GR
4799	416	2 GR	1 OT	3 GR	3 GR	2 GR
1282	378	2 GR	1 OT	1 OT	1 OT	2 GR
770	317	1 OT	1 OT	1 OT	1 OT	1 OT
4797	303	2 GR	1 OT	1 OT	3 GR	1 OT
4774	294	2 GR	1 OT	2 GR	2 GR	2 GR
4778	294	2 GR	1 OT	3 GR	1 OT	1 OT
4792	277	2 GR	1 OT	2 GR	1 OT	1 OT
4767	236	2 GR	3 GR	2 GR	2 GR	2 GR
4798	226	2 GR	1 OT	1 OT	2 GR	2 GR
1286	224	2 GR	3 GR	1 OT	3 GR	2 GR
4758	216	1 OT	3 GR	1 OT	1 OT	1 OT
4744	215	2 GR	3 GR	1 OT	2 GR	2 GR
4800	211	2 GR	3 GR	1 OT	1 OT	2 GR
...	...	...	...	...	...	...
...	...	...	...	...	...	...

▶3. Generate Vector Sequences and Labels

VECTOR NUMBER	COUNT	MASK VALUE* 14 NOV	7 APR	13 MAY	31 MAY	6 JUL	VECTOR LABEL**	VECTOR SEQUENCES
1026	2122	2	1	1	1	1	1	OT 01111 GR 10000
4766	1510	2	2	2	3	2	2	GR 11111
1799	703	2	3	3	3	2	2	GR 11111
1797	661	2	2	3	3	2	2	GR 11111
4763	529	2	3	2	3	2	2	GR 11111
4799	416	2	1	3	3	2	3	GR 10111 OT 01000
1282	378	2	1	1	1	2	4	OT 01110 GR 10001
4789	361	2	1	2	3	2	3	GR 10111 OT 01000
4790	342	2	2	2	2	2	2	GR 11111
1284	326	2	1	1	3	2	5	GR 10011 OT 01100
770	317	1	1	1	1	1	6	OT 11111
4797	303	2	1	1	3	1	4	OT 01101 GR 10010
4774	294	2	1	2	2	2	3	GR 10111 OT 01000
4778	294	2	1	3	1	1	4	OT 01011 GR 10100
4792	277	2	1	2	1	1	4	OT 01011 GR 10100
4767	236	2	3	2	2	2	2	GR 11111
4798	226	2	1	1	2	2	5	GR 10011 OT 01100
1286	224	2	3	1	3	2	3	GR 11011 OT 00100
4758	216	1	3	1	1	1	1	OT 10111 GR 01000
4744	215	2	3	1	2	2	3	GR 11011 OT 00100
4800	211	2	3	1	1	2	5	GR 11001 OT 00110
4796	192	2	1	1	2	1	4	OT 01101 GR 10010
...								
...								
...								

* 1 = OT
2 = MX = GR
3 = GR

** 1 = Other
2 = Grain
3 = Grain
4 = Other
5 = Grain
6 = Other

▶4. Edit Vector Sequences and Labels

COVER CLASS		SEQUENCES (PARTIAL LIST)				
1	OT	OT 11110	11101	11011	01111	10111
		GR 00001	00010	00100	10000	01000
2	GR	GR 11111				
3	GR	GR 11011	01111	11101	10111	11110
		OT 00100	10000	00010	01000	00001
4	OT	OT 01110	10110	11010	11001	01011
		GR 10001	01001	00101	00110	10100
5	OT	GR 11001	10011	10101	11100	01101
		OT 00110	01100	01010	00011	10010
6	OT	OT 11111				

▶5. Perform Final Classification Run

Figure 10. Processing flow for variable polygon partitioning. See text for description.

Figure 11. Three-color display for 13 May greenness vs. brightness scatterplot. Red represents irrigated grain, green non-irrigated grain, and blue non-grain.

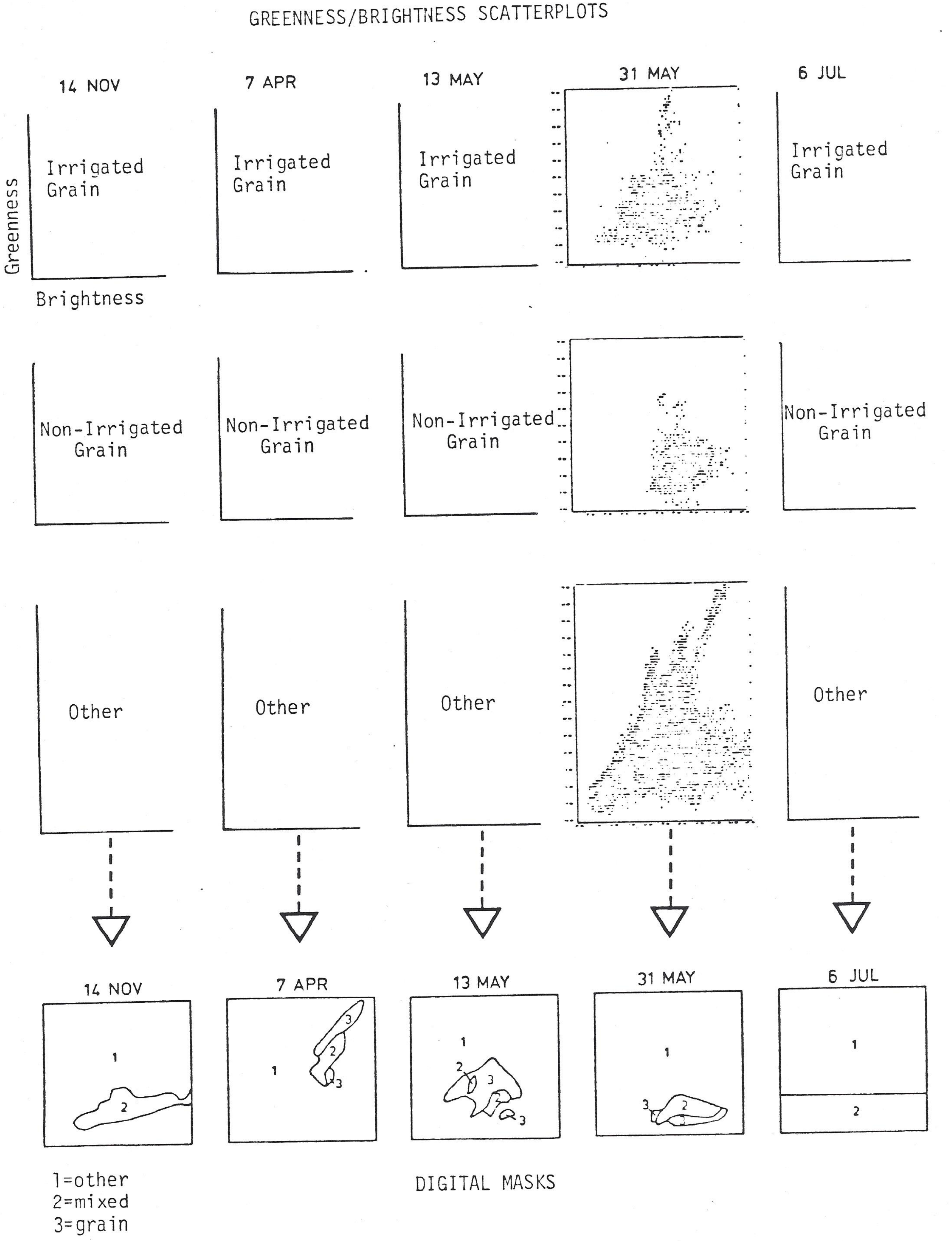


Figure 12. An example of variable polygon mask generation.

A fifth digital mask was added that corresponded to the July date, and was used as a final criterion for grain separation. This last polygon mask labelled all pixels above a greenness value of ten as "other", and all pixels below this discriminant as "grain". This threshold value was derived from an examination of the greenness values for all grain fields in the sample segments in July; all grain fields at that time contained either (a) stubble, (b) bare soil, or (c) emerging summer crops. Therefore, former grain fields should contain pixels with very low greenness values in July.

In order to simplify the vector classification process, both the irrigated and non-irrigated grain polygons were labelled as "grain", as were all the mixed polygons (corresponding to the areas of overlap of grain and non-grain pixels in greenness-brightness space). Therefore, all pixels on each date were considered by the computer to be either grain or non-grain. (Although nearly all non-grain pixels coincided with grain's greenness-brightness distribution on one or two dates, none overlapped on all dates, permitting separation in most cases.)

For each date, the computer used the appropriate polygon mask to sort all the pixels in the sample segments based on their relative polygon location in greenness-brightness space. This resulted in five classification bands (polygon assignment bands). The full set of five polygon assignment bands were then used by the computer to track all pixels in the data set through the multidate sequence (Figure 13). As a result, each pixel had a string of five class labels assigned to it, one label ("grain" or "other") per date. This string of labels, or vector, was calculated for each pixel in the data set, and put into a vector file where each unique vector was stored separately. A counter was incremented for every repeated vector (see Table 11).

The computer then grouped vectors according to the ratio of grain to non-grain labels in each vector. For example, vectors containing three grain labels (out of a possible five) were grouped into one grain class, while vectors with four grain labels were grouped into a second grain class. Finally, vectors with five grain labels were put into a third grain class. Conversely, there were also three "other" classes possible. Table 12 shows how vectors were grouped, based on label occurrence. (The automatic grouping of vectors into classes could have been done interactively, rather than accept the "winner take all" decision logic. This option will be discussed below in Section (b).

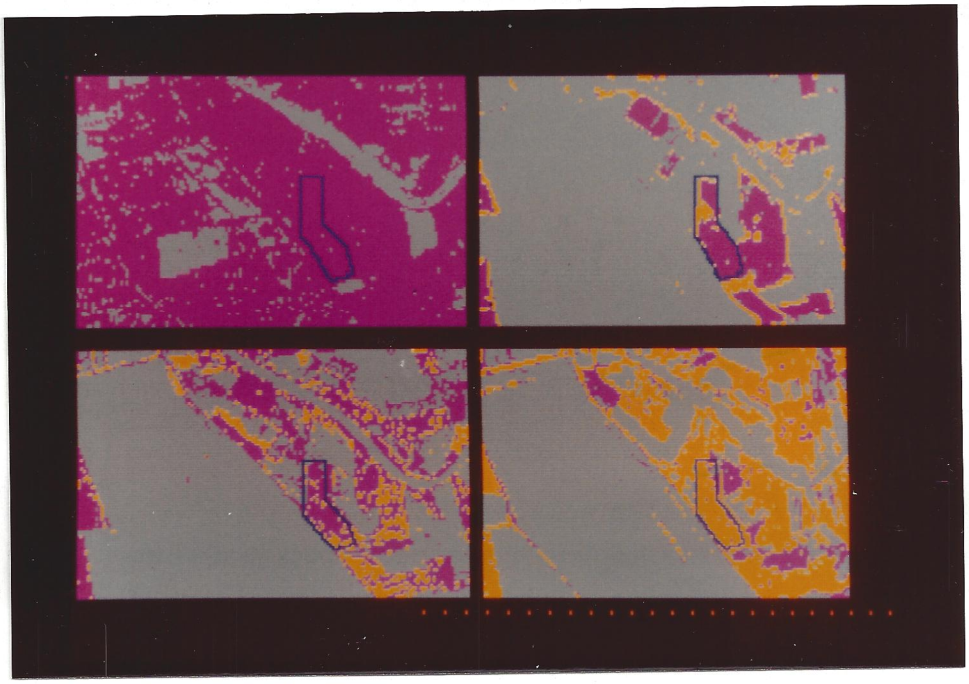


Figure 13. Sequence of polygon assignment bands used to assign a vector label to a sample field (outlined in blue). On each date, purple represents grain, yellow represents mixed grain/other, and grey represents other.

