General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

SPACE
SCIENCES
LABORATORY

ANALYSIS OF THE QUALITY OF IMAGE DATA ACQUIRED BY THE
LANDSAT-4 THEMATIC MAPPER AND MULTISPECTRAL SCANNER

Principal Investigator
Professor Robert N. Colwell

Project Leader
Stephen D. DeGloria

Remote Sensing Research Program
Space Sciences Laboratory
University of California
Berkeley, California 94720

Period of Performance

January 3, 1983 - December 31, 1984

Final Technical Report

UNIVERSITY OF CALIFORNIA, BERKELEY

December 31, 1984

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Analysis of the Quality of Image Data Acquired by the Landsat-4 Thematic Mapper and Multispectral Scanner		5. Report Date 31 December 1984
		6. Performing Organization Code
7. Author(s) R. N. Colwell and S.D. DeGloria		8. Performing Organization Report No.
9. Performing Organization Name and Address Remote Sensing Research Program Space Sciences Laboratory University of California Berkeley, California 94720		10. Work Unit No.
		11. Contract or Grant No. NAS5-27377
		13. Type of Report and Period Covered Final Technical Report
12. Sponsoring Agency Name and Address		
		14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

The spatial, geometric, and radiometric qualities of Landsat-4 Thematic Mapper (TM) and Multispectral Scanner (MSS) data were evaluated by interpreting, through visual and computer means, film and digital products for selected agricultural and forest cover types in California. Multispectral analyses employing Bayesian maximum likelihood, discrete relaxation, and unsupervised clustering algorithms were used to compare the usefulness of TM and MSS data for discriminating individual cover types.

Some of the significant results from these studies were as follows: (1) for maximizing the interpretability of agricultural and forest resources TM color composites should contain spectral bands in the visible, near-reflectance infrared and middle-reflectance infrared regions, namely TM 4 and TM 5, and must contain TM 4 in all cases even at the expense of excluding TM 5; (2) using enlarged TM film products, planimetric accuracy of mapped points was within 91 meters (RMSE east) and 117 meters (RMSE north); (3) using TM digital products, planimetric accuracy of mapped points was within 12.0 meters (RMSE east) and 13.7 meters (RMSE north); and (4) applying a contextual classification algorithm to TM data provided classification accuracies competitive with Bayesian maximum likelihood.

17. Key Words (Selected by Author(s)) Landsat, remote sensing, thematic mapper, multi-spectral scanner, geometric processing, extra band recommendations, and spectral information extraction		18. Distribution Statement	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 144	22. Price*

ANALYSIS OF THE QUALITY OF IMAGE DATA ACQUIRED BY THE
LANDSAT-4 THEMATIC MAPPER AND MULTISPECTRAL SCANNER

Principal Investigator
Professor Robert N. Colwell

Project Leader
Stephen D. DeGloria

Remote Sensing Research Program
Space Sciences Laboratory
University of California
Berkeley, California 94720

Final Technical Report

December 31, 1984

EXECUTIVE SUMMARY

The spatial, geometric, and radiometric qualities of Landsat-4 Thematic
Mapper (TM) and Multispectral Scanner (MSS) data were evaluated by
interpreting, through visual and computer means, film and digital products
and statistical data for selected agricultural and forest cover types in
California. Multispectral analyses employing Bayesian maximum likelihood,
discrete relaxation, and unsupervised clustering algorithms were used to
compare the usefulness of TM and MSS data for discriminating individual
cover types. The promotion of the routine and practical uses of Landsat-4
data was accomplished by distributing data to a limited number of co-
investigators, and by presenting our results to many groups and individuals,
both domestic and foreign, who are potential users of such data.

The spectral and spatial quality of TM and MSS data was evaluated by
analyzing summary statistics and film products using both qualitative and
quantitative interpretation methods. Significant results obtained were:
(1) for maximizing the interpretability of agricultural and forest
resources Tri color composites should contain spectral bands in the visible,
near-reflectance infrared and middle-reflectance infrared regions, namely
TM 4 and TM 5, and must contain TM 4 in all cases even at the expense of
excluding TM 5; (2) MSS film composites were more interpretable than all
TM film composites for certain cover types, and those TM film composites
without TM 4 for all forest cover types; (3) TM spatial resolution
permited improved discrimination of road and stream networks, area and shape
of agricultural fields and forest clearings, and significant cover type
boundaries; and (4) TM color composites using the thermal band (TM 6)
along with TM 4 and a visible band were useful for discriminating
temperature and moisture conditions, and sites suitable for possible
reforestation activities.

The geometric quality of TM film and digital products was evaluated by
measuring coordinates of known surface features on TM image and digital
products and on USGS published maps. These paired observations between TM
coordinates (point and line) and map coordinates (UTM east and north) were
related using standard linear least-squares regression techniques.
Prediction of map coordinates for an independent set of "test" points using
first-order polynomials and analysis of variance of residual errors yielded
the following significant results: (1) using enlarged TM film products,
planimetric accuracy of mapped points was within 91 meters (RMSE east) and
117 meters (RMSE north); (2) using TM digital products, planimetric
accuracy of mapped points was within 12.0 meters (RMSE east) and 13.7 meters
(RMSE north); and (3) TM scene-to-scene registration was within 0.778
pixels (RMSE points) and 0.596 pixels (RMSE lines). The differences in RMS
errors were a result of reduced control point measurement errors when using
the digital product at full resolution as opposed to the manual
interpretation and digitizing of control points on the enlarged photo
product.

The multispectral analysis of TM and MSS data using computer-assisted techniques yielded the following significant results: (1) major forest vegetation types were consistently separated by applying a clustering algorithm to four bands of TM data, but discrimination of individual conifer species was not possible using this data set alone; (2) TM data at either full or degraded spatial resolution achieved higher classification accuracies than MSS data when comparable bands were used; (3) use of all TM bands significantly improved classification accuracy; and (4) use of a contextual classification algorithm provided classification accuracies competitive with Bayesian maximum likelihood.

Promotion of the routine and practical use of Landsat-4 and -5 data was accomplished by presenting numerous seminars, publishing our research results in symposia proceedings and journals, and distributing data to and participating in the research of selected co-investigators in academia, state government, and the private sector.

In summary, the overall utility and quality of Landsat-4 TM and MSS data are excellent, and sufficient for meeting most of the mapping, inventory, and monitoring needs of the renewable resource specialist.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

R.N. Colwell, Principal Investigator
S.D. DeGloria, Project Leader
Remote Sensing Research Program

1.1 Background

The launch of Landsat-4, the fourth in a series of land observing satellites, on 16 July 1982, followed by the launch of Landsat-5 on 1 March 1984, by the National Aeronautics and Space Administration (NASA) marked a significant milestone in the era of Earth observations from space. Previous Landsats have allowed Earth scientists and resource specialists to detect, identify, and quantify natural resources over large areas on a repetitive basis. This discrimination is possible using one or more spectral bands of the Multispectral Scanner (MSS) in the visible and near-infrared regions of the electromagnetic spectrum. The MSS detects reflected energy in four discrete spectral bands (green, red, and two reflectance infrared bands) from a ground area of approximately 83 meters square over a 185km swath every 16 days. This reflected energy is encoded into 64 gray levels per band. Based on the demand for more detailed resource management and exploitation information, a new electro-optical sensor, called the Thematic Mapper (TM), was developed by NASA which represented a vast improvement in spectral and spatial resolution and radiometric performance over the MSS. The TM sensor detects reflected energy in six narrow bands in the visible and reflectance infrared regions, and emitted energy in one broad band in the thermal infrared region of the spectrum. This energy is detected, pixel-by-pixel, from a ground area of approximately 30 meters square over a 185km swath every 16 days. The reflected and emitted energy are encoded into 256 gray levels (Engel, et al., 1983a, 1983b). Both the MSS and TM sensors are currently operating on both Landsat-4 and Landsat-5. The major objectives of the Landsat-4 and -5 missions are summarized in Table 1.1 (NASA, 1984a). NASA (1984b) provides a concise summary of the capabilities of the current Landsat system and the future role these satellites will play in characterizing the Earth's natural resources and in promoting the wise management of these diminishing resources. Tables 1.2 and 1.3 summarize the basic Landsat-4 orbit and image format specifications, and actual spectral regions and ranges of both the TM and MSS bands, respectively.

With the successful implementation of the Landsat-4 and Landsat-5 Earth observation satellite systems, NASA is conducting a sponsored program to quantitatively evaluate the spectral, spatial, geometric, and radiometric quality of the data acquired by both satellites. The Landsat Image Data Quality Assessment (LIDQA) program is serving to provide the user community with the most comprehensive evaluation of a newly implemented satellite system in the history of civil Earth observations from space.

Of the many anticipated uses of the data acquired by the Landsat series of satellites, the improvement in the efficiency of conducting renewable resource mapping and inventory activities is of great importance to both public and private agencies. This increase in efficiency will occur if the satellite data: (1) can be implemented by the user group in film or digital format; (2) can provide consistent information within user-specified limits of performance, and (3) are cost-competitive with other forms of remotely sensed data. To enhance the probability that the satellite data will be implemented, potential user groups must: (1) have both a personal and organizational interest in the data, (2) be aware of the data's advantages and limitations, and (3) have a level of expertise to properly evaluate the quality of the data for meeting that organization's information needs. To this end, personnel from the Remote Sensing Research Program at the Space Science Laboratory, University of California – Berkeley along with other research cooperators investigated the spectral, spatial, geometric, and radiometric qualities of Landsat-4 Thematic Mapper and Multispectral Scanner data, in film and digital format, for meeting the mapping and inventory needs of the renewable resource management community.

As part of the LIDQA team, our responsibilities under the NASA-funded program were to: (1) conduct research evaluating the quality of Landsat-4 TM and MSS imagery in the area of this investigator's expertise, namely renewable resource survey, and (2) promote the routine and practical use of these data for the betterment of mankind (NASA, 1984a). Our research focused primarily on the evaluation of film products generated from computer compatible tapes for the purpose of discriminating natural targets in agricultural and forest environments. This final technical report summarizes in the following nine chapters the results of our investigations.