Table 11. Partial vector list. Vector labels for each date refer to the polygon location of the pixel on that date. The polygon masks have been included for illustrative purposes.

VECTOR NUMBER	COUNT	1 14 NOV	2 7 APR	3 13 MAY	4 31 MAY	5 6 JUL
1026	2122	2 GR	1 OT	1 OT	1 OT	1 OT
4766	1510	2 GR	2 GR	2 GR	3 GR	2 GR
1798	703	2 GR	3 GR	3 GR	3 GR	2 GR
1797	661	2 GR	2 GR	3 GR	3 GR	2 GR
4763	529	2 GR	3 GR	2 GR	3 GR	2 GR
4789	416	2 GR	1 OT	3 GR	3 GR	2 GR
1284	378	2 GR	1 OT	1 OT	1 OT	2 GR
770	317	1 OT	1 OT	1 OT	1 OT	1 OT
4797	309	2 GR	1 OT	1 OT	3 GR	1 OT
4774	294	2 GR	1 OT	2 GR	2 GR	2 GR
4778	284	2 GR	1 OT	3 GR	1 OT	1 OT
4792	277	2 GR	1 OT	2 GR	1 OT	1 OT
4767	236	2 GR	3 GR	2 GR	2 GR	2 GR
4798	226	2 GR	1 OT	1 OT	2 GR	2 GR
1286	224	2 GR	3 GR	1 OT	3 GR	2 GR
4758	216	1 OT	3 GR	1 OT	1 OT	1 OT
4744	215	2 GR	3 GR	1 OT	2 GR	2 GR
4800	211	2 GR	3 GR	1 OT	1 OT	2 GR
...	...	...	...	...	...	...
...	...	...	...	...	...	...

Table 12. Partial vector list, with vector labels and associated sequences.

VECTOR NUMBER	COUNT	MASK VALUE* 14 NOV	7 APR	13 MAY	31 MAY	6 JUL	VECTOR LABEL**	VECTOR SEQUENCES	
1026	2122	2	1	1	1	1	1	OT 01111	GR 10000
4766	1510	2	2	2	3	2	2	GR 11111	
1798	703	2	3	3	3	2	2	GR 11111	
1797	661	2	2	3	3	2	2	GR 11111	
4763	529	2	3	2	3	2	2	GR 11111	
4789	416	2	1	3	3	2	3	GR 10111	OT 01000
1282	378	2	1	1	1	2	4	OT 01110	GR 10001
4788	361	2	1	2	3	2	3	GR 10111	OT 01000
4790	342	2	2	2	2	2	2	GR 11111	
1284	326	2	1	1	3	2	5	GR 10011	OT 01100
770	317	1	1	1	1	1	6	OT 11111	
4797	309	2	1	1	3	1	4	OT 01101	GR 10010
4774	294	2	1	2	2	2	3	GR 10111	OT 01000
4778	284	2	1	3	1	1	4	OT 01011	GR 10100
4792	277	2	1	2	1	1	4	OT 01011	GR 10100
4767	236	2	3	2	2	2	2	GR 11111	
4798	226	2	1	1	2	2	5	GR 10011	OT 01100
1286	224	2	3	1	3	2	3	GR 11011	OT 00100
4758	216	1	3	1	1	1	1	OT 10111	GR 01000
4744	215	2	3	1	2	2	3	GR 11011	OT 00100
4800	211	2	3	1	1	2	5	GR 11001	OT 00110
4796	192	2	1	1	2	1	4	OT 01101	GR 10010
...									
...									
...									

* 1 = OT
2 = MX = GR
3 = GR

** 1 = Other
2 = Grain
3 = Grain
4 = Other
5 = Grain
6 = Other

As a final step, the machine then used its grouped vector classes to classify all the pixels in the data set. The class map was intersected with the ground truth labels in order to assess classification accuracy. The results are shown in Table 13a-c.

(b) Labelling Refinement

As can be seen in the results of the first classification run (Table 13a-c), there was a substantial overcommittment of non-grain pixels to grain. Based on the intersection of the class map with the ground data mask, the analyst was able to refine vector labelling in the following ways:

- An examination of the crop composition of grain class #5 indicated that it contained relatively few actual grain pixels. It was consequently decided that three grain labels in the five element vector sequence were insufficient to warrant assigning that vector to a grain class; there were simply too many non-grain pixels that fell into a grain (or mixed) polygon on <u>three</u> of the five dates.

- Based upon ancillary 1981 crop calendar information for the Sacramento Valley, the analyst made further refinements within grain class #3. Vectors that were assigned to that class contained four grain labels and one non-grain label. Normally, this would be an acceptable vector for grain, given the variability in emergence and turning (senescence) dates of grain. However, certain vectors were not acceptable grain vectors, and these were edited from grain class #3. These were vectors with a non-grain label for 7 April (when grain is "greennest" and most distinctive), or 31 May (which could signify a native vegetation vector). As previously mentioned, a non-grain label on 6 July was also unacceptable. On this basis, three vectors were deleted from the class (Table 14).

A second classification run was then performed on the edited file; the results are shown in Table 15a-c.

(c) Final Labelling Refinement

As a final step in the analysis, all individual vectors were intersected with the ground data in order to identify those vectors that contained mixtures of

Table 13A. Percentage of a given vector class occupied by each ground category using automatic vector labelling

	Vector Class					
Ground Class	1-OT	2-GR	3-GR	4-OT	5-GR	6-OT
Grain	3.1	71.3	35.2	11.9	19.2	6.0
Tomato	8.0	3.7	4.2	22.7	10.1	0.9
S.Beet	7.8	1.3	3.5	3.3	5.1	0.9
Corn	11.3	1.6	3.9	10.2	4.1	12.6
Native V.	0.5	14.4	24.3	5.8	21.3	0.3
Rice	47.1	0.1	0.3	3.5	1.2	46.4
Alfalfa	5.4	1.1	1.0	7.5	1.9	2.5
Misc.Orch.	7.5	1.3	5.5	16.0	12.9	21.5
Urban	1.9	1.4	5.4	4.7	7.4	0
Other	7.4	3.8	16.7	14.4	16.8	8.9
Class total	100.0	100.0	100.0	100.0	100.0	100.0
Percent total	20.0	28.6	18.8	15.6	14.7	2.2

Vector Class Grain Estimate = 62.1%

Ground Class Grain Estimate = 32.5%

Table 13B. Proportion of ground class assinged to each vector class

		Vector Class		
		Grain	Other	
Ground Class	Grain	.92	.08	1.0
	Other	.48	.52	1.0

Table 13C. Ground class composition within each vector class

		Vector Class	
		Grain	Other
Ground Class	Grain	.48	.07
	Other	.52	.93
		1.00	1.00

Table 14. Boxed vector sequences for grain class #3 indicate anomalous grain vectors that were deleted prior to the second classification run. The class label for class #5 was also changed to the more appropriate label, rather than accept the automatic label (GR).

COVER CLASS		SEQUENCES (PARTIAL LIST)					
1	OT	OT	11110	11101	11011	01111	10111
		GR	00001	00010	00100	10000	01000
2	GR	GR	11111				
3	GR	GR	11011	01111	$\boxed{11101}$	$\boxed{10111}$	$\boxed{11110}$
		OT	00100	10000	$\boxed{00010}$	$\boxed{01000}$	$\boxed{00001}$
4	OT	OT	01110	10110	11010	11001	01011
		GR	10001	01001	00101	00110	10100
5	OT	GR	11001	10011	10101	11100	01101
		OT	00110	01100	01010	00011	10010
6	OT	OT	11111				

Table 15A. Percentage of a given vector class occupied by each ground category after vector sequence editing

	Vector Class		
Ground Class	2-GR	3-GR	1,4,5,6-OT
Grain	71.3	44.7	10.0
Tomato	3.7	5.1	7.2
S.Beet	1.3	4.8	3.3
Corn	1.6	4.7	5.3
Native V.	14.4	23.0	7.6
Rice	0.1	0	11.2
Alfalfa	1.1	1.1	2.7
Misc.Orch.	1.3	4.2	7.2
Urban	1.4	1.5	3.2
Other	3.8	10.9	42.3
Class total	100.0	100.0	100.0
Percent total	28.6	4.7	66.7

Vector Class Grain Estimate = 33.3%

Ground Class Grain Estimate = 32.5%

Table 15B. Proportion of ground class assigned to each vector class

		Vector Class		
		Grain	Other	
Ground Class	Grain	.69	.31	1.00
	Other	.16	.84	1.00

Table 15C. Ground class composition within each vector class

		Vector Class	
		Grain	Other
Ground Class	Grain	.68	.15
	Other	.32	.85
		1.00	1.00

crop pixels differing with the label automatically assigned. Vectors were reassigned where appropriate. A third classification was then performed on the file, using the edited vectors. The results are shown in Table 16a-d. Figure 14 enables visual comparison of the vector classification with the JES ground labels for two of the sample segments.

(d) Results (Refer to Tables 13, 15, and 16)

Although the grain accuracy was 92 percent after the first automatic classification run, nearly half (48 percent) of the non-grain pixels had also been labelled grain. The reassignment of dubious grain vector sequences to non-grain vector labels, along with the reassignment of grain class #5 to non-grain, brought the grain accuracy down to 69 percent; however, the non-grain commission error also dropped from 48 percent to 16 percent. By the third run, in which mixed vectors were individually edited, grain classification accuracy rose to 72 percent, with a non-grain commission error of 16 percent.

Taken together, the grain classes were composed of 75 percent grain pixels by the third run, up from 48 percent on the initial classification run. The two grain classes represented 31.3 percent of the total area in the 17 sample segments. The corresponding ground truth measurement for these same 17 segments was 32.5 percent. Weighted overall classification accuracy was 82.5 percent.

d. Conclusions

The final version of the polygon vector classification procedure combined portions of both the Fixed and Variable Polygon Partitioning techniques. The first four dates of the Tasselled Cap-transformed MSS data were partitioned using the variable approach, since the grain and non-grain distributions for those dates contained a pattern of overlap not easily described by a set of fixed partitions. From the fixed approach, the final procedure borrowed the concept of a fixed threshold as the last criterion for grain separation. This threshold merely defined a greenness boundary below which all grain pixels should occur at the end of the growing season.

The results of the final polygon vector classification were generally encouraging, indicating the potential for characterizing grain (and other crop types) based on spectral greenness behavior throughout the season. Since the final run, some modifications have been suggested for further implementation. These include:

1. Improve segment registration to the Tasselled Cap data: minor shifts in segment locations should re-

Table 16A. Percentage of a given vector class occupied by each ground category after vector label editing

	Vector Class		
Ground Class	2-GR	3-GR	1,4,5,6-OT
Grain	77.4	55.9	9.0
Tomato	3.9	5.8	7.2
S.Beet	2.1	6.0	3.1
Corn	2.5	5.8	5.0
Native V.	6.1	10.7	10.8
Rice	0.3	0	11.2
Alfalfa	1.1	1.8	2.7
Misc.Orch.	1.0	3.9	7.4
Urban	1.4	1.2	3.2
Other	4.2	8.1	40.4
Class total	100.0	100.0	100.0
Percent total	27.7	3.6	68.7

Vector Class Grain Estimate = 31.3%

Ground Class Grain Estimate = 32.5%

Table 16B. Proportion of ground class assigned to each vector class

		Vector Class		
		Grain	Other	
Ground Class	Grain	.72	.28	1.00
	Other	.12	.88	1.00

Table 16C. Ground class composition within each vector class

		Vector Class	
		Grain	Other
Ground Class	Grain	.75	.13
	Other	.25	.87
		1.00	1.00

Table 16D. Weighted ground class accuracy

		% correct	% area	% area correct
Ground Class	Grain	72.0	32.5	23.4
	Other	88.0	67.5	59.4
				82.8

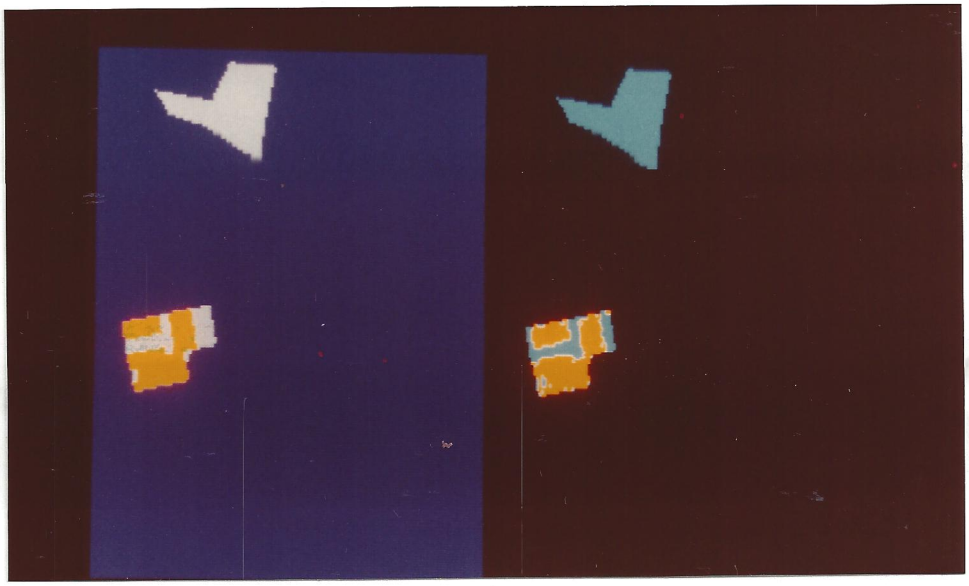


Figure 14. Comparison between ground truth (left) and polygon vector classification for two segments. Blue represents non-grain fields, while gold denotes grain.

duce some of the errors introduced by adjoining fields, levees, roads, etc.

2. Increase training sample size: Supplemental training segments will be added in those strata in which there were few JES segments.

3. Modify the manner in which the scatterplots are displayed for polygon partitioning: improvement will be made in the consistency with which count density levels are displayed across training categories.

Once these changes have been made to the data set and procedure, the polygon vector classification procedure will be tested on the Yolo County 1981 data set and the 1982 data set associated with Landsat scene 48-33.

3. Small Grains Classification Technique Development Using the 1976 Sacramento Valley Data Set

This effort was carried out in conjunction with the initial polygon vector classification work for multicrop application. Results for small grains are summarized in that section.

B. Development of Procedures for Classification of Multiple Crops

1. Formulation of a Conceptual Framework for Multiple Crop Classification in California

Review of previous Landsat-related agricultural work in California suggested that achievement of long-range Cooperative Study objectives would likely require new or augmented classification procedures. Furthermore, these new procedures would probably require stepwise decision logic, possibly similar to current analyst-oriented approaches. Decisions at each step in such a classification process could be limited to information, and based upon classification rules, specific to the decision problem at hand. This approach was expected to potentially

(a) simplify individual classification problems,

(b) improve classification accuracy, particularly for crops of medium to low frequency, and

(c) in the long run, minimize cost to achieve fixed performance levels.

Figure 15 shows a current view of this classification framework. Its key feature is a classification level structure corresponding to successively more detailed information based on successively more difficult decisions. This framework should allow the generation of products of the kind suggested by

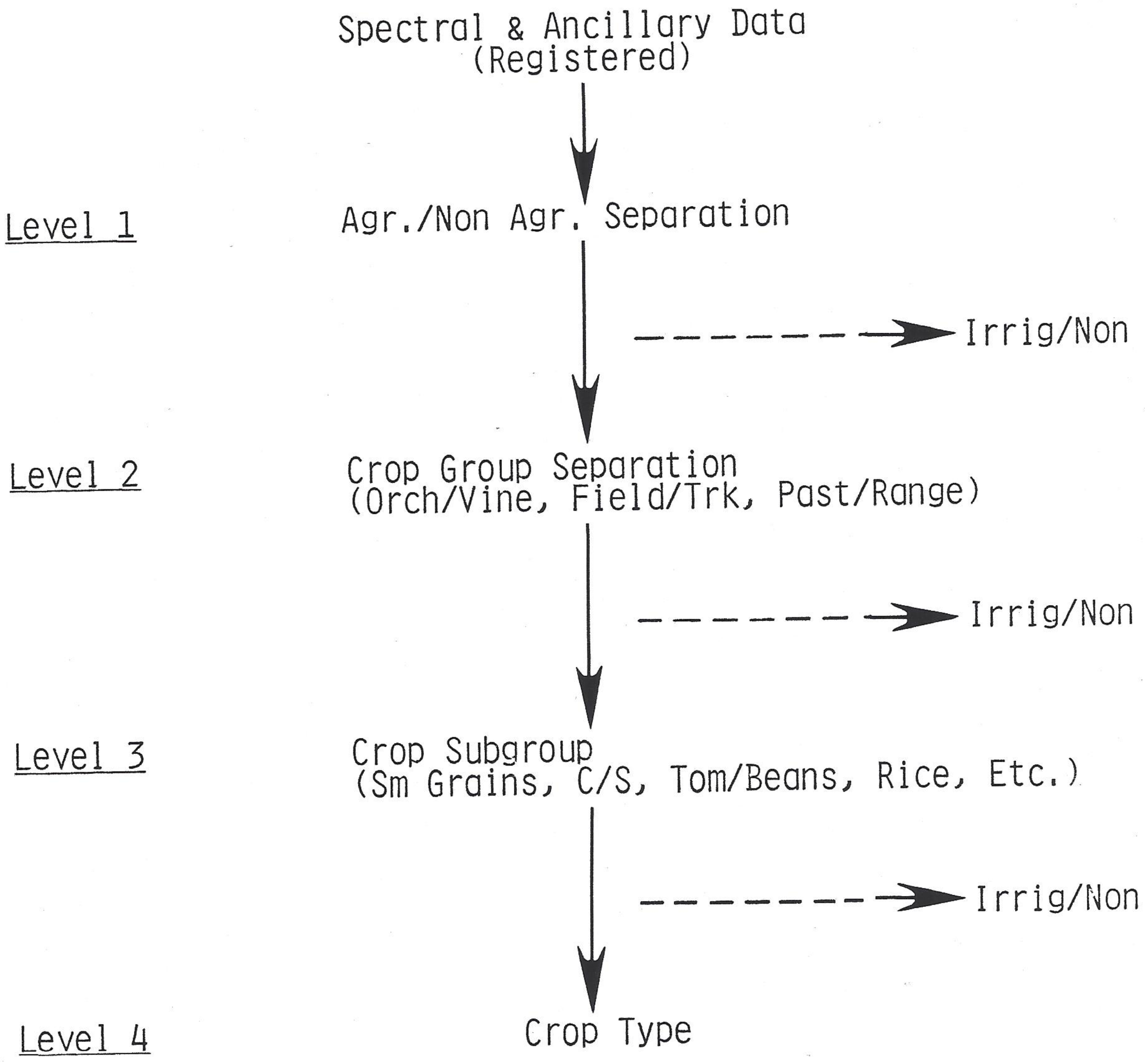


Figure 15. Current View of Classification Frame

Bosecker (1982) in Figure 16, and should enhance the ability to identify and clearly define research and implementation issues at each classification level and for each estimation/mapping objective during the process of development.

Figure 17 shows a candidate classification flow using this approach. Note that such a framework enables modular substitution of alternative classification procedures at each level. In the example of Figure 17, a number of alternatives are shown at each level. These range from polygon classification related procedures at the top levels to conventional and profile classification at the middle to lower levels. Together they form a "tool box" of Boolean and statistical decision procedures that can be drawn out for test as progress in the development process dictates.

2. Classification Technique Development Using the 1976 Sacramento Valley Data Set

Initial classification technique development focused on polygon classification. It was felt that the biggest short- and midterm payoffs in the development of a working classification framework would come by emphasizing work on these kinds of techniques during the early phases of the Cooperative effort.

a. The Data Set

Since 1982 data was not to be available for development purposes before the first or second quarter of 1983, a decision was made to use the 1976 Sacramento Valley data set for this purpose. This data set, obtained in support of the California Irrigated Lands Project, contained a very good through-the-season Landsat acquisition history as well as area-wide ground data. In particular, the 1976 Sacramento Valley data set consisted of five dates of Landsat MSS data including 3 May, 30 May, 26 June, 28 August, and 4 October, 1976. The study area covered an area 30 minutes of latitude by 30 minutes of longitude (representing sixteen 7-1/2 minute quadrangles) and contained several agricultural strata (see Figure 18). Wall-to-wall ground data had been collected by the California Department of Water Resources for the entire study area.

b. Data Preparation

All MSS bands were registered to a north-south ground coordinate system based on the 7-1/2 minute quadrangle maps, with each pixel being approximately 0.50 hectare. The DWR ground data maps were digitized and registered to this map base. A greenness indicator, the ratio of MSS 7 to MSS 5, was computed for each pixel, on each date. Two-dimensional scatterplots were produced by plotting MSS 7/5 ratios for Date i versus Date i + 1 utilizing all pixels with the