1.2 Research Objectives

To fulfill our research responsibilities under the LIDQA program, several objectives were defined for the characterization of Landsat-4 TM and MSS data quality. These research objectives were:

(1) to develop a basic understanding of the Landsat-4 TM and MSS data in terms of spectral, spatial, geometric, and radiometric characteristics;

(2) to develop a basic understanding of data products in computer compatible tape (CCT) and film formats, and help to minimize any data handling problems potential users may encounter with either format;

(3) to determine the extent to which agricultural and forest cover types and conditions can be detected and identified on Landsat-4 film and digital data products; and

(4) to promote the use of Landsat-4 data by a diverse user community active in the management of renewable natural resources.

To accomplish the objectives as stated above, these investigators and a number of research cooperators evaluated Landsat-4 data acquired over a number of preselected study sites in California. The research cooperators and their respective institutions were: Mr. Ralph Bernstein and Dr. Silvano

Di Zenzo, IBM Corporation, Palo Alto and Rome Scientific Centers, respectively (Chapter 5); Mr. Jay Baggett, California Department of Water Resources, Sacramento (Chapter 7); Professor Thomas Dickert, Center for Environmental Design Research, University of California - Berkeley (Chapter 8); and Dr. Walter Carnahan, Indiana State University, Terre Haute, Indiana (Chapter 9).

The study sites selected for our analyses are shown in Figure 1.1. Agricultural sites were located in San Joaquin, Sacramento, Yolo, Solano, Sutter, Colusa, Glenn, Butte, and Tehema Counties; forestry sites were located in Plumas and Santa Clara Counties; and the coastal zone watershed study site was located in Santa Cruz County.

The research results of this investigation are summarized by chapter and authored by those individuals taking the lead responsibility for conducting and documenting the research.

1.3 Acknowledgements

1.4 References

Engel, J.L. and O. Weinstein. 1983a. The thematic mapper - an overview. IEEE Trans. Geosci. Remote Sensing, Vol GE-21(3):258-265.

Engel, J.L., J.C. Lansing, Jr., D.G. Brandshaft, and B.J. Marks. 1983b. Radiometric performance of the Thematic Mapper. Proc. 17th Int. Symp. Remote Sensing Environ. 25 p.

NASA. 1982. Landsat-D to ground station description. Goddard Space Flight Center Interface Contr. Doc. #435-D-400, Rev.#4, February. Greenbelt, MD. 91 p, Appendices.

NASA. 1984a. Landsat-4 Science Investigations Summary. NASA Con. Publ. #2326, Vol. I,II.

NASA. 1984b. A prospectus for Thematic Mapper research in the Earth sciences. NASA Tech. Memo #86149. Goddard Space Flight Center. 66 p.

Figure 1.1 Location of the agricultural, forestry, and coastal watershed study sites in northern California used by personnel from the Remote Sensing Research Program, University of California – Berkeley and other research cooperators, to assess the spatial, geometric, spectral, and radiometric quality of the data acquired by the Landsat-4 Thematic Mapper and Multispectral Scanner sensors.

Table 1.1 Major objectives of the Landsat-4 and Landsat-5 missions (NASA, 1984).

● To provide for system-level feasibility demonstrations in concert with NOAA and user agencies to define the characteristics of an operational system

● To assess the capability of the thematic mapper (TM) and associated systems to provide information for Earth resources management

● To provide for continued availability of multispectral scanner (MSS) data

● To provide a transition for both domestic and foreign users from MSS data to the higher resolution and data rate of the TM

● To permit continued foreign participation in the program

Table 1-2 Basic Landsat-4 spacecraft orbit and scene format constants (NASA, 1982).

SPACECRAFT

Sun Synchronous	
Nominal Altitude	705.3 km
Nominal Swath Width	185.0 km
Orbit Inclination	98.2°
Orbit Period	98.9 min
Orbits per Day	14.5
Repeat Period	16 days
Sidelap (at equator)	7.6%
Sidelap (at 40° Latitude)	29.0%
Descending Node (local time)	0930-1000

SCENE

MSS (partially processed image: CCT-AM)

Nominal number of pixels/line	3240
Number of lines	2400
Nominal interpixel distance	57.0m
Nominal interline distance	82.7m

MSS (fully processed image: CCT-PM)

Number of pixels/line	3548
Number of lines	2983
Interpixel distance	57.0m
Interline distance	57.0m
Signal quantization levels	127

TM (fully processed image: CCT-PT)

Number of pixels/line	6967
Number of lines	5965
Interpixel distance	28.5m
Interline distance	28.5m
Signal quantization levels	256

7

Table 1.3 Spectral range of the TM and MSS sensors on Landsat-4 (NASA, 1984 a,b).

THEMATIC MAPPER

Band	Spectral Region	Spectral Range, μm
1	Blue	0.45 - 0.52
2	Green	0.53 - 0.61
3	Red	0.62 - 0.69
4	Near reflectance IR (photographic IR)	0.78 - 0.91
5	Middle reflectance IR	1.57 - 1.78
7	Far reflectance IR	2.10 - 2.35
6	Thermal IR	10.42 - 11.60

MULTISPECTRAL SCANNER

Band	Spectral Region	Spectral Range, μm
1	Green	0.50 - 0.61
2	Red	0.60 - 0.70
3	Near reflectance IR	0.70 - 0.81
4	Near reflectance IR (photographic IR)	0.81 - 1.02

8

Chapter 2

SPECTRAL QUALITY OF LANDSAT-4 THEMATIC MAPPER
AND MULTISPECTRAL SCANNER DATA

S. D. DeGloria, A. S. Benson, and K. J. Dummer
Remote Sensing Research Program

2.0 Introduction

The primary goal of our research related to the evaluation of the spectral quality of Thematic Mapper (TM) and Multispectral Scanner (MSS) data was to determine the extent to which both digital and film products can discriminate agricultural and forest cover types and conditions important to resource survey and management objectives. The major objectives for this phase of our research investigation were to: (1) develop a basic understanding of the Landsat-4 TM and MSS data regarding spectral, spatial, geometric, and radiometric characteristics, CCT and film formats and products, and special problems in data handling; (2) determine the advantages and limitations of selected TM color composite film products for detecting and identifying agricultural and forest cover types; and (3) evaluate the quality of TM data in comparison to simultaneously-acquired MSS data in film product format for discriminating forest cover types.

Our study sites in California which have been used for the past 40 years represent a variety of agricultural and forestry conditions and provide numerous opportunities for evaluating the quality of any sensor system. California has the most diversified and productive agricultural system in the world including over 220 commercial crops growing on over 25 million hectares; varied agronomic practices; diverse field sizes and shapes; and variable soil and landform conditions. California is also richly endowed with over 16 million hectares of forestland which is the largest extent of forests in the contiguous United States. California leads the nation in forest products sold and the number of people employed by the timber industry, but ranks only third in the volume of standing timber (256 billion board feet). Our forestry study sites are located primarily in the western mixed conifer type, one of the most productive types in California.

The approach for accomplishing our research objectives related to the evaluation of Landsat-4 TM and MSS film product quality was as follows:

(1) Acquire the first available Landsat-4 scenes for our active study areas in California's Central Valley and northern Sierra Nevada. An active study site is one in which ongoing projects are collecting detailed field data consisting of ground and aerial photography, recording quantitative descriptions of specific field or stand conditions, and utilizing existing land survey maps.

(2) Acquire, process, and catalog large-scale, small format color and color infrared oblique aerial photography of selected areas within these scenes to document major agronomic and forest stand conditions at the time of the overpass. This aerial photography is generally acquired at a scale between 1:1000 and 1:4000 using commercially available Ektachrome film and a hand-held dual 35mm camera system. Acquisition generally occurs coincidently with the satellite overpass of interest if suitable notification of an overpass is provided to the investigator by the responsible agency or by existing acquisition schedules.

(3) Compile the available ground data for the area imaged to reconstruct, as accurately as possible, the environmental conditions prevalent at the time of the overpass. The sources of ground data for our research included: (a) Land Use Survey Maps of the California Department of Water Resources (DWR), published at a scale of 1:24,000, which included individual field boundaries which are outlined and labeled; (b) the U.S. Department of Agriculture's (USDA) monthly statistical summaries of crops and climate; (c) field crew notes and data, as compiled by personnel of our Remote Sensing Research Program (RSRP) at the University of California - Berkeley; and (d) forest cover type maps prepared under U.S. Forest Service sponsorship, published at a scale of 1:24,000, which included type boundaries and polygon labels corresponding to the forest cover type classification system in use by the Plumas National Forest.

(4) Produce TM and MSS black-and-white and color composite products for interpretation based on a linear mapping process using an interactive digital image color display and a Matrix color graphics camera.

(5) Locate the field data and aerial photographic coverage on the TM and MSS imagery.

(6) Relate the environmental conditions to the TM and MSS spectral data in both analog (film) and digital (numeric) formats.

(7) Determine the interpretability of major agricultural crops and forest cover types using these Landsat-4 image products based on established interpretation techniques (Colwell, 1978; Benson and Dummer, 1983). Interpretability requires that the image products allow both the detection and correct identification of features of interest. Detection requires, at a minimum, the simple recognition or awareness that a feature is present. To identify the feature requires a synthesis of spectral, spatial, textural, and temporal characteristics (Hay, 1982).

Our objectives call for the interpretation of diverse agricultural and forest cover types using various image products. Such products have both advantages and limitations in that photo-like imagery provides the best format for evaluating spatial and textural characteristics, but it has limited usefulness for analyzing detailed spectral and temporal characteristics (Hay, 1982). When an image is created from digital data, whether in black-and-white or in color, the data must be compressed into a

10

limited number of gray (or color) levels thus obscuring subtle spectral differences of a feature of interest. In addition, color variation due to photographic processing may confound spectral characteristics needed for consistent multi-temporal analysis (Odenweller and Johnson, 1982). Although these conditions suggest that some of our analyses of Landsat data quality can best be made using the original numeric data rather than images produced from these data , we have chosen to evaluate film products as we believe there will be an increased interest in using high quality film products in the operational sector as spatial resolution of current and proposed satellite systems continues to increase.