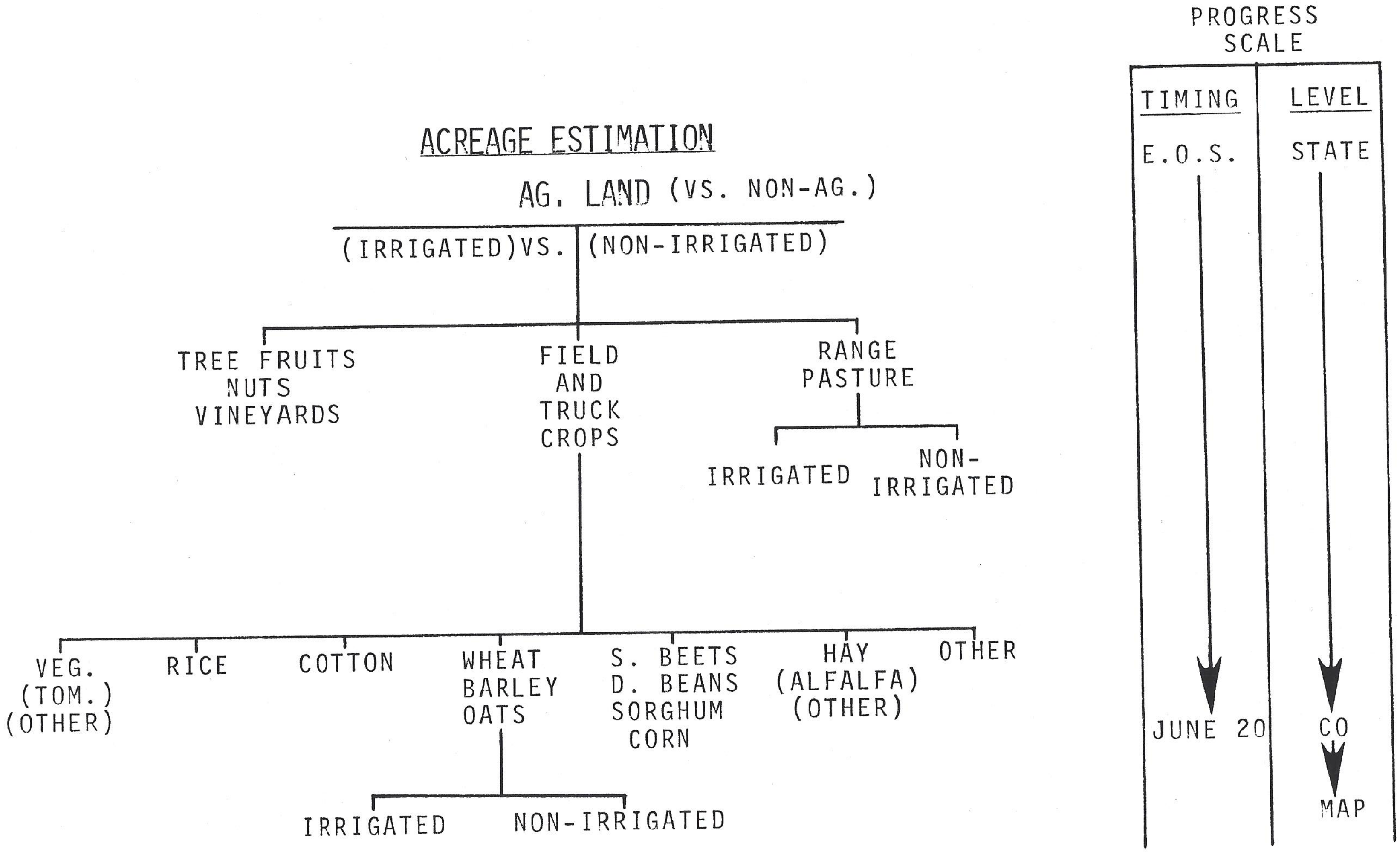


Figure 16. A Suggested Hierarchy of Remote Sensing Activity
(Reproduced with permission of Ron Bosecker, California SSO)

Candidate Classification Flow

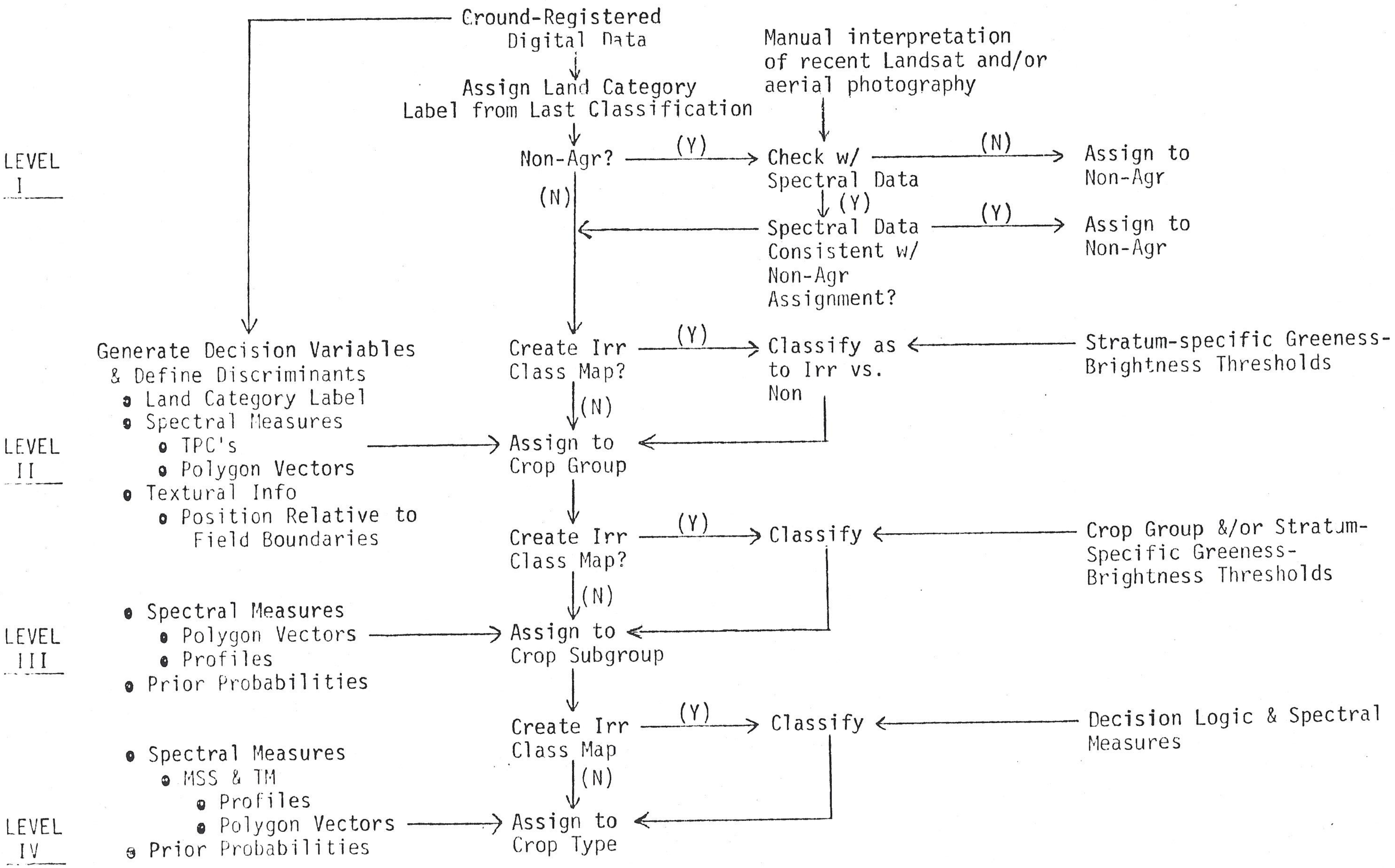


Figure 17. Candidate Classification Flow

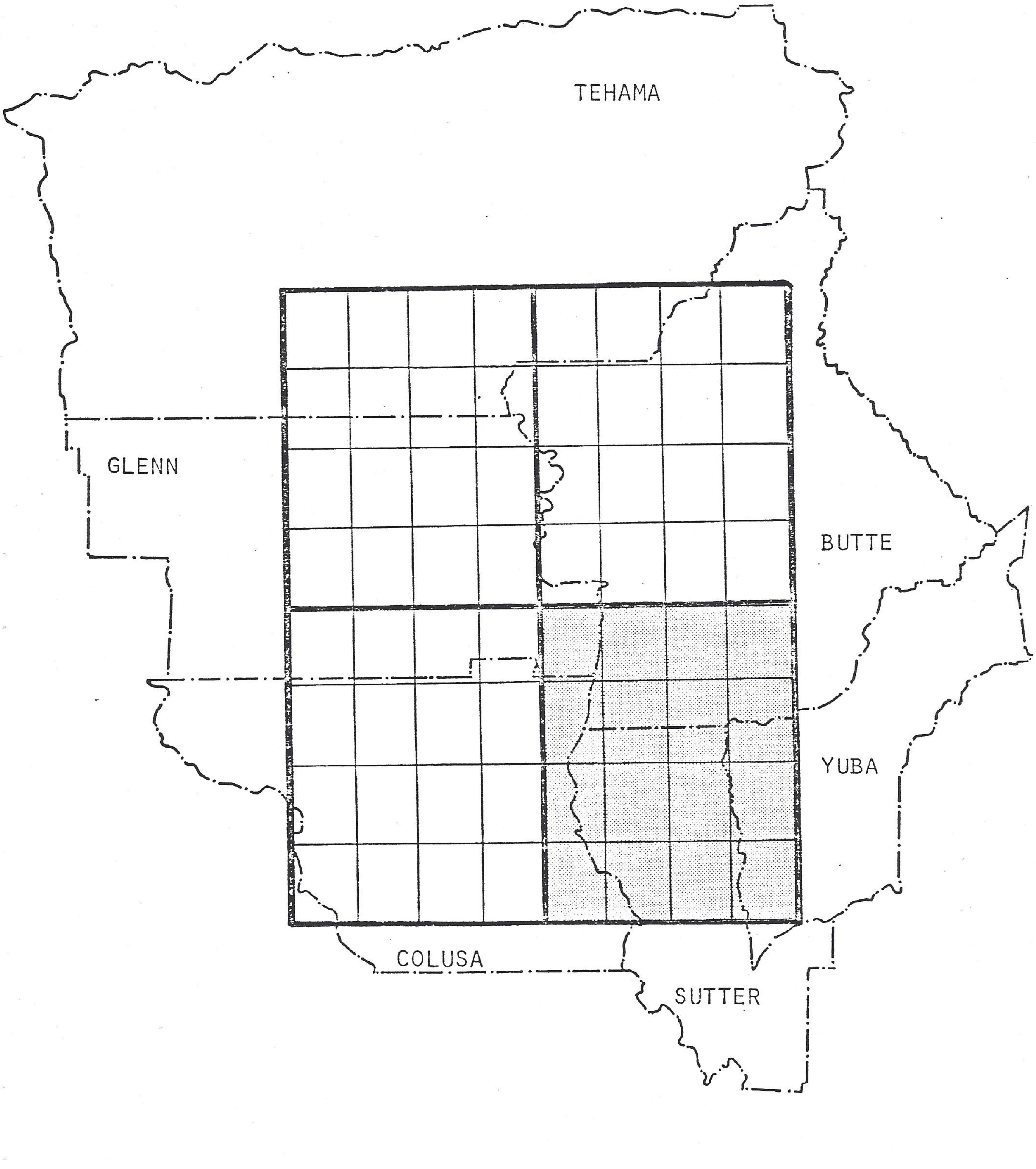


Figure 18. Sacramento Valley Study Area. Statistics were obtained for 4 30' blocks; classification was performed on the shaded block.

study area. Scatterplots were generated in this manner for the following date combinations: 3 May versus 30 May, 30 May versus 26 June, 26 June versus 28 August, and 28 August versus 4 October. Also, using the ground data as a mask, separate scatterplots were produced for grain, rice, and tomatoes. These crops were selected for the initial test because they are important crops in the Sacramento Valley and are of importance to both USDA and DWR.

c. Scatterplot Partitioning

1) Fixed Polygon Partitioning

Under the fixed polygon approach, each scatterplot was partitioned by first locating the soil line for the MSS 7/5 ratio. This "line" represented the MSS 7/5 ratio value above which green vegetation becomes spectrally detectable. A 7/5 threshold value of 1.10 was used as the soil line for each date (Hay, Thomas, et al 1977). In addition, a partition was placed along the arm of equal 7/5 ratio value (the line of equal greenness) to discriminate between vegetation increasing in greenness and vegetation decreasing in greenness. This partitioning resulted in five polygons per date as shown previously in Figure 7.

These polygon boundaries were placed over the scatterplots for each crop type on each date combination. Labels were then assigned by the analyst to each scatterplot polygon based on the presence or absence of each of the three crops of interest (small grains, rice, or tomatoes) or the presence of "other". For example, on the first date, polygon 1 contained tomatoes, rice, and grain; polygon 2 contained tomatoes and rice; polygon 3 contained other; polygon 4 contained other; and polygon 5 contained tomatoes and grain.

A map was produced for each date combination showing the polygon assignment for each pixel within the test site (see Figure 19). Polygon vectors representing unique sequences of polygon assignments over the four scatterplots were formed and all pixels having the same sequence were then assigned the same vector label. Thus, for example, all pixels classified to polygons 1, 1, 2, 5 (Table 17) over the four date combinations would be grouped together. A new map was produced showing these polygon vector assignments (see Figure 20).

The vector classification was labelled automatically by the computer using the crop occurrence information shown in Table 17. Crop occurrence was coded for each unique sequence of polygons as a string of ones and zeroes denoting, respectively, the presence or absence of a given crop in each polygon. For example, if grain,

Figure 19. Fixed polygons are shown in upper right. Polygon Assignment Bands for 3 May vs 30 May, 30 May vs 26 June, 26 June vs 28 August, counter-clockwise from upper left.

Table 17: A partial list of polygon vectors, associated crop occurrence in each polygon (denoted as presence ("1") or absence ("0") for each crop type) and vector classes automatically assigned based on crop occurrence.

VECTOR NUMBER	POLYGON VALUE FOR SCATTERPLOT				CLASS NUMBER	CROP OCCURRENCES IN POLYGONS
	1	2	3	4		
1	1	1	1	1	1	TO 1111 GR 1111 RI 1101
2	1	1	1	2	2	GR 1111 TO 1110 RI 1100
3	1	1	2	3	3	TO 1110 RI 1101 GR 1100 NO 0001
4	1	1	2	4	4	TO 1110 RI 1101 GR 1101
5	1	1	2	5	5	TO 1111 RI 1101 GR 1101
6	2	5	1	2	6	TO 1110 GR 0111 RI 1000
7	1	2	3	3	4	TO 1110 RI 1110 GR 1100 NO 0001
8	1	2	3	4	7	RI 1111 TO 1110 GR 1100
9	1	2	3	5	8	TO 1111 RI 1111 GR 1101
10	2	5	2	3	3	TO 1110 NO 0001 RI 1000 GR 0100
11	2	5	2	4	3	TO 1110 RI 1001 GR 0100
12	2	5	2	5	5	TO 1111 RI 1001 GR 0101
13	1	2	4	3	4	TO 1110 RI 1110 GR 1100 NO 0001
14	4	5	1	1	6	TO 0111 GR 0111 NO 1000 RI 0001
15	2	3	3	4	9	RI 1011 TO 1010 GR 0100

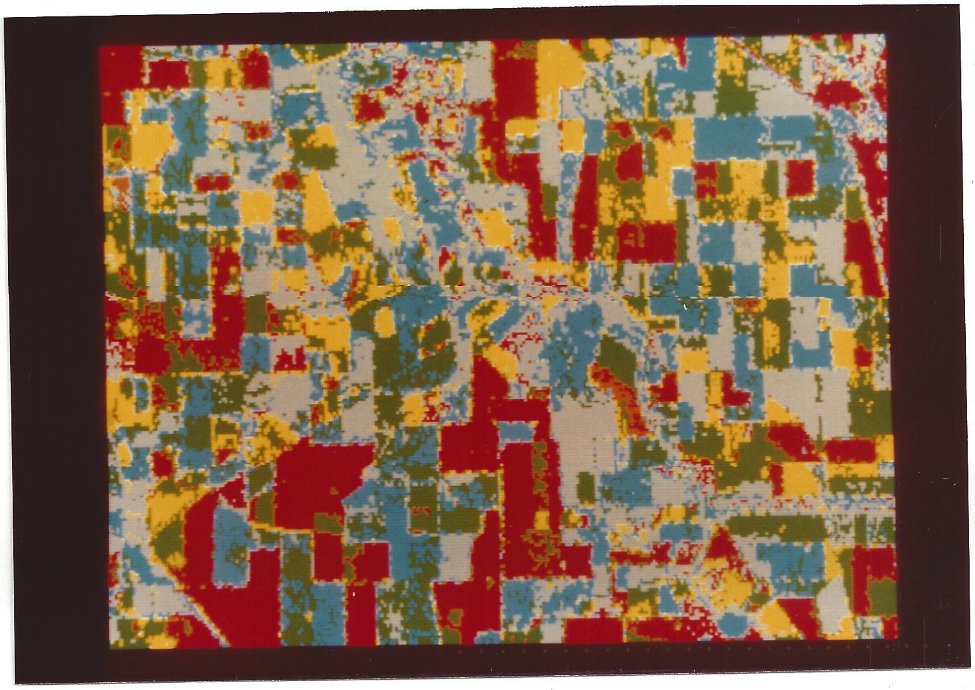


Figure 20. Classification using fixed polygons. Red is rice, blue is small grains, yellow is tomatoes, green is mixed, gray is other.

rice, and tomatoes all occurred in polygon 1 on the first scatterplot, then a "1" was placed in the first position of the zero-one string for each crop type for that vector. Each of the resulting 22 unique polygon vectors was given a label based on the number of occurrences of each cover type as shown in Table 18.

The performance of the fixed polygon approach was disappointing. Few polygon vectors represented single ground categories (see Table 19). The fixed polygons, as defined, did not adequately capture the actual temporal behavior of the crops of interest.

Comparing the crop makeup of the polygon vectors shown in Table 19 with the automatic labels shown in Table 18, it can be seen that the automatic labelling procedure did not account for the actual crop mix within the vectors. Based on the intersection with the registered ground data, the classes were then relabelled as illustrated in Table 20.

The 22 vectors were grouped into four new classes based on the user labels: small grains, rice, tomato, and other. The mixed vectors were grouped with the crop that predominated. Some vectors were relabelled after examination of the class map. These vectors contained a plurality of "other" but the pixels in that class fell along roads and field boundaries. They were relabelled to rice, grain, or tomatoes if the remaining pixels were predominantly that crop type. Resulting classification accuracies associated with this grouping are shown in Tables 22 and 23. These figures were derived from the pixel counts given in Table 21.

Although the proportion of grain and rice classified correctly by the polygon classification was high (.90 and .92, respectively), the class purities shown in Table 23 were lower than desired. Examination of the scatterplots and the polygon boundaries suggested that the fixed polygon approach did not account for the actual behavior of the data, and that a customized boundary setting procedure would probably be required to improve classifier performance.

2) Custom Polygon Partitioning

The same dates of Landsat MSS 7/5 ratio and the same scatterplots by date combination were used in the development and test of a custom polygon classification procedure, where each crop of interest was classified separately. As of January 1983, testing of this classification approach had been limited to small grains and rice.

Table 18. Automatic polygon vector labelling for the fixed polygon classification procedure.

VECTOR CLASS	AUTOMATIC LABEL
1	tomato, small grains
2	small grains
3	tomato
4	tomato, rice
5	tomato
6	tomato, small grains
7	rice
8	tomato, rice
9	rice
10	tomato, rice, small grains
11	tomato, rice
12	other, tomato
13	tomato
14	tomato, rice, small grains
15	small grains
16	other
17	rice
18	tomato, rice, small grains
19	other, tomato, rice, small grains
20	other, rice, small grains
21	other, rice
22	other

TABLE 19. The percentage of a given vector class occupied by each ground category using automatic vector labelling.

VECTOR CLASS	% GRAIN	% RICE	% TOMATO	% OTHER
1	49	3	1	47
2	10	12	1	77
3	4	10	1	85
4	1	40	3	56
5	6	6	14	74
6	81	2	1	16
7	0	88	0	11
8	1	36	23	40
9	0	90	0	10
10	3	6	44	47
11	1	39	1	60
12	6	2	0	92
13	2	3	1	94
14	3	2	52	43
15	52	3	2	43
16	1	2	0	97
17	1	6	0	93
18	8	7	1	84
19	20	5	1	74
20	1	14	2	83
21	2	3	1	94
22	2	2	0	96

Table 20. Automated and user labels for polygon vectors.

VECTOR CLASS	AUTOMATIC LABEL	USER LABEL
1	tomato, grain	grain mix
2	grain	other
3	tomato	rice mix
4	tomato, rice	rice mix
5	tomato	other
6	tomato, grain	grain
7	rice	rice
8	tomato, rice	rice mix
9	rice	rice
10	tomato, rice, grain	tomato mix
11	tomato, rice	other
12	other, tomato	other
13	tomato	other
14	tomato, rice, grain	tomato
15	grain	grain mix
16	other	other
17	rice	other
18	tomato, rice, grain	other
19	tomato, rice, grain, other	other
20	other, rice, grain	other
21	other, rice	other
22	other	other

Table 21. Pixel counts for polygon classes.

		VECTOR CLASS		
GROUND CLASS	GRAIN	RICE	TOMATO	OTHER
grain	45133	1239	291	3621
rice	2379	97640	409	5678
tomato	1271	5258	4700	2538
other	34727	46595	4382	100007

Table 22. Proportion of ground class assigned to each vector class.

		VECTOR CLASS			
GROUND CLASS	GRAIN	RICE	TOMATO	OTHER	
grain	.90	.02	.01	.07	1.00
rice	.02	.92	.01	.05	1.00
tomato	.09	.38	.34	.19	1.00
other	.19	.25	.02	.54	1.00

Table 23. Ground class composition within each vector class.

		VECTOR CLASS		
GROUND CLASS	GRAIN	RICE	TOMATO	OTHER
grain	.54	.01	.03	.03
rice	.03	.65	.04	.05
tomato	.01	.03	.48	.02
other	.42	.31	.45	.90
	1.00	1.00	1.00	1.00

To define the polygon boundary for small grains, the grains-only scatterplot for the first date combination was displayed on a color monitor and a boundary was drawn around the distribution of grain pixels using an interactive cursor. The area outside the polygon was labelled "not grain" and was given a mask value "1" and the area inside the polygon was labelled "grain" and given a mask value "2". These classes were written to a mask band in the data file. Then the rice, tomato, and "other" scatterplots for the same date were displayed. Using the small grains mask band just created, all scatterplot points labelled "not grain" were masked out. This action left a display of only non-grain scatterplot points occupying the same partition as grain. These remaining points were outlined labelled "mixed" and given a mask value "3". This process was performed for each of the date combinations, resulting in a mask band showing areas of grain, not grain, and mixed for each scatterplot (see Figures 21 and 22).