This chapter is divided into three major sections relating to our results on the qualitative analysis of film products in both our agricultural and forest study sites (Sections 2.1 and 2.2) and on the quantitative analysis of TM and MSS film products for the interpretation of selected forest cover types. The materials contained in Sections 2.1 and 2.2 are enhancements to material previously published (DeGloria, 1984; DeGloria, et al., 1984). These enhancements include the addition of color illustrations, and discussions highlighting the advantages and limitations of various spectral band combinations of TM data for discriminating agricultural and forest cover types.

2.1 Spectral Quality of TM Film Products in Agricultural Environments

2.1.1 Introduction

Landsat MSS film products have represented an important part of agricultural surveys by serving as a stratification base. Areas of similar crop types and agronomic practices are delimited on photographic images on the basis of their spectral and spatial characteristics in a single-date or multi-temporal mode. Because the detection and identification of different agronomic conditions is a critical step in an effective stratification procedure, the high quality TM film products from the Landsat-4 and -5 systems should play an even more valuable role in this regard as opposed to those from the MSS systems.

The spectral quality of TM data was evaluated by determining the extent to which natural targets in agricultural environments could be discriminated using specially generated and commercially available film products. The specific approach for this phase of our investigation was to: (1) extract spectral statistics for specific agricultural cover types, (2) compare these data to selected photographic reproductions of the same area representing various spectral band combinations, and (3) determine which color composite provides the most information for stratification purposes.

2.1.2 Experimental Methods

Two agricultural study sites in California were selected for analysis (Figure 1.1). The first site is located in San Joaquin County and served as our early agricultural site. The area is located approximately 110 km east of San Francisco and contains a diversity of cropping practices. Land use for the area was recently mapped by the California's Department of Water Resources (DWR)(1982). The second site is located in Yolo County, and served as our agricultural study site for the detailed analysis of TM and MSS data quality using coincidentally acquired field data and aerial photography.

The first Landsat-4 scene for the site in San Joaquin County was acquired on 8 December 1982 (#84014518082, WRS Path 43, Row 34). The image data were received by the Prince Albert Station, Canada, and forwarded to NASA-Goddard Space Flight Center (GSFC) for processing. The TM data were processed by the Landsat Assessment System at GSFC to produce a "P" tape (CCT-PT) in which the data were radiometrically and geometrically corrected. The simultaneously acquired MSS data were purchased from the EROS Data Center (EDC) as an "A" tape (CCT-AM) in which the data were only radiometrically corrected. No geometric corrections were applied to the MSS data.

The first Landsat-4 scene for the site in Yolo County was acquired on 12 August 1983 (#84039218143, WRS Path 44, Row 32). The TM data were transmitted to GSFC via the Tracking and Data Relay Satellite System (TDRSS) and the TDRSS receiving station at White Sands, New Mexico. The TM data were processed by the Thematic Mapper Image Processing System at GSFC as a "P" tape. The simultaneously acquired MSS data were purchased from the EDC as a "P" tape (CCT-PM).

Large-scale (1:4,300 - 1:13,000) color and color infrared oblique aerial photography was acquired over both study sites using a dual 35mm camera system operated from a light aircraft. The purpose of this photography was to document the environmental conditions prevalent at the time of the overpass. It was used in conjunction with available ground data to document crop type and conditions, cultivation and irrigation practices, and other surface conditions which contributed to the spectral characteristics of the scene. The aerial photography for the San Joaquin site was acquired on 20 January 1983, approximately six weeks following the overpass. Ground conditions at the time of the overpass were reconstructed from published ground and climatic data (DWR, 1982; USDA, 1983) and field notes made available by our RSRP project staff. An example of the land use map used to document agricultural cover type conditions is shown in Figure 2.1. Aerial photography and ground data for the Yolo study site were acquired coincidentally with the Landsat-4 overpass.

Based on the available ground data and aerial photography for the study sites, the TM and corresponding MSS digital data were extracted from the full scene CCT's, and written to separate disk files for subsequent statistical analysis and color composite production.

Within the study sites, agricultural fields of known cover type and condition were located in the digital data set, annotated, and statistics calculated for each individual sample field. Both digital numbers (DN) (or counts) were extracted from these sample fields and spectral radiance values were used for statistical analysis, plots, and image generation. Spectral radiance for each spectral band was calculated using the following relationship (Dozier, 1983; NASA, 1984):

$$L_i = \frac{DN_i - \hat{O}_i}{\hat{G}_i}$$

where

L_i = Spectral radiance, $mWcm^{-2}-sr^{-1}-\mu m^{-1}$, for TM Band i
DN_i = Digital Number, in counts, for TM Band i
$\hat{O}_i$ = Average Offset, in counts, for TM Band i
$\hat{G}_i$ = Average Gain, in counts/$mWcm^{-2}-sr^{-1}\mu m^{-1}$, for TM Band i
i = Spectral band index.

The average band offset and average band gain values are found for each scene in the Radiometric Calibration Ancillary Record, Leader File, Imagery Logical Volume (NASA, 1983). The summary statistics for the areas representing each study site are presented in digital numbers and spectral radiance; sample field statistics and plots are in digital numbers only.

The digital numbers for the two study sites were input to a standard statistical program to compute means, variances, coefficients of variation (CV), minimum and maximum values, skewness, kurtosis, and covariance and correlation matrices for the TM and MSS data. These summary statistics provided an overview of the frequency distributions by band for both agricultural study sites and were used at a later time to optimize color mapping tables for various image products. The CV, expressed as a percent,

13

was used to represent the relative variation about the mean, and was calculated using the mean ($\bar{x}$) and standard deviation (s) of individual cover types [(s/$\bar{x}$)*100]. This relative measure of variability allows for an easy comparison of spectral band means and their associated variance which may exhibit a wide range of values. Though both covariance and correlation values were calculated, only the correlation matrix is presented. For the purpose of evaluating the relationship between spectral bands, the values of the covariance matrix are difficult to interpret due to the large changes in the magnitude of the digital number between bands. The product moment correlation coefficient (r) is calculated from the band variances, and provides a standard form for evaluating these inter-band relationships. For the digital numbers extracted from individual fields, only the means, variances, and CV's were calculated. Through the use of various band combinations from the TM and MSS data sets for both agricultural study sites, color additive photographic images were produced to visually represent the digital data and to illustrate the spectral variability of selected agricultural types.

2.1.3 Results and Discussion

The summary statistics for the 255,328-pixel San Joaquin study sites are shown in Tables 2.1 and 2.2. Of the seven TM bands, Band 7 was the most variable (CV = 33.6, γ_1 = 2.7, γ_2 = 44.2, and range = 0-255) as a result of highly variable reflectance phenomenon exhibited by the ground cover conditions in this area. Band 6 was the least variable band (CV = 1.3) due to the low magnitude and narrow range of radiant temperature values of the variable ground cover types. Even in this agricultural scene during a period of low sun angle in the winter season, we found saturated pixels in both Band 5 and Band 7. These saturated pixels were a result of the high reflectance of natural targets such as concrete slabs used as loading areas and coarse textured soils. The interband correlation matrix for the seven TM bands is shown in Table 2.2. TM bands 1, 2, and 3 had a high positive correlation due to the nature of ground conditions, such as bare soil and dormant vegetation stands. The relatively low interband correlation between TM bands 3, 4 and 5 resulted from the combined influence of: (1) vegetated and non-vegetated surfaces absorbing red radiation (TM 3); (2) plant canopy structure and other surface properties reflecting near infrared radiation (TM 4); and (3) plant canopy moisture and other moist surfaces absorbing middle reflectance infrared radiation (TM 5). Although lower interband correlations can be found among other combinations of bands, namely TM 3, 5, and 6, and TM 4, 5, and 6, these combinations include the thermal band. Although TM 6 has high radiometric quality, its lower spatial resolution would decrease the value of the resulting composite film product for discriminating small, but important, surface features. Thus, a color composite of visible and infrared spectral bands of low interband correlation and high spatial resolution would combine to make an image product suitable for meeting most agricultural interpretation needs.

Individual field means and covariances (CV) for the TM and MSS bands are shown in Table 2.3 and plotted in Figure 2.2 Examination of these data illustrate how the addition of the TM middle reflectance infrared band (TM 5) has significantly enhanced our ability to discriminate different crop types. In the visible (TM 1, 2, 3) and near reflectance infrared (TM 4) bands, there are no significant differences between the field means for the sugar beet and alfalfa fields. In the middle reflectance infrared band,

14

however, there are significantly different mean spectral responses resulting
from the increased absorption of the solar radiation due to the higher leaf
water content of the sugar beet canopy in contrast to that of the alfalfa
canopy. This differential reflectance due to variations in leaf water
content has been previously documented under laboratory conditions (Tucker,
1980). Discrimination of orchard crops from moist bare soil on this date of
imagery is also enhanced with the addition of the TM 5 spectral data. The
mean spectral response of both cover types is similar for TM 1-4, but
dissimilar in TM 5. The bare soil has a higher mean spectral response in
TM5 than does the orchard crop, providing an improved capability for
detecting and identifying these cover crops. Differences in mean spectral
response between TM 4 and TM 5 for other cover types are graphically
displayed in Figure 2.2 Two color additive composites illustrating these
spectral relationships among cover types are shown in Figure 2.3: Figure
2.3a is a color composite using TM 4, 3, and 2; Figure 2.3b is a color
composite using TM 5, 3 and 2; and Figure 2.3c is a color aerial oblique
photograph. In Figures 2.3a and b, the composites were constructed by
projecting TM 4, 3, and 2 and TM 5, 3, and 2 through a red, green, and blue
filter, respectively, using a Matrix color graphics camera and 35mm
Ektachrome film.

By looking closely at Figures 2.3a and b in concert with the large-
scale color oblique photograph (Figure 2.3c, taken six weeks after the
overpass) and the detailed land use map prepared by the DWR (Figure 2.1), we
can see the significance of the spectral statistics graphically displayed in
Figure 2.2. The "L-shaped" feature in the top central portion of both TM
images is a set of three fields. Based on our interpretation of the aerial
photo (upper right corner) and field maps, and on a subsequent field visit,
we found that the fields on the left and lower right were sugar beet fields,
and the field in the center was an alfalfa field. The traditional color
composite (Figure 2.3a) shows no visual difference between these fields,
which is also reflected in the spectral statistics displayed in Figure 2.2.
In Figure 2.3b, which replaces the reflectance infrared band (TM 4) with the
middle infrared band (TM 5), there is a significant visual difference
between the sugar beet and alfalfa fields which is also graphically
displayed in Figure 2.2.