As in the fixed polygon approach, each pixel in the study area was assigned to the appropriate polygon for each date combination. The resulting sequence of polygon assignments (the polygon "vector") was obtained for each pixel. Vectors were then grouped into classes according to the sequence of polygon labels through the four scatterplots. Pure vectors (all four polygon assignments having the same crop or cover type) were automatically assigned to that label. Mixed vectors (vectors having a mix of labels over the four scatterplots) were labelled according to their dominant ground class composition. The ground class composition was determined by intersection of the polygon classification with the digitized ground data (see Figures 23 and 24).

The entire procedure was repeated to classify rice by first displaying the rice scatterplot, then scatterplots for the other cover types, and noting areas of intersection. Polygon numbering proceeded as described previously for grains, producing a scatterplot mask of rice, not rice, and mixed for each date.

3) Results

The results for the custom polygon approach are shown in Tables 24 and 25. The accuracy figures and the class purity were encouraging. Within-ground class accuracy was 61 percent for grain and 78 percent for rice. The proportion of "other" classified incorrectly (committed) to grain and rice was three

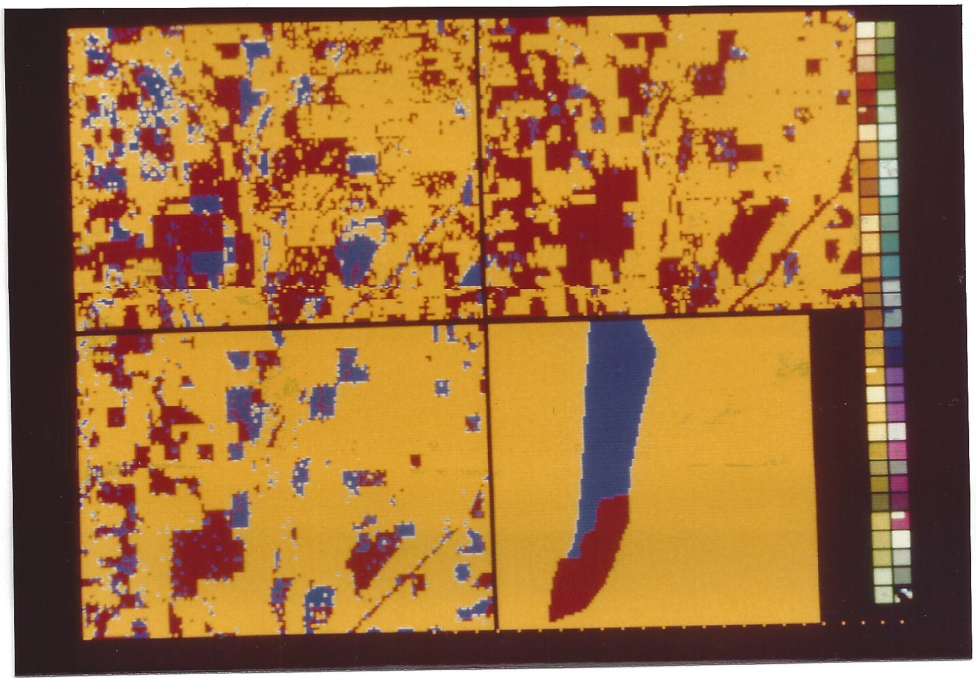


Figure 21. Mask for small grains for 3 May vs. 30 May in lower right - blue = grain, brown = mixed, yellow = other. Polygon assignments for 30 May vs. 26 June, 3 May vs. 30 May, and 26 June vs. 28 August shown counter-clockwise from upper right.

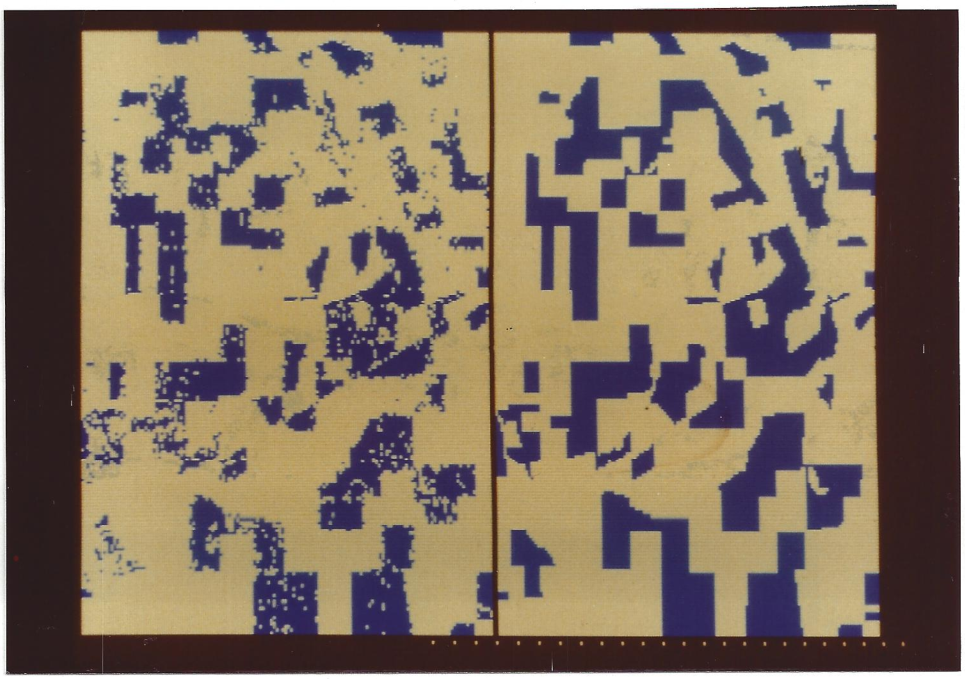


Figure 22. Classification for grain shown on left, digitized ground data on right. Grain is blue, other is gray.

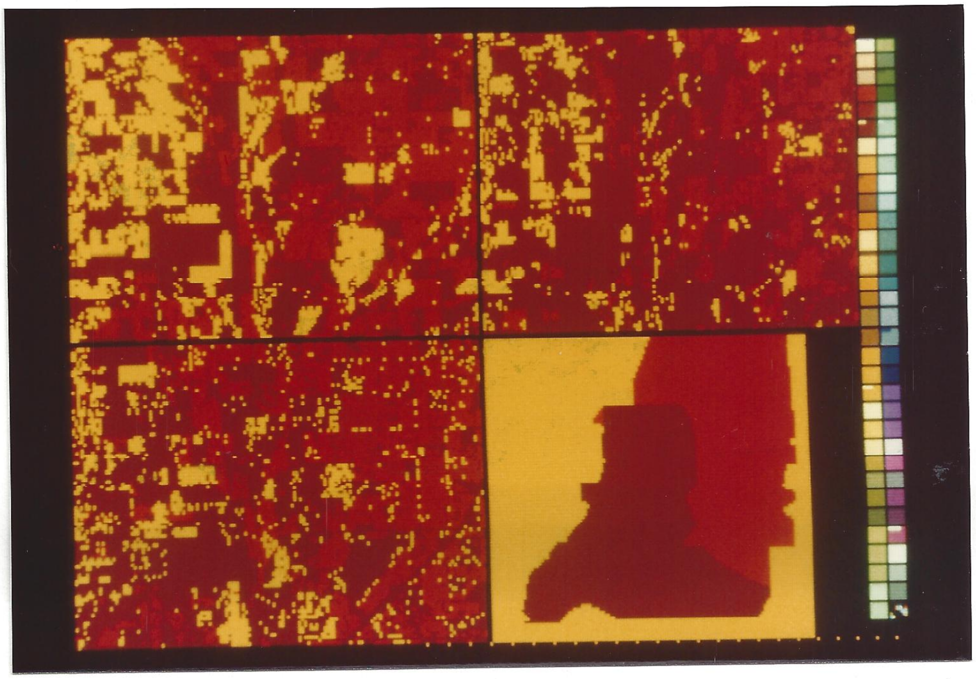


Figure 23. Mask for rice for 26 June vs. 28 August shown in lower right; rice = red, mixed = brown, other = yellow. Polygon assignment bands for 30 May vs. 26 June, 3 May vs. 30 May, and 26 June vs. 28 August shown counter-clockwise from upper right.

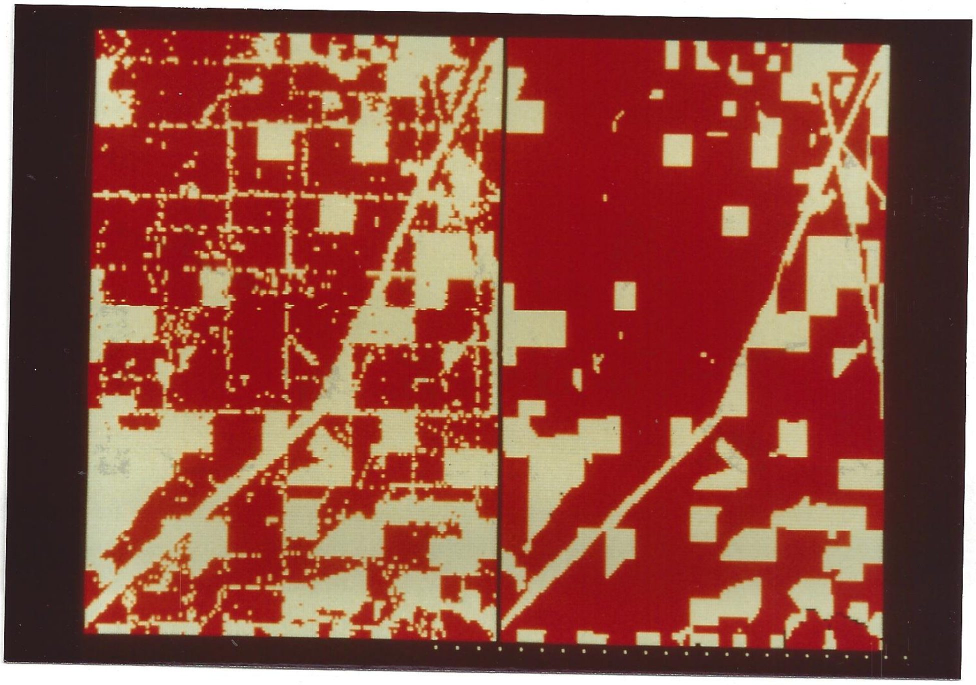


Figure 24. Classification for rice shown on left; digitized ground data on right. Rice is red, other is gray.

Table 24. Results for small grains classification.

A. PIXEL COUNTS FOR CLASSIFICATION

	VECTOR CLASS			
GROUND CLASS	GRAIN	NOT GRAIN	MIXED	TOTAL
Grain	30491	13661	6132	50284
Not Grain	6974	149622	73488	230984
Total	37465	163283	80520	281268

B. GROUND CLASS COMPOSITION WITHIN EACH VECTOR CLASS

	VECTOR CLASS		
GROUND CLASS	GRAIN	MIXED	NOT GRAIN
Grain	.81	.08	.08
Not Grain	.19	.92	.92
Total	1.00	1.00	1.00

C. GROUND CLASS COMPOSITION WITHIN EACH VECTOR CLASS WHEN "MIXED" IS CALLED "NOT GRAIN"

	VECTOR CLASS	
GROUND CLASS	GRAIN	NOT GRAIN
Grain	.81	.08
Not Grain	.19	.92
Total	1.00	1.00

D. PROPORTION OF GROUND CLASS ASSIGNED TO EACH VECTOR CLASS

	VECTOR CLASS		
	GRAIN	NOT GRAIN	
Grain	.61	.39	1.00
Not Grain	.03	.97	1.00

E. WEIGHTED ACCURACY

	% CORRECT	% OF AREA	% OF AREA CORRECT
Grain	61	18	10.98
Not Grain	97	82	79.54

WEIGHTED ACCURACY = 90.52

Table 25. Classification accuracies for rice.