Close examination of these two composite images illustrates both the
advantages and the limitations of including TM 5 and excluding TM 4 in any
film or digital product for purposes of agricultural survey. We can
discriminate the sugar beets from the alfalfa on this particular winter
scene using Figure 2.3b; however, important, subtle color distinctions of
actively growing vegetation, as recorded on TM 4, are missing from the
composite. The most important signature lost in this composite which
excludes TM 4 are those related to emerging small grains plants in fields
distributed throughout the images. As we have subsequently learned in both
agricultural and forested environments, color composites must contain near
reflectance infrared data (TM 4) to be of any significant value to
vegetation survey, even at the expense of excluding the middle infrared
data (TM 5).

This assessment of the improved quality of TM film and digital products
when TM 4 is used was developed using a larger sample of Landsat-4 TM data
acquired over our Yolo County agricultural study site during peak crop
production. On 12 August 1983, a simultaneous TM and MSS data set were

acquired over the intensively cultivated Sacramento Valley, California (WRS, Path 44, Row 33). Both extensive field data and large-scale aerial photography were acquired coincident with the Landsat-4 overpass due solely to the advanced notice provided to these investigators by the Landsat Project Office at GSFC that the California scenes may be acquired as part of an operational test of the TDRSS (Masuoka, 1983). Because of this successful test of the TDRSS and the advanced notice by NASA, a data set was collected which will significantly contribute to our characterization of the Landsat system for many years to come.

For the present investigation related to the evaluation of the spectral quality of TM data as represented in film products, only small portions of the agricultural scene were selected for analysis so as not to expend project resources required for other phases of our research effort. We encourage and invite other investigators with an interest in quantifying the extent to which TM data can improve our ability to classify and map agricultural resources to work with us in the analysis of this extremely valuable data set in future sponsored research efforts. This type of investigation was external to our current objectives and will have to await these future opportunities. The preliminary analysis of two classification methods using TM and MSS data for crop type discrimination described in Chapter 5 is just one example of such research activity which can be conducted on a much larger data set covering most of the Sacramento Valley for which oblique aerial photography and detailed measurements were made on nearly 300 individual fields coincident with the August overpass.

For the small study area used to generate the film products, statistics including covariance and correlation matrices for a sample of TM data for all seven bands were calculated in order to optimize the setting of the color mapping tables used to generate film products using a linear contrast enhancement (see Section 2.3.4). The correlation matrix for this sample is shown in Table 2.4. A series of TM color additive composites for this area is shown in Figure 2.4: (a) TM 4, 3, and 1, (b) TM 5, 3, and 1, and (c) TM 5, 4, and 3, and (d) a color oblique aerial photograph, taken coincident with the overpass, illustrating the cover conditions in the uppermost portion of the TM composites. The unlabeled arrow in Figure 2.4a indicates the look direction of Figure 2.4d.

Several significant features are apparent when you carefully interpret these composites in Figures 2.4 a, b and c which reveal the relative quality of the three composites. These features are annotated on the composites to facilitate interpretation by the reader:

(1) Major agricultural cover types prevalent in these composites are bare soil, corn, rice, sugar beets, and tomatoes. The block of six fields in the center of Figure 2.4d consists of sugar beets with the top three appearing lighter green than the bottom three. This is due primarily to crop growth stage since row direction is the same for all fields. The variability in color might also be attributed to differences in plant varieties and irrigation practices. The field to the extreme left is a corn field with fallow fields to the top, bottom, and left.

(2) The arrow at point "a" on the composite images illustrates the spectral differences between the corn field and the adjacent sugar

16

beet fields. When using TM 4, 3 and 1 (Figure 2.4a) or TM 5, 3 and 1 (Figure 2.4b), one can easily differentiate the two crop types; this is not the case when using TM 5, 4 and 3 where there is insignificant change in the green colors representing a crop type change. In addition, at point "a" there is spectral confusion between sugar beets and bare soil on TM 5, 3, and 1 while no such confusion exists on the other two composites which contain TM 4. Other such differences can be detected throughout these image products where the composites containing the TM 5 middle infrared band, with or without the TM 4 reflectance infrared band, may occasionally improve the discrimination of certain crops. These image products represent more of the moisture relationships between cover types, whether vegetation is present or not, than the differences between vegetative cover types independent of relative surface moisture conditions.

Based on an evaluation of these and other image products for this scene, the best color composite of reflected energy continues to be one which contains the near reflectance (photographic) infrared band with one or two visible bands. Given the decrease in the spatial and spectral quality of composites constructed with Landsat-4's TM 2 band with its failed detector, the best composite would be one which used TM 4, 3, and 1 projected through red, green, and blue filters, respectively (Figure 2.4a).

2.2 Spectral Quality of TM and MSS Film Products in Forested Environments

2.2.1 Introduction

The objectives for this phase of our research were to: (1) develop a basic understanding of the TM data of forested environments with regard to spectral characteristics and variability, spatial resolution, and radiometric sensitivity, and (2) determine the extent to which major forest cover types can be detected and identified on TM film and digital products. The approach used was similar to that used in the agricultural evaluation where spectral statistics and film products were used to evaluate image quality.

2.2.2 Experimental Methods

The Landsat-4 scene used for our analysis was acquired on 12 August 1983 (#84039218143, WRS Path 44, Row 32). The TM data were transmitted to Goddard Space Flight Center (GSFC) via the Tracking and Data Relay Satellite System (TDRSS) and the TDRSS receiving station at White Sands, New Mexico. The TM data were processed by the Thematic Mapper Image Processing System at GSFC as a "P" tape.

Coincidentally with the overpass, large-scale (1:4300 - 1:13,000) color and color infrared oblique aerial photography was acquired over the Plumas National Forest study site (Figure 1.1) using a dual 35mm camera system operated from a light aircraft. This photography was used to document the environmental conditions prevalent at the time of the overpass. These photos are used in conjunction with available ground data to document forest canopy, non-forest cover, and understory conditions. The aerial photography was acquired within one hour of the overpass from 1200-1300 hours PDT. Ground conditions were documented based on interpretation of these photographs, the field data sheets from an intensive field study conducted in the area during August 1982, and these investigators' knowledge of forest conditions and management practices in the area.

Based on the available ground data and aerial photography for this study site, the TM and corresponding MSS digital data were extracted from the full scene CCT's. A 1200-by-1200 pixel block area (1,440,000 pixels) was extracted from the CCT-PT using the IBM image processing system at the Palo Alto Scientific Center with the assistance of Mr. Ralph Bernstein and Mr. Jeff Lotspiech. MSS data for the same area were extracted from the CCT-PM, and both data sets were written to separate disk files for analysis.

The digital numbers for the entire block area were input to a standard statistical program to compute means, variances, coefficients of variation (CV), minimum and maximum values, skewness, kurtosis, and covariance and correlation matrices for the TM and MSS bands. The summary statistics provided the overview of the frequency distributions, by band, of the TM spectral data, that was needed to optimize the color mapping tables used to produce various image products used for interpretation.

Forest stands of known cover type and conditions were located in the digital data sets, annotated, and statistics calculated for each stand. For the digital number extracted from the individual stands, only means, variances, and coefficients of variation were calculated. Both black-and-

white and color composite images were produced to: (1) visually represent the digital data, (2) illustrate the spectral variability of selected forest cover types, and (3) determine the extent to which major forest cover types could be detected and identified on TM digital and film products.

2.2.3 Results and Discussion

The summary statistics for the forestry study site are shown in Tables 2.5 and 2.6. The statistics in Table 2.5 include band means, variances, CV's, minimum and maximum values, and skewness and kurtosis for the frequency distribution of the 1,440,000 pixels per band. TM 5 and TM 7 are the most variable bands and can be considered normally distributed as is the least variable band, TM 6. Bands 1, 5, and 6 contain saturated pixels due to: (1) the high albedo of granitic rock outcrops (TM 1); high albedo of dry, bare soil on both level terrain and south facing slopes (TM 5); and the high radiant temperature of the granitic rock and bare soil (TM 6). The interband correlations in Table 2.6 show the familiar high positive correlation between TM 1, 2, and 3; and the low correlation between TM 4 and TM 1, 2, 3, 5, and 7. Both the middle- and far-reflectance infrared bands in the summer forested scene show high positive correlations with the thermal band. This is due to the close association between the radiant temperature responses of forest cover types and their relative moisture status. Generally speaking, the low reflective properties of vegetative canopies in TM 5 and 7 are directly related to actively growing vegetation under relatively high moisture conditions. As canopy moisture content increases, there is more absorption of the short-wave radiation by the canopy and a correspondingly lower radiant temperature of the canopy due to dissipation of heat energy through evapotranspiration. Conversely, conditions in a forested environment which have highly reflective properties in TM 5 and 7 (granitic outcrops and dry, bare, coarse-textured soils) also have high radiant temperatures due to low moisture content.

Table 2.7 contains the band means and coefficients of variation (standard deviation/mean x 100) for individual vegetated and non-vegetated cover types for the seven TM bands. As both sensors were simultaneously operating during this overpass, band means and CV's for three of the four MSS bands which most closely correspond to the relevant TM bands are also included. Although portions of both MSS 3 (0.70-0.81µm) and MSS 4 (0.81-1.02µm) spectral ranges fall within portions of the TM 4 (0.78-0.90µm) spectral range, MSS 4 was selected so as to avoid that highly variable region of the vegetation response curve between 0.70-0.80µm where the spectral reflectance of vegetation goes from a minimum at 0.645µm (peak chlorophyll absorption) to a maximum at 0.880µm (peak near infrared reflectance). Based on field plot data and synoptic surveys conducted in this area for the past several years, areas of known cover type conditions were located in the spectral data and extracted from the disk files for statistical analysis. For this phase of our investigation, only homogeneous stands of vegetation and other cover types, all of variable areal extent, were selected for analysis. Future analysis will include incorporating the spectral data with quantitative forest stand data (basal area, average stand height, age, diameter breast height, and timber site index) which were collected from 130 variable ground plots within the Meadow Valley 7½' topographic quadrangle.

The data from the six reflective bands are graphically displayed in Figure 2.5. For each band, the mean digital number (in counts), -1 standard deviation about the mean, is plotted for each cover type sampled. Figure 2.6 serves to illustrate these differences in the reflectance characteristics of forest cover types using both TM and MSS color composites and oblique aerial photography. Important cover types are annotated on the image and described below.

(1) A major cover type discrimination that foresters would like to make is between true fir stands, which are of relatively low economic value, and mixed conifer stands. By reviewing the spectral statistics, one finds no apparent spectral difference between these two major cover types. The only band in which we see some difference is in the thermal band (TM 6). This results from the higher topographic position that true fir occupies. The resulting lower radiant temperature of the true fir canopy allows the spectral discrimination between these two types using TM 6.

(2) Another discrimination of importance is between mixed conifer stands and those dominated by hardwood species, mainly black oak (Quercus kelloggii), which is an indicator of high timber site. Statistics from a mixed stand dominated by black oak are plotted under "Hardwood". From the statistical data, a major difference in spectral response is seen in TM 4, as one would expect. This is also evident when viewing the TM image products where an increasing percentage of hardwood species within a mixed conifer stand is reflected in increasing light red tones on the image distinguishing these cover types from the "pure" conifer stands having a dark red tone.

(3) Discrimination of trees from shrubs and herbaceous vegetation is quite obvious as seen by examining both the spectral plots and film products.

(4) Differentiation between wet and dry meadows can be accomplished using both TM 4 and TM 5 due to the differences in high photographic infrared signature (TM 4) and low middle reflectance infrared (TM 5) resulting from dense herbaceous growth under high canopy moisture conditions in the wet meadows.

(5) Dry, bare soil and rock outcrops exhibit the highest TM 5 and TM 6 radiance with some areas on south facing slopes causing pixel saturation (DN = 255) on both bands.

(6) The spectral response of deep and relatively clear water bodies in this area are as expected with increasing absorption of solar radiation as wavelength increases.

The four additive color composites in Figure 2.6 which used selected TM and MSS bands provide a good indication of the relative improvements of TM data over MSS in terms of both spatial and spectral resolution: (a) TM 4, 3 and 2; (b) MSS 4, 2, and 1; (c) TM 5, 3 and 2; (d) TM 6, 4 and 3; and (e) a stereo-pair of small format (35mm), low altitude color oblique aerial photography taken coincident with the Landsat-4 overpass. The composites cover a ground area of approximately 5200 ha and are reproduced here at a

scale of approximately 1:65,000. Each composite was produced by color coding the bands with red, green and blue filters, respectively, using a Matrix Color graphics camera. Five major forest cover type conditions as annotated on Figure 2.6a only, are: A - bare soil which results from clearing forest vegetation for expanding a ponderosa pine plantation; B - high density mixed conifer (greater than 70 percent crown closure) including a variable species mix of ponderosa pine, sugar pine, white fir, and incense cedar; C - low density Jeffrey pine (less than 20 percent crown closure) on soil with low fertility; D - wet (high water table) and dry meadows at/or shortly after peak forage production; and E - broadleaf riparian forest vegetation (greater than 90 percent crown closure) including cottonwood, willow, and big leaf maple.

The improvement in the spatial resolution of the TM sensor over the MSS sensor can be seen by comparing Figures 2.6a and 2.6b, respectively, on which the same surface is represented. Significant detail on the TM image which cannot be consistently identified on the MSS image includes the location of road and stream networks, area and shape of small forest clearings, definition of vegetation boundaries, and discrimination of small homogeneous vegetation units within a matrix of complex cover types. With the clear definition of significant topographical features and small stream networks on the TM imagery, these image products are more suited for field navigation than the MSS film products. A composite which includes TM 5 (Figure 2.6c) would be best for road and stream network delineation, because the contrast between a dry road surface and moist vegetation surface and between a stream course and adjacent features is higher on this band than on a composite containing TM 4 without TM 5 (Figure 2.6a).

The improvement in spectral resolution of the TM sensor over the MSS sensor and the improvements which can be expected by using various TM spectral band combinations for stratifying major forest conditions can be seen in Figure 2.6. By visually comparing the forest cover type conditions annotated on Figure 2.6a with their appearance on the other image types, two significant conclusions can be made that also have been supported by the graphic display in Figure 2.5:

(1) Image composites which contain TM 5 in combination with two other visible bands, i.e. TM 2 and 3, are inferior to those which contain TM 4 for discriminating between important cover types which display high infrared reflectivity (shrubs, hardwoods, and meadows). This becomes evident after comparing the detail in Figure 2.6a (TM 4, 3 and 2), in which a variety of these cover types are easily discriminated due to their bright red tones to the lack of detail in Figure 2.6c (TM 5, 3 and 2) for these same cover types. As examples, the riparian forest area at "E", a network of small wet meadows near the upper left center portion of the images, and the dense shrub field in the upper right center of the images. The apparent lack of detail of these examples in Figure 2.6c is due to the high level of absorption of the solar radiation in TM 5 (1.57 - 1.78μm) by vegetation having a high leaf water content. The optimum combination of reflective TM bands for discriminating most forest cover types in this region of

California, would be TM 5, 4 and 2, TM 5, 4 and 3, or TM 5, 4 and 1. A composite containing any of these band combinations would be able to exploit both the reflective and absorptive properties of the diverse cover types.

(2) <u>Discrimination of meadow and hardwood cover types is enhanced with the addition of TM 5 spectral data.</u> These two cover types are spectrally confused when using TM 1-4 alone. They can be separated spectrally, however, because of the change in magnitude of the mean spectral response of both types in TM 5. The discrimination of granitic rock from bare soil is also enhanced when TM 5 spectral data are used.

Based on analysis of these and other images and spectral statistics, we do see an improvement in our ability to discriminate major forest cover types with the addition of the middle-reflectance infrared data which cannot already be done using the other reflectance bands. Future evaluation of the TM spectral bands in other forest environments should be carried out, however, to see if these conclusions still hold true.

(3) The discrimination of thermal differences of forested cover types is important for assessing regional soil and plant moisture conditions and for mapping those sites suitable for reforestation programs. The data being acquired by the broad thermal infrared band (TM 6: $10.42-11.66\mu m$) are the highest quality thermal data, both in terms of spectral and spatial resolution, that have been acquired by an Earth-orbiting land remote sensing satellite. The TM6 data have added a new dimension to the characterization of forest resources on image products. This is evident on the image shown in Figure 2.6d, in the thermal band (TM 6) has been composited with two reflective bands (TM 4 and 3). Although the coarse spatial resolution of TM 6 (120m versus 30m) precludes the discrimination of some features, such as small forest clearings and small water bodies, the large scale radiant temperature differences between major forest cover types is obvious. In this composite, those cover types which are highly infrared reflective in TM 4 and also have a high moisture status (low radiant temperature) are shown in tones of light green (wet meadows, riparian forest, and dense shrub fields). At the other extreme, those cover types with relatively low infrared reflectance in TM 4 and which also have a low moisture status (high radiant temperature) are shown in light red to pink (bare soil, dry meadows, and low density coniferous forests). Intermediate cover types, such as high density conifer forest on northerly aspects, are shown in dark green tones. <u>By the combining of the thermal band of low spatial resolution with the reflective bands of high spatial resolution, a unique and useful image product is produced. This "thermal" composite should include, however, TM 4 as one of the reflective bands and not TM 5 due to the high positive correlation between TM 5 and TM 6 (Table 2.6).</u> The most useful composite, therefore, would be one which combined TM 5 and 4 or TM 6 and 4 with one of the remaining visible bands (TM 3, 2, or 1).

22

The improved spatial resolution of TM data over MSS data in a forested environment is illustrated again by the two "color infrared" composites shown in Figure 2.7: (a) TM 4, 3, and 2; and (b) MSS 4, 2 and 1; accompanied by a color aerial oblique photography taken approximately one hour after the satellite overpass (c). Differences in spatial quality can be seen on those locations annotated on the images by an "A", and include road networks, shorelines, and key cover type boundaries. The effect of differential solar illumination as a result of variable relief is shown at "B", where a shrub community is present on both the north and south aspects. The north aspect shrubs appear dark intone on the TM and MSS composites due to reduced solar illumination, but light in tone on the oblique aerial photography due to both changes in sun angle and camera look angle. Important cover type boundaries at "C" can be discriminated on both composites, but they are better represented on the TM composite with its improved spatial resolution.

2.2.4 Summary

A systematic analysis of both image and numeric data for a forestry study site in the northern Sierra Nevada, California, has yielded the following significant results:

(1) TM 1, 5, and 6 contain saturated pixels (DN = 255) for areas of high albedo and high radiant temperature (bare soil and rock outcrops).

(2) TM 5, 7, and 6 show high positive interband correlation for vegetative surfaces due to the close association between absorption of solar radiation, as recorded on TM 5 and 7, and low radiant temperatures as a result of the cooling effects of evapotranspiration.

(3) The reflective bands (TM 1-5 and 7) do not permit a clear discrimination between true fir and mixed conifer stands on the basis of spectral characteristics alone. Use of both the thermal band (TM 6) and elevational gradients will improve this important discrimination.

(4) The high spectral, spatial, and radiometric resolution characteristics of the reflective TM bands permit the discrimination of forest stands with increasing densities of black oak, an indicator of highly productive forest sites.

(5) The improved spatial resolution of the TM sensor over that of the MSS allows for improved detection of road and stream networks, area and shape of small forest clearings, significant vegetation boundaries, and small homogeneous units within a matrix of complex cover types.

(6) Color composites which include TM 5 are best suited for discriminating those small forest features essential for both navigation and mapping purposes.

(7) Color composites which include TM 5 with other visible bands (TM 1-3) are inferior for discriminating important forest types which are highly reflective in the photographic infrared (TM 4).

(8) The optimum combination of reflective TM bands for discriminating forest cover types would be either TM 5, 4, and 2 or TM 5, 4, and 3 where both the reflective and absorptive properties of the diverse cover types can be exploited.

(9) The thermal infrared band (TM 6) contains valuable information on the radiant temperature differences between major cover types. Image products combining the spectral qualities of both the thermal and reflective bands and the spatial qualities of the reflective bands provide an optimum multispectral approach for the stratification of major forest cover types.