A. PIXEL COUNTS FOR CLASSIFICATION

		VECTOR CLASS		
	RICE	NOT RICE	MIXED	TOTAL
Rice	65536	23867	16703	106106
Not Rice	0	170154	5008	175162
Total	65536	194021	21711	281268

B. GROUND CLASS COMPOSITION WITHIN EACH VECTOR CLASS

		VECTOR CLASS	
	RICE	NOT RICE	MIXED
Rice	1.00	.12	.77
Not Rice	.00	.88	.23
Total	1.00	1.00	1.00

C. GROUND CLASS COMPOSITION WHEN "MIXED" CALLED "RICE"

	VECTOR CLASS	
	RICE	NOT RICE
Rice	.94	.12
Not Rice	.06	.88
Total	1.00	1.00

D. PROPORTION OF GROUND CLASS ASSIGNED TO EACH VECTOR CLASS

	VECTOR CLASS		
	RICE	NOT RICE	
Rice	.78	.22	1.00
Not Rice	.03	.97	1.00

E. WEIGHTED ACCURACY

	% CORRECT	% OF AREA	% OF AREA CORRECT
Rice	78	38	29.6
Not Rice	97	62	60.1
		WEIGHTED ACCURACY =	89.7

percent in both cases. Ground class composition of the grain polygon vector class was 81 percent grain and 19 percent non-grain; the non-grain class was 92 percent non-grain and 8 percent grain. The rice polygon vector class was composed of 94 percent rice and 6 percent non-rice, while the non-rice vector included 88 percent non-rice and 12 percent rice. Overall, the area-weighted average classification average of grain and non-grain was 90.5 percent, and of rice and non-rice was 89.7 percent.

Although a large number of mixed pixels were still present in both the grain and rice classification, examination of the class map showed that a significant proportion of these occurred along roads and field boundaries.* Unfortunately, the ground data did not include all the field boundaries. The traditional method of tabulation for DWR has been to cut and weigh to determine crop or land use acreage, so boundaries between fields of the same crop or cover type had been eliminated to facilitate cutting.

---

*The impact on the results reported here was undoubtably to lower grain and rice class accuracies. Improved ground file field definition will be sought in future test and development activity.

C. SUMMARY

Development of classification techniques based on temporal sequences of scatterplot partitions was initiated this past summer and fall. Initial results employing pre-specified polygon partitions were poor in both the small grains-only and three crop examples. Subsequent work emphasizing training data-dependent scatterplot partitioning showed improved, though as yet not satisfactory, classification performance.

V. ANTICIPATED 1983 WORK

A. Development of Procedures for Classification of Small Grains

Work on the development of polygon classification techniques will continue through the first half of 1983. Beginning this summer, a test is planned on the 1981 Yolo County data set, followed by a test over Landsat frame 48-33 on the 1982 data set. These tests will be used to compare and cull in stepwise fashion small grains techniques developed by Cooperative Study participants. Evaluation criteria will include relative sampling efficiency, class map accuracy, relative cost and time requirements, relative difficulty, data requirements, and anticipated repeatibility of performance.

On the basis of results forthcoming from these tests, recommended procedures for early and end-of-season classification and estimation of small grains will be developed by the Cooperative Study participants. Examples of results and products will be prepared for presentation to DWR and USDA.

B. Development of Procedures for Classification of Multiple Crops

The proposed overall pattern of development and test work parallels that described for small grains. However, the schedule will lag that of small grains, and the number of development issues will undoubtably be larger. It is anticipated that polygon vector, ground truth masking, and decision logic development will receive the greatest emphasis during 1983. The latter will entail the development of an initial classification procedure flow, incorporating candidate procedures as necessary to obtain required land use separtions.

LITERATURE CITED

Beck, L. H. 1982. Crop separation information for Landsat pass selection. University of California at Berkeley Cooperative Study document, 10 September. 28 pp.

Bosecker, R. 1982. Personal communication.

Cicone, R., C. Hay, R. Horvath, M. Metzler, O. Mykolenko, J. Odenweller, and D. Rice. 1981. Users manual for the U.S. baseline corn and soybean segment classification procedure. AgRISTARS procedures report FC-E1-00712, Environmental Research Institute of Michigan, Ann Arbor, Michigan, and Space Sciences Laboratory, University of California, Berkeley, California.

Hay, C. M., R. W. Thomas, L. H. Beck, A. S. Benson, C. E. Brown, J. C. Claydon, J. B. Odenweller, E. J. Sheffner, and G. Smith. 1977. Development of techniques for producing static strata maps and development of photointerpretive methods based on multitemporal Landsat data. Annual Report. NASA Contract NAS9-14565, Principal Investigator, R. N. Colwell. University of California at Berkeley. Space Sciences Laboratory, Series 19, Issue 1. December.

Hay, C. M., C. A. Kuretz, J. B. Odenweller, E. J. Sheffner, B. L. Wood. 1979. Development of AI procedures for dealing with the effects of episodal events on crop temporal spectral response; and development of AI guidelines for corn and soybean labeling. Annual Report. NASA Contract NAS9-14565, Principal Investigator, R. N. Colwell. University of California at Berkeley, Space Sciences Laboratory. November.

Hall, R. G. 1982. Satellite remote sensing: an integral tool in acquiring global crop production information. In: Proceedings of the Eighth International Symposium on Machine Processing of Remotely Sensed Data. LARS-Purdue University, West Lafayette, Indiana, July. pp. 10-22.

Kauth, R. J. and G. S. Thomas. 1976. The Tasselled Cap - a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. In: Proceedings of the 1976 Symposium on Machine Processing of Remotely Sensed Data. LARS-Purdue University, West Lafayette, Indiana, June 29 - July 1. pp. 4B-41 to 4B-51.

Kauth, R. J., P. F. Lambeck, W. Richardson, G. S. Thomas, and A. P. Pentland. 1979. Feature extraction applied to agricultural crops as seen by Landsat. In: Proceedings of Technical Sessions, Volume II, The LACIE Symposium. NASA-Johnson Space Center Document JSC-16015, Houston, Texas, July. pp. 705-721.

Odenweller, J. B., C. M. Hay, and B. L. Wood. 1983. Development, implementation, and evaluation of methods incorporating expertise in analysis of remotely sensed and other data. Final Report. NASA Contract NAS9-16413, Principal Investigator, R. N. Colwell. University of California at Berkeley, Space Sciences Laboratory. (In Press).

Rice, D. P., E. P. Crist, and W. A. Malila. 1980. Applicability of selected wheat remote sensing technology to corn and soybeans. Final Report. NASA Contract No. NAS9-15082. Environment Research Institute of Michigan, Ann Arbor, Michigan, Report No. 124000-9-F, January. 41 pp.

Wall, S. L., R. W. Thomas, C. E. Brown, J. D. Baggett, E. H. Bauer, M. Eriksson, L. R. Tinney, G. B. Sawyer, C. L. Ferchaud, and C. A. Grigg. 1981. Irrigated lands assessment for water management applications pilot test. Annual Report. NASA Cooperative Agreement NCC 2-54, Principal Investigator, R. N. Colwell; Co-Investigator J. E. Estes. University of California, Space Sciences Laboratory, Series 22, Issue 23. 333 pp.

APPENDIX I: PRELIMINARY* DEFINITION OF EXPERIMENTS DESIGNED TO SUPPORT ACHIEVEMENT OF THE OVERALL COOPERATIVE LANDSAT STUDY OBJECTIVES

I. Experiment #1a: Small Grains Classification Technique Development

A. Objective:

Develop a digital technique for mapping small grains analogous to the manual procedure developed in 1981.

B. Purpose:

Provide an accurate method of creating a small grains stratum for

1. direct use by DWR in their land use mapping program
2. ratio estimation of wheat, barley, and oats acreage by USDA
3. later use as a prestratification tool in direct classification of wheat, barley, and oats for use by both USDA and DWR
4. use as a prestratification tool in mapping irrigated/non-irrgated grains as a category, and later individually by crop type
5. possible use as a small grains flag for JES enumerators and/or as a year-specific sample or count unit stratification tool

C. Approach:

1. Select one Landsat frame for 1981 containing a variety of small grains situations within counties included in the 1981 DWR-APT small grains task
2. Register 3 to 4 dates for this frame
3. Develop date-specific classification procedure for small grains
4. Develop Boolean procedure for combining date-specific information to generate a small grains classification
5. Compare resulting digital class map accuracy, cost, time, and flexibility with corresponding manual procedure

II. Experiment #1b: Large Area Test of Small Grains Classification and Area Estimation

A. Objectives:

1. Develop county, regional, and statewide procedures for small grains, wheat, barley, and oats acreage estimation using the Landsat small grains stratification technique from Experiment #1a
2. Do likewise for estimating acreage of small grains, wheat, barley, and oats irrigated
3. Begin development of classification procedures for separating wheat, barley, and oats within the small grains spectral stratum

---

*January 1982

B. Purpose:

1. Improve precision of statewise acreage estimates for wheat, barley, oats and their total

2. Significantly improve precision of county level estimates of wheat, barley, oats, and their total

3. Improve estimate precision at all levels for acreage of wheat, barley, and oats irrigated

4. Provide an accurate map of small grains average, and ultimately a map of wheat, barley, and oats

C. Approach:

1. Select 2 or 3 Landsat frames for 1982 and register minimum number of dates required for adequate small grains classification

2. Register Landsat dates to one another, and then register the resulting set to a north-south projection

3. Acquire June Enumerator Survey (JES) data for small grains (crop type, irrigation code) using revised recording form

4. Digitize county, basin, and field boundaries, together with field identification (type, irrigation code); register to north-south coordinate system defined earlier

5. Generate a small grains-only classification using Boolean technique developed in Experiment #1a and modified as necessary for location and/or date differences

6. Specify regression or ratio estimators for estimation of wheat, barley, oats, and their total acreage using registered JES data and Landsat small grains classification

7. Generate acreage estimates and associated estimates of error

8. Evaluate with respect to

    a. estimate precision at county and county group levels

    b. projected estimate of statewide precision

    c. map accuracy

    d. sensitivity of above to missing Landsat dates

    e. cost

    f. time required with particular reference to when results might be available in an operational environment

III. Experiment #2a: Multiple Crop Classification Procedures

A. Objectives:

1. Develop initial greeness-brightness profile classification procedure

2. Evaluate and refine DWR-APT crop classification procedure (This task shared with DWR-APT effort)

B. Purpose

1.. Develop and cull classification techniques prior to large-area test on 1982 data in experiment #2b

2. Establish requirements for Landsat data normalization

C. Approach:

1. Use present 1976 north-south registered Landsat and ground data set for a 1° x 1° block in the Sacramento Valley; augment Landsat dates as necessary from current library at U.C. Berkeley

2. Determine Landsat data normalization procedure required for effective use of greeness-brightness bands; likely adaption of XSTAR algorithm for haze and sun-angle correction

   a. data sampling issue

   b. new Landsat data format issue

3. Specify greeness-brightness Boolean procedure based on knowledge of cropping practices and upon lessons learned in LACIE and AgRISTARS

4. Evaluate classification accuracy of greeness-brightness technique; modify technique as necessary and re-evaluate

5. In cooperation with NASA-Ames, continue to evaluate and refine the classification technique applied to the Sacramento Valley 1976 data set as part of the DWR-APT.