(10) Color image composites should not contain both TM 5 and TM 6 data due to the high positive correlation between them in this forested environment. If spatial resolution must be maximized, a composite containing TM 5 is recommended.

We conclude that the TM data are more than sufficient for providing much of the renewable resource information needed to meet inventory objectives. The TM data should prove extremely valuable for: (1) estimating the area of forest cover types, (2) updating land use survey maps, and (3) determining the size, shape, and location of individual forest clearings and water sources.

2.3. Evaluation of Thematic Mapper and Multispectral Scanner Composite Film Products for the Identification of Forest Cover Types

2.3.1 Introduction

Selective photo interpretation keys were constructed for three TM images and one MSS image covering the Plumas National Forest study site, California. The keys were developed for eight forest resource categories present in the study area. These keys were used to train a pool of photo interpreters prior to the testing of the different image products. The objective of these tests was to quantify the interpretability, for forest resource applications, of the three TM and one MSS image types.

2.3.2 The Plumas Forest Study Site

The Plumas Forest study site is located in Plumas County, California, approximately 265 km northeast of San Francisco. The area has been defined as part of the Sierra Nevada Mixed Conifer forest cover type by the Society of American Foresters (1980). It contains a diversity of forest cover types ranging from stands of red and white fir (Abies magnifica and A. concolor, respectively), to stands of mixed conifer dominated by ponderosa pine (Pinus ponderosa), Douglas-fir (Pseudotsuga menziesii), and/or sugar pine (P. lambertiana). Several other cover types are also prevalent and include low density Jeffrey pine (P. jeffreyi) stands on soils derived from ultramafic parent material; hardwood stands on wet and dry sites; dense shrub fields; wet and dry meadows; bare soil; granitic rock outcrops; and large water bodies.

2.3.3 Forest Resource Categories

Based on our 40 plus years of experience in this study area, eight forest resource categories were defined for which the selective photo interpretation keys were developed: high density conifer, low density conifer, hardwood-conifer mix, brush, wet and dry site hardwood, wet and dry meadow, bare ground, and rock (Table 2.8).

2.3.4 Thematic Mapper and Multispectral Scanner Images Used to Develop Keys

The prints of four color image types, derived from the coincident Landsat-4 digital MSS and TM data acquired on 12 August 1983, were used to develop the respective photo interpretation keys: (1) MSS bands 4, 2 and 1, (2) TM bands 4, 3 and 2, (3) TM bands 5, 3 and 2, and (4) TM bands 5, 4 and 3. For each image type, the three bands were color coded with red, green, and blue color filters, respectively, on the RSRP color monitor using a linear mapping function to simulate color infrared imagery on the RSRP's interactive image processing system. The resulting color images were copied to photographic film using a Matrix color graphics camera. The image specifications are summarized in Table 2.9.

The linear mapping functions that were applied to the four Landsat-4 images tested in this study were computed in the following manner. For each of the three bands to be used to make an image, a histogram was made using all of the digital values of scene brightness. Then, from this histogram,

the arithmetic mean ($\bar{x}$) and standard deviation (s) were calculated. The mapping interval was calculated for the respective color gun intensity levels (eight intensity levels for each color gun on the RSRP color monitor) as follows:

$$\text{Mapping interval} = \frac{\text{Range of values} + 1}{\text{Number of intensity levels}}$$

where : Range of Values = Maximum step value – Minimum step value;

Maximum step value = $\bar{x}$ + 3*s or the maximum observed step value, which ever is less; and

Minimum step value = $\bar{x}$ – 3*s or the minimum observed step value, whichever is greater.

Based on our interpretation experiences using both this enhancement procedure and other commonly used procedures, we prefer linear mapping enhancement. This product uses the available color levels efficiently, provides an image that closely simulates color infrared photography if the appropriate TM bands are used, and preserves many of the subtleties which occur in transition zones between and within wildland vegetation types. On the other hand, the alternative enhancements, which usually employ a histogram equalization algorithm, produce a high contrast image which may be more appropriate to geologic applications than to wildland vegetation analysis.

These four image types were selected from the many possible combination for the following reasons:

(1) MSS 4, 2 and 1 represented those images that have been available since the Landsat system was first launched in 1972. Through a comparison of the interpretability of this image to the following TM image types, we can make some inferences as to improvement in interpretability due to increased spatial and spectral resolution.

(2) TM 4, 3 and 2 had the same spectral sensitivity of MSS 4, 2 and 1 but with twice the spatial resolution.

(3) TM 5, 3 and 2 allowed us to compare the contribution made to an image if the middle reflectance infrared band (5) is used in place of the more traditional photographic infrared band (4).

(4) TM 5, 4 and 3 image appeared to contain more spectral information than the other TM image types based on the correlation matrix (Table 2.10).

26

2.3.5 Construction of the Photo Interpretation Keys

As defined by the Interservice Committee on Photo Interpretation Keys, a photo interpretation key is reference material designed to facilitate the rapid and accurate identification of features from an examination of their photographic images (Colwell, 1952). All photo interpretation keys are based upon diagnostic features (tone, color, texture, etc.) of the photo images of the objects or conditions to be identified. Depending on the manner in which the features are organized, two general types are recognized: (1) selective keys are arranged so that the photo interpreter simply selects the example corresponding to the image type he is trying to identify, and (2) elimination keys are arranged so that the photo interpreter follows a step-by-step process that leads to the elimination of all items except the one that he is trying to identify.

We developed only selective keys for the four Landsat-4 images for the following two reasons. Because of the relatively poor spatial resolution of satellite imagery, association and location are important image characteristics that must be used to identify a forest resource category; these two characteristics are presented most easily in selective keys. Secondly, the primary use of these keys was to familiarize skilled interpreters with the expected spectral responses of the resource categories on the different image types; again, the range of these responses can be presented most effectively in selective keys. These keys are presented in Figures 2.8 through 2.11.

2.3.6 Photo Interpretation Test Design

On each of the four image types, ten points were randomly allocated for each of the eight resource categories; this resulted in 80 test points per image type. The points were randomly located on each of the image types so that "interpreter-recall" would be minimized. Six interpreters were given the photo interpretation tests during one 60-minute session. (See Appendix A for the written training information given to the interpreters prior to testing.) After the testing was completed, the answer sheets were corrected and error matrices were constructed for each interpreter by image type. The resulting 24 matrices were then aggregated to produce an error matrix based on 480 interpreter responses for each image type.

Percent correct and percent commission errors were calculated for the error matrices as follows:

$$\text{Percent Correct} = \frac{\text{number of correct interpretation responses for a resource category}}{\text{total number of that resource category present}} \times 100$$

$$\text{Percent Commission Error} = \frac{\text{number of incorrect interpretation responses for a resource category}}{\text{total number of that resource category indicated by the interpreter(s)}} \times 100$$

27

In addition, for each matrix, a Kappa statistic was calculated as follows:

$$\text{Kappa} = \frac{\text{Diagonal}/N - (\text{row total} \times \text{column total})/N2}{1 \qquad - (\text{row total} \times \text{column total})/N2}$$

The Kappa statistic, which is a non-parametric measure of agreement between "ground truth" and photo interpretation labels, was used to rank the error matrices. The rankings were considered to be significantly different, or not, based on the values obtained through the use of the following relationship:

$$\text{Delta Kappa} = \frac{\text{ABS}[\text{Kappa}(i) - \text{Kappa}(j)]}{[\text{Variance Kappa}(i) + \text{Variance Kappa}(j)]^{\frac{1}{2}}}$$

If the Delta Kappa calculated between two matrices exceeded 1.96, we concluded that the Kappa values were significantly different at the 95 percent confidence level; if the calculated value was less than 1.96 we concluded that the Kappa values were not significantly different and that the image types represented by those matrices were equally interpretable. A complete description of the use of the Kappa statistic can be found in Congalton, et al., 1981. An example of an error matrix showing the percent correct, percent commission error, and Kappa calculations is given in Table 2.11.

It should be emphasized that the Kappa statistic cannot be used alone when evaluating the usefulness of different image types. Although it is a very helpful statistical tool for the ranking of image types, because it combines errors of omission and commission, it only measures the degree of association along the diagonal of the error matrix. It is very possible that other associations may exist in a matrix. For example, an interpreter may consistently misidentify the high density conifer category as the hardwood category, and vice versa. While such an error would result in a low Kappa value for the image type, the low value may be more a reflection on interpreter skill than it is on the information content of the image. In this simple case, the interpreter, with the available training aids and background experience, has confused the identifying characteristics of one "tree" category with those of another. Since he has done it consistently, it is merely a matter of retraining him. Therefore, in order to properly evaluate an image type for a given application, it is very important to look not only at the Kappa statistic, but also at the percent correct value and the percent commission errors, as given by the individual entries in the error matrices.

2.3.7 Results and Recommendations

Summaries of the test results are given in Tables 2.12a and 2.12b. Table 2.12a lists the ranked image types based on the Kappa statistics calculated for all eight resource categories, and Table 2.12b lists the ranked image types based on the individual resource categories. In these two tables, lines have been drawn to the right of the Kappa-ranked images. These lines represent those image types that are not significantly different

at the 95 percent confidence level. In Table 2.12b, the percent correct (%C) and percent commission error (%CE) associated with each image type have been included for each resource category. The two types of error matrices used to calculate the values given in Table 5a are presented in Tables 2.13 through 2.16: the first matrix (a) represents the interpretability of the image type based on the six interpreters, and the second matrix (b) represents the interpretability of the image type based on the "best" photo interpreter. The data from this second matrix should be used when evaluating the information content of the respective image types, for in operational applications of these images, the most able interpreter would be given the task of information extraction.

By examining the rankings in Table 2.12a, we concluded that image types TM 5, 4and 3, MSS 4, 2 and 1, and TM 5, 3 and 2 were statistically more interpretable than image type TM 5, 3 and 2. This indicates that any image type for this mixed conifer forest environment must contain a band that is sensitive to the photographic infrared region of the electromagnetic spectrum. Even the MSS image, despite its relatively poor spatial resolution, proved to be more interpretable than higher spatial resolution TM 5, 3 and 2 which had no band sensitive to the photographic infrared region. There may have been a relatively high negative correlation between these two infrared bands because, in this wildland environment, those vegetative features that exhibit high reflectance in the photographic infrared region due to a turgid leaf would exhibit a correspondingly low response in the water absorption band. Yet, there also appears to be some complementary aspects betweeen these two bands, in that all six interpreters commented that they considered TM 5, 4 and 3 to be the most easily interpretable of the four tested. This appears to be true, particularly when examining the brush fields and hardwood stands on this image type, for many more subtle differences are present than with conventional color infrared image TM 4, 3 and 2.

It is interesting to note that for four of the individual resource categories, the MSS 4, 2 and 1 image surpassed the higher spatial resolution TM images in absolute terms: low-density conifer, hardwood, bare soil, and rock. With the exception of the hardwood category, the other three categories represent non-vegetated categories which can be identified more easily on the poor spatial resolution image which covers a larger areal extent on the individual test images. Why the MSS image was more interpretable than the others for the hardwood category remains an enigma, for this category is typically located in riparian zones which should be more easily detected and identified on the higher resolution imagery.

29

2.4 References

Benson, A.S. and K.J. Dummer. 1983. Analysis of photo interpretation test results for seven aerospace image types of the San Juan National Forest, Colorado. Proc. 49th ASP Annual Meeting, Washington, D.C., pp 121-130.

California Department of Water Resources. 1982. Land use survey maps, Vernalis and Ripon quadrangles. Sacramento.

Colwell, R.N., 1952. Report on Commission VII (Photographic Interpretation) to the International Society of Photogrammetry. Photogrammetric Engineering, Vol. 18, no.3. pp 375-400.

Colwell, R.N., 1978. Interpretability of vegetation resources on various image types acquired from earth-orbiting spacecraft. J. Appl. Photogr. Eng. 4(3):107-117.

Congalton, R. G., R.A. Mead, R. G. Odewald, and J. Heinen, 1981. Analysis of forest classification accuracy. Cooperative Research Report between U.S. Forest Service and Virginia Polytechnic Institute and State University. No. 13-1134. Blacksburg, Virginia 24061.

DeGloria, S. D. 1984. Spectral variability of Landsat-4 Thematic Mapper and Multispectral Scanner data for selected crop and forest cover types. In Transactions on Geoscience and Remote Sensing, Vol. GE-22, No.3, May. p 308.

DeGloria, S. D., A. Benson, K. Dummer, and E. Fakhoury. 1984. Evaluation of Thematic Mapper data for mapping forest, agricultural and soil resources. J. Adv. Space Research, Pergammon Press (in press).

Dozier, J. 1983. TM data processing and correction. Landsat Sensor Design and Operation Course Syllabus. Lect. 12. Hughes Santa Barbara Research Center and U.C. Santa Barbaraa Extension. p. 12-20.

Hay, C. M. 1982. Remote sensing measurement technique for use in crop inventories. In Remote Sensing for Resource Management, C. Johannsen and J. Sanders, editors. Soil Conservation Society of America, Ankeny, Iowa. pp 420-433.

Masuoka, P. 1983. Personal communication.

NASA. 1983. Thematic mapper computer compatible tape (CCT-AT, CCT-PT). Interface Cont. Doc. LSD-ICD-105. Goddard Space Flight Center, Greenbelt, MD. Rev. A. December. 184 p.

NASA. 1984. A prospectus for Thematic Mapper research in the Earth sciences. NASA Tech. Memo #86149. Goddard Space Flight Center, Greenbelt MD. 66p.

Odenweller, J. B. and K. I. Johnson. 1982. Crop identification using Landsat temporal-spectral profiles. <u>Proc</u>. Symp. Machine Processing Remotely Sensed Data, Purdue University, West Lafayette, IN. pp 469-475.

Society of American Foresters, 1980. Forest cover types of the United States. F.H. Eyre, editor. Society of American Foresters, Washington, D.C.

Tucker, C.J. 1980. Remote sensing of leaf water content in the near infrared. Remote Sensing Environ. 10:23-32.

U.S. Department of Agriculture. 1983. California crop-weather. Statistical Reporting Service, Sacramento, vol. 2, nos. 21, 22, and 28, December 82 and January 83.
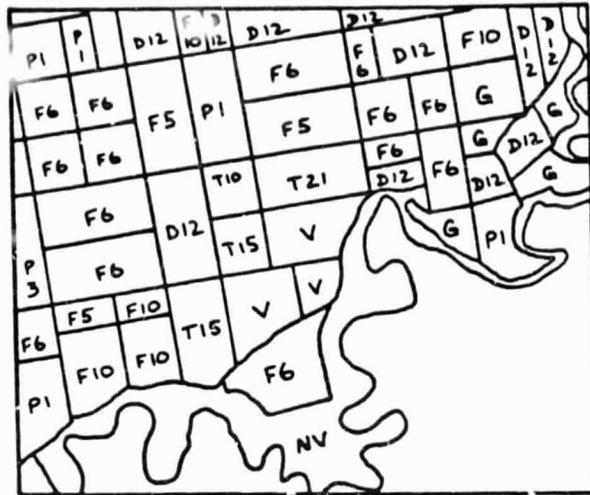
LEGEND

Symbol	Cover Type
D12	Almonds
F5	Sugar Beets
F6	Corn
F10	Beans
G	Grain and Hay Crops
P1	Alfalfa
P3	Mixed Pasture
T10	Onions and Garlic
T15	Tomatoes
T21	Peppers
V	Vineyard
NV	Native Vegetation

Figure 2.1 Portion of the agricultural land use map, map symbols, and legend used for the San Joaquin agricultural study site in California.

32
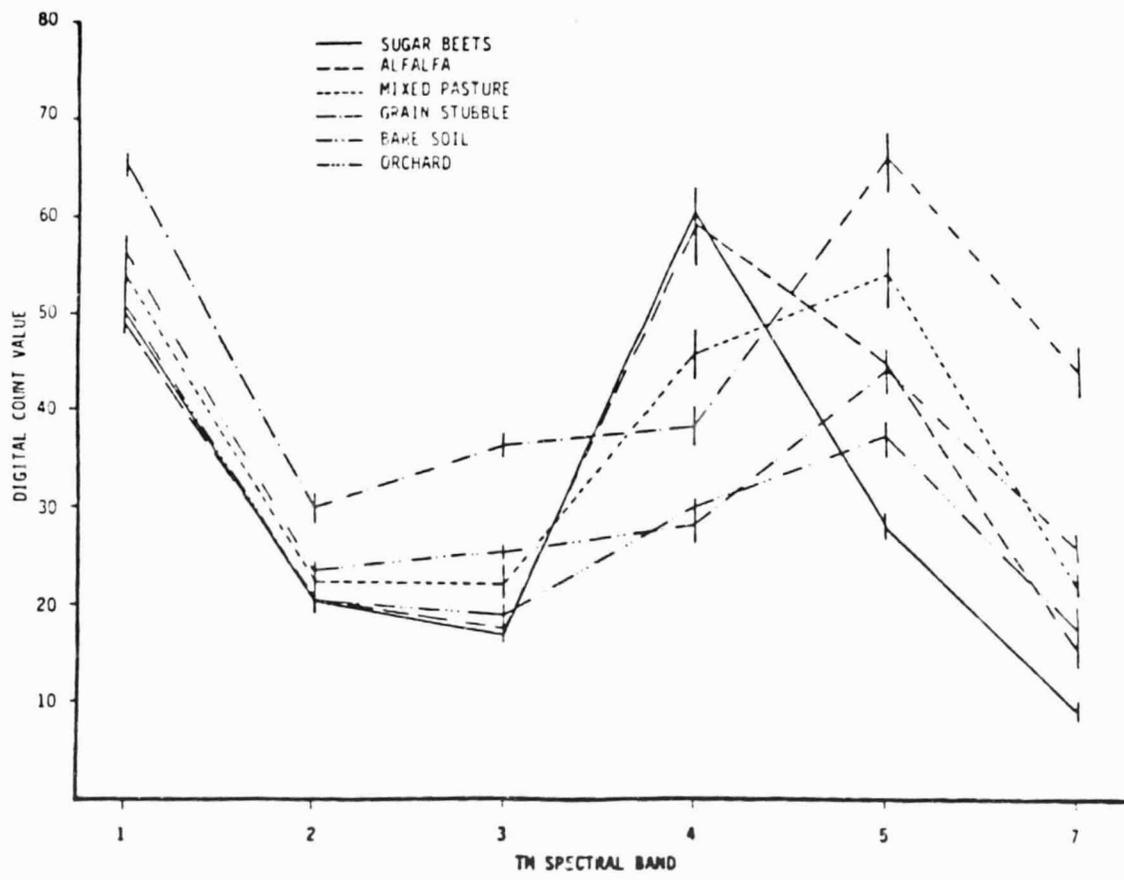
Figure 2.2 Plot of the Digital Number values for each of the reflective TM bands. Data values represent field means ($\pm$ 1 standard deviation as indicated by vertical lines) using one or more fields per cover condition. Values are tabulated in Table 2.2
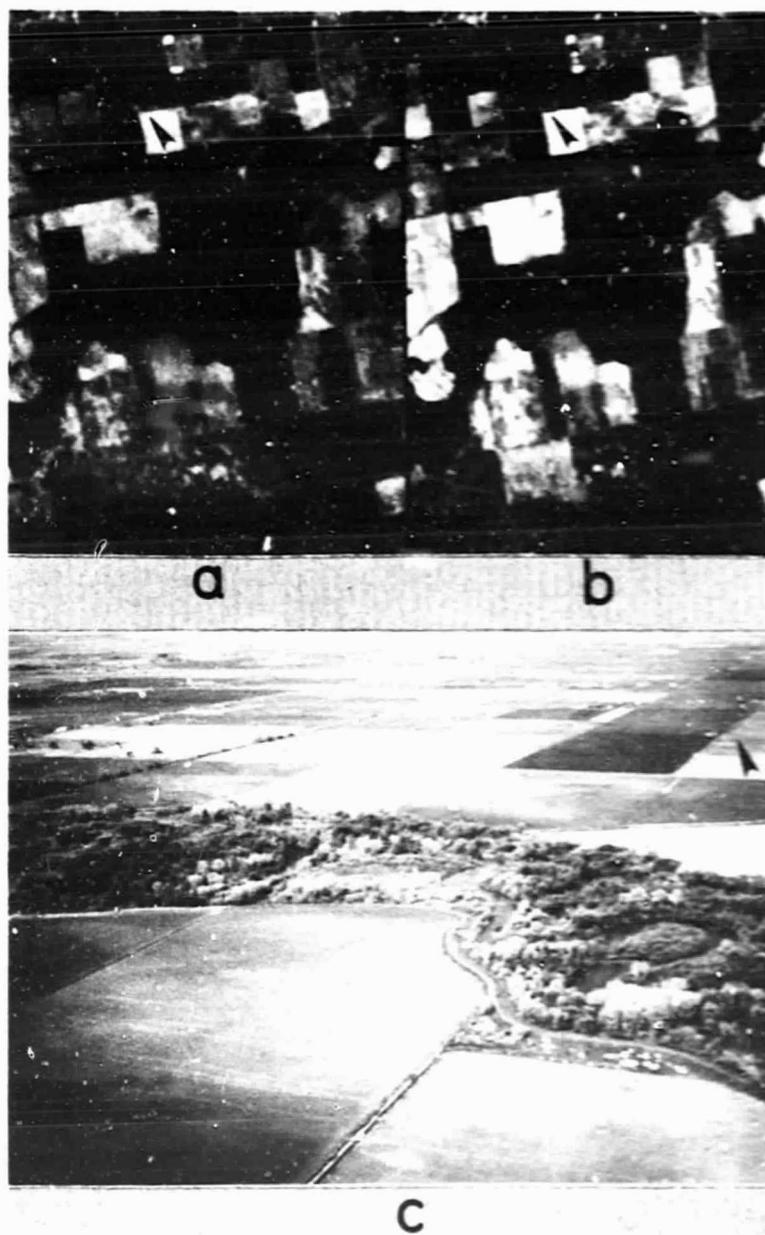
ORIGINAL PAGE
COLOR PHOTOGRAPH