6. Identify classification procedure options suitable for testing and further development in experiment #2b

IV. Experiment #2b: Large Area Development and Test of a Multiple Crop Classification and Area Estimation Procedure

A. Objectives:

1. Development and test of a joint Federal-State end-of-season (EOS) cover type mapping and area estimation procedure

2. Determine prior-to-EOS map accuracy and estimate precision for simulated June, July, and August reporting dates

3. Provide an initial test of prior-to-EOS classification procedures (from experiment #3a) for area estimation

4. Develop

    a. initial protocol for data sharing between agencies

    b. initial protocol for sharing of processing functions (if appropriate) between agencies

    c. initial map product generation procedures capable of supporting both Federal and State requirements

B. Purpose:

1. Improve both EOS and prior-to-EOS crop acreage estimate precision at state, regional, and county levels

2. Provide a cost-effective and flexible land use mapping capability for the California DWR and possibly other state users

3. Demonstrate cost savings potential and cooperative mechanism for joint Federal and State remote sensing-based agricultural inventory

C. Approach:

1. Specify general aspects of the experimental design

2. Select sets of counties within 3 frame area for development of multicrop estimation and mapping capability

3. Specify set of Landsat classification procedures to be evaluated and/or developed; tentatively these are

    a. current USDA-RSB procedure

    b. a version of the DWR-APT procedure undergoing evaluation on the 1976 Sacramento Valley data set

    c. new greeness/brightness time profile procedure taking best advantage of recent research

4. On the basis of previous research, local crop calendar data, and classifier requirements, specify and acquire Landsat digital and transparency data for a 3 frame area; (see also Landsat data requirements section of Idea Document #1, December 1, 1981)

5. Register Landsat dates using Editor

6. Register Landsat data sandwich to north-south coordinate system

7. Obtain 1982 ground data

   a. JES (revised forms) plus revisit check

   b. DWR land use survey data for Shasta, Tehama, San Joaquin, and San Benito counties

   c. supplementary segment data if required in small grains or multicropped areas (may not be required)

8. Edit ground data

9. Digitze USDA California sample frame boundaries, JES sample segment boundaries, JES and DWR field boundaries, county and other administrative boundaries, other land use related boundaries, (e.g. masking categories - riparian, wildland, etc.)

10. Register boundary and land use data to registered Landsat file(s)

11. Generate crop/land use classification map according to each classification procedure; perform for EOS, then resources permitting, apply in successive fashion, August, July, and June cutoff dates for use of Landsat data

12. Specify estimators by reporting unit, and then generate crop acreage estimates according to these estimators

13. Evaluate results with respect to

    a. crop acreage estimate precision at each reporting level by cutoff date

    b. projected estimate of statewide precision by cutoff date

    c. estimate bias within counties mapped by DWR by cutoff date

    d. map accuracy by cutoff date including analysis of omission/commission errors

    e. sensitivity of above to missing Landsat dates for a given cutoff date

    f. cost

    g. expected time required for operational inventory

14. Develop recommendations for implementation and/or future experiments

    a. data set preparation procedure

    b. classification procedure

    c. sample frame

    d. estimators

    e. map product generation techniques

15. Develop recommendations regarding interagency cooperation for

    a. data acquisition and sharing

    b. map product standardization

    c. sharing of product generation

    d. map product exchange

    e. coordinate use of sample frame for producing area estimates

V. Experiment #3a: Early Season Classification Technique Development

A. Objective:

Develop techniques for early and prior-to-end of season crop type classification for an initial set of important crops*

1. Cotton
2. Rice
3. Alfalfa
4. Pasture
5. Other?

B. Purpose:

Develop classification procedures most likely to cost-effectively support USDA June, July, and/or August estimation objectives.

C. Approach

1. Define and conduct an exploratory analysis of early season classification procedures on the 1976 Sacramento Valley data set; this analysis will provide the primary input for early season rice and alfalfa classification procedures evaluated for area estimation purposes in Experiment #2b

2. Determine local crop calendars and cropping practices for crops of interest within 3 Landsat frame areas for use with 1982 data

3. On the basis of this review, select test sites (e.g. counties) in 2 or possibly 3 frames for use with 1982 data

4. Obtain

    a) at least one Summer and one Fall 1981 Landsat date for each frame

---

*In addition to small grains drops treated in Experiments #1a and #1b

   b) selected Winter, Spring and Summer 1982 Landsat dates for each frame depending on the crop calendars for the crops of interest in each test site

5. Register Landsat dates
   a) 1982 dates first priority
   b) 1981 dates second priority
6. Obtain 1981 and 1982 JES and selected DWR ground map data, and register to Landsat base (either Landsat coordinate system, or north-south coordinate system used in Experiment #2b)
   a) 1982 data first priority
   b) 1981 data second priority
7. Create and register test site and land use strata boundaries to coordinate base
8. Resources permitting, obtain and register ancillary (soil association, field boundary) to coordinate base
9. Resources permitting, obtain early season phone or mail survey data and convert to prior probabilities of crop type presence
10. Specify, develop, and test early season, and prior-to-end of season classification strategies
11. In conjunction with Experiment #2b, evaluate with respect to
    a) estimate county level precision (including documentation of correlation and relative efficiency statistics)
    b) projected regional precision
    c) estimate bias
    d) map accuracy
    e) cost
    f) realistic operational time requirements

APPENDIX II: DEVELOPMENT OF RECOMMENDATIONS FOR ACQUISITION OF 1982 LANDSAT DATA

In order to facilitate the selection of Landsat MSS data for the Cooperative Study, it was important to predict which overpasses would provide the best opportunities for crop type separation. This problem was approached by first constructing historical crop calendars.

SRS Crop and Weather Reports, dating back to 1978, were used for this purpose, as well as back issues of the California-Arizona Farm Press. Also important were interviews with personnel from the U.C. Cooperative Extension Service, the U.S. Soil Conservation Service, and the California Department of Water Resources.

The crop calendar for 1978 was selected as an average one upon which to base relative crop development variations. Early 1982 crop development trends (as obtained from Coop transect data and the sources cited above) were used to adjust the relative (1978) crop development patterns to the 1982 growing season.

The resulting, predicted 1982 crop calendar was in turn used to generate (Beck 1982) a table (Table II.1) showing the expected presence or absence of green canopy for each major crop type. The discriminant used was the Landsat MSS "threshold of detection" for growing vegetation, which corresponds roughly to 20 percent ground cover (see, for example, Rice, et al, 1980).

The multitemporal sequence of green canopy cover unique to each crop type could then be exploited for crop type separation. Based on Table II.1, a set of crop type separation matrices were developed, each matrix corresponding to a 1982 Landsat 3 overpass. Figure II.1 shows two examples of the matrix. Each dot represents a pair of crops that were expected to be potentially separable on that particular overpass, based on green canopy presence. An analyst, using the full set of matrices, then constructed a list of recommended acquisitions for identifying each major crop type of interest; the list varied with number of acquisitions allotted for purchase (see Table II.2).

	J	F	M	A	M	J	J	A	S	O	N	D
CORN - GRAIN	0	0	0	0	0	1	1	1	0	0	0	0
CORN - SILAGE	0	0	0	0	0	0OR1	0OR1	1	1	1	0	0
GRAIN	1	1	1	1	1OR0	0	0	0	0	0	0	0OR1
ALFALFA	1	1	1	1	1	1	1	1	1	1	1	1
D BEANS	0	0	0	0	0	1	1	1	1	0	0	0
COTTON	0	0	0	0	0	0	1	1	1	0	0	0
TOMATOES	0	0	0	0	1	1	1	1OR0	1OR0	1OR0	0	0
RICE	0	0	0	0	0	1	1	1	1	0	0	0
O.W. BEETS	1	1	1OR0	1OR0	1OR0	1OR0	1OR0	1OR0	1OR0	1OR0	1OR0	1
NEW BEETS	0	0	0	0OR1	1	1	1	1	1OR0	1OR0	0	0
SUNFLOWERS	0	0	0	0	0	1	1	1	1	0	0	0

Table II.1. Matrix representing the presence (1) or absence (0) of green canopy cover for the 1982 growing season. The discriminant used for determining canopy presence is 20% ground cover.

2 SEPTEMBER

	GRAIN	ALFALFA	TOMATOES	CORN	RICE	D BEANS	S BEETS	SUNFLWRS
COTTON	●			●				
GRAIN		●		●	●	●		●
ALFALFA				●				
TOMATOES								
CORN						●	●	
RICE								
D BEANS								
S BEETS								

20 SEPTEMBER

	GRAIN	ALFALFA	TOMATOES	CORN	RICE	D BEANS	S BEETS	SUNFLWRS
COTTON	●			●	●			
GRAIN		●		●	●	●		●
ALFALFA					●			●
TOMATOES					●			
CORN					●	●		
RICE						●	●	●
D BEANS								
S BEETS								

Figure II.1. Examples of matrices indicating (by the presence of a dot) which crop types were potentially separable on specific 1982 Landsat pass dates based on the information provided in Table 2.

Table II.2. Initial Landsat acquisition recommendations developed from simultaneous inspection of Landsat pass-specific matrices of the kind shown in Figure 2.

	OPTIMUM SEPARATION GIVEN...			
CROP	4 DATES	3 DATES	2 DATES	1 DATE
COTTON				26 OCT
GRAIN		8 NOV 81 11 APR 4 JUN	8 NOV 81 11 APR	
ALFALFA[1]			8 NOV 81 11 APR	
TOMATOES	8 NOV 81 11 APR 4 JUN 16 AUG	11 APR 4 JUN 16 AUG	11 APR 16 AUG	
CORN		16 AUG 2 SEP 8 OCT	16 AUG 2 SEP	
RICE				20 SEP
D BEANS		11 APR 2 SEP 8 OCT	11 APR 2 SEP	
S BEETS[2]	8 NOV 81 4 JUN 2 SEP 8 OCT	8 NOV 81 2 SEP 8 OCT	8 NOV 81 8 OCT	
SUNFLWRS[3]			17 MAY 22 JUN	

(Footnotes on following page.)

FOOTNOTES FOR TABLE II.2

[1] Alfalfa should be green (or have green stubble) on most acquisitions throughout the year, so any spread of dates should suffice for alfalfa separation.

[2] Sugar beets are extremely difficult to typify, as they are planted and harvested almost continuously throughout the year. Acquisitions purchased in an attempt to separate this crop, based on canopy presence/absence alone except in its overwintered form, would probably be insufficient, and buy little information in return.

[3] Only a comparatively small fraction of the agricultural acreage in the Central Valley is dedicated to sunflowers, and the bulk of them are grown in the Delta. Therefore, they are only confusors in that specific area, and only of limited concern.