Figure 2.3 Two color TM additive composites of a portion of the San Joaquin County agricultural study site: (a) TM 4, 3, and 2, and (b) TM 5, 3, and 2. These composites were made by color coding the bands with red, green, and blue filters, respectively. The color aerial oblique (c) was acquired six weeks after the Landsat-4 overpass.

34
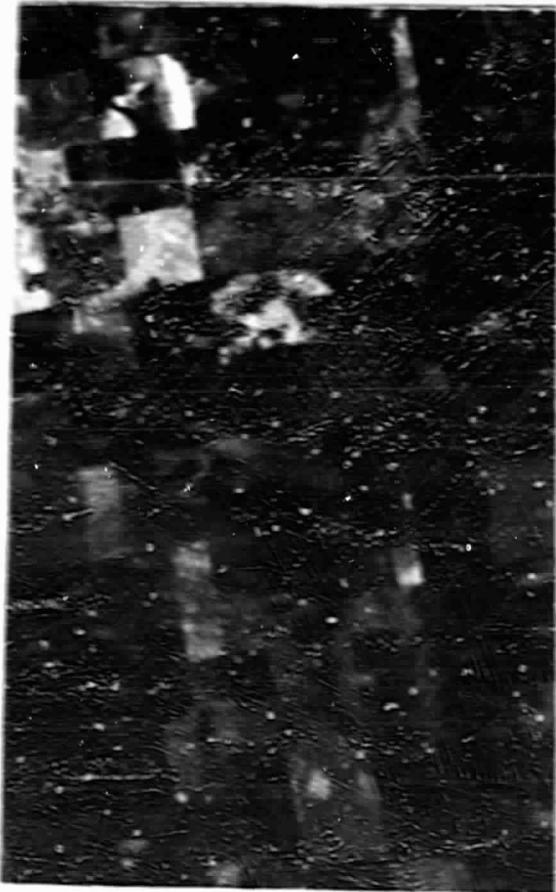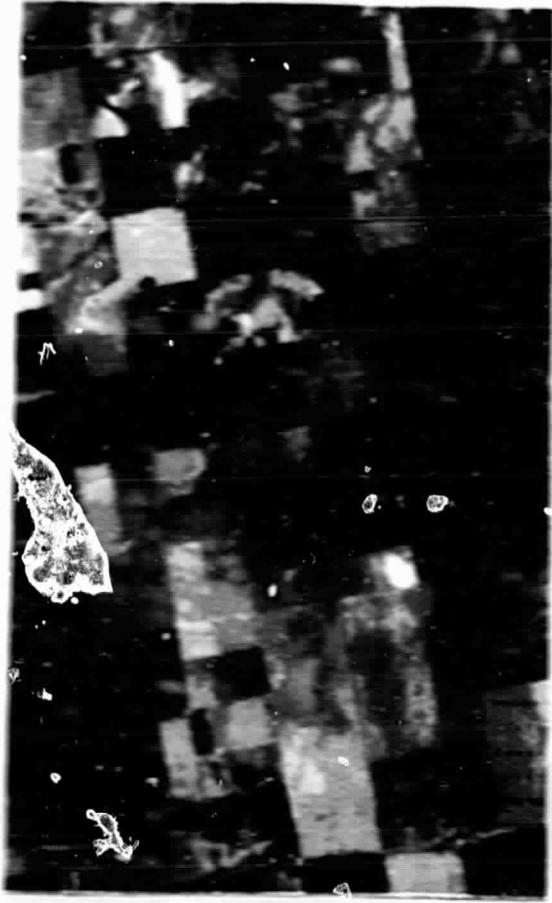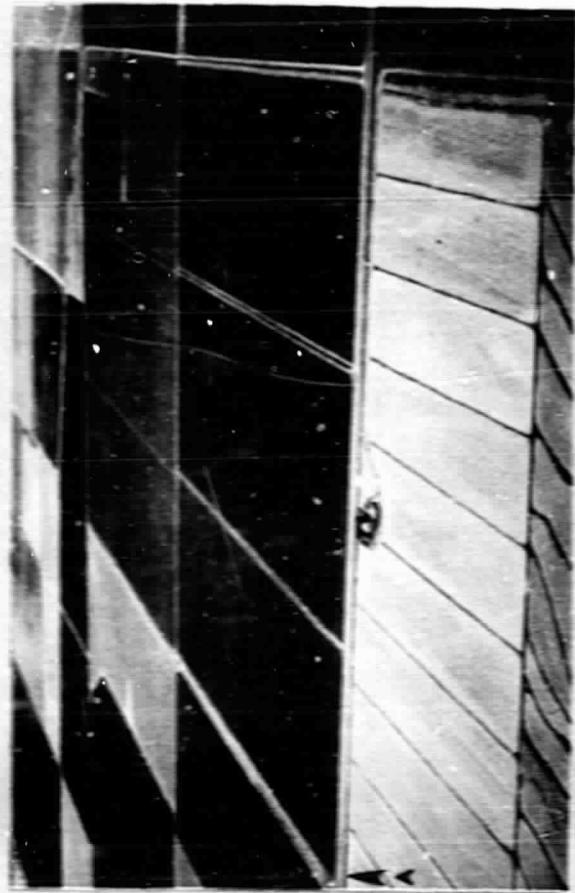
Figure 2.4 Three TM color additive composites for the Yolo County agricultural site: (a) TM 4, 3, and 1, (b) TM 5, 3 and 1, and (c) TM 5, 4, and 3. The color oblique aerial photograph (d) was taken coincident with this Landsat-4 overpass on 12 August 1983.

35
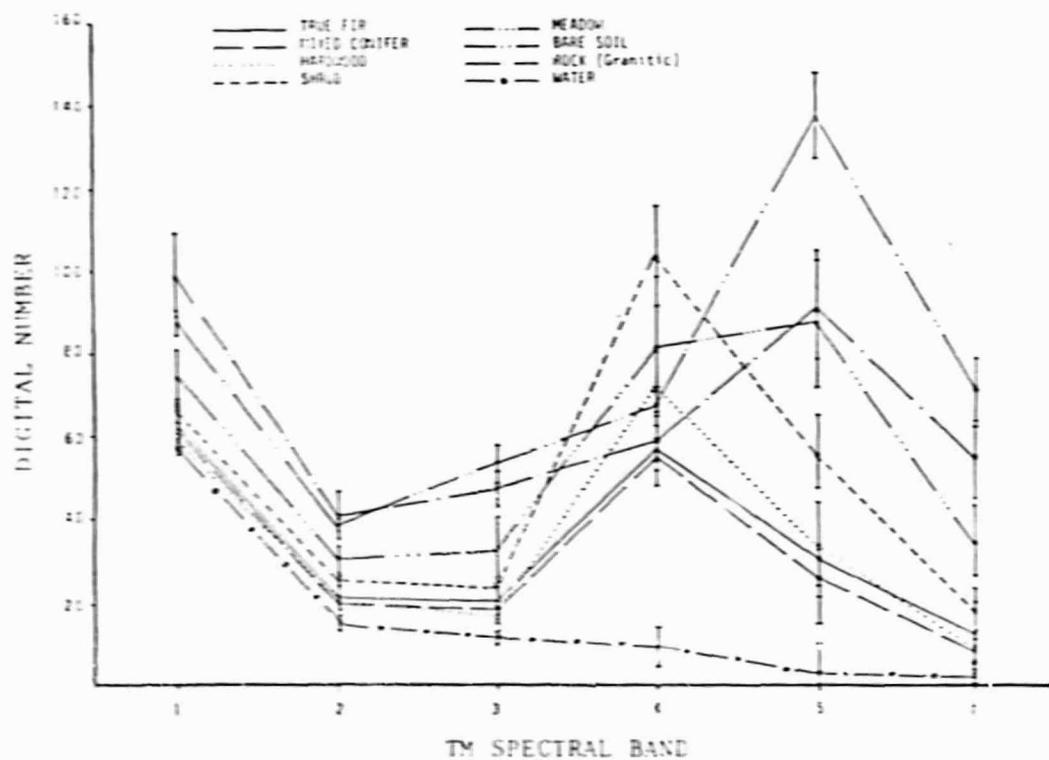
Figure 2.5 Plot of DN values (±1 standard deviation as indicated by the vertical lines and cross marks) for each of the reflective TM bands for selected forest cover types on the Plumas Forest study site.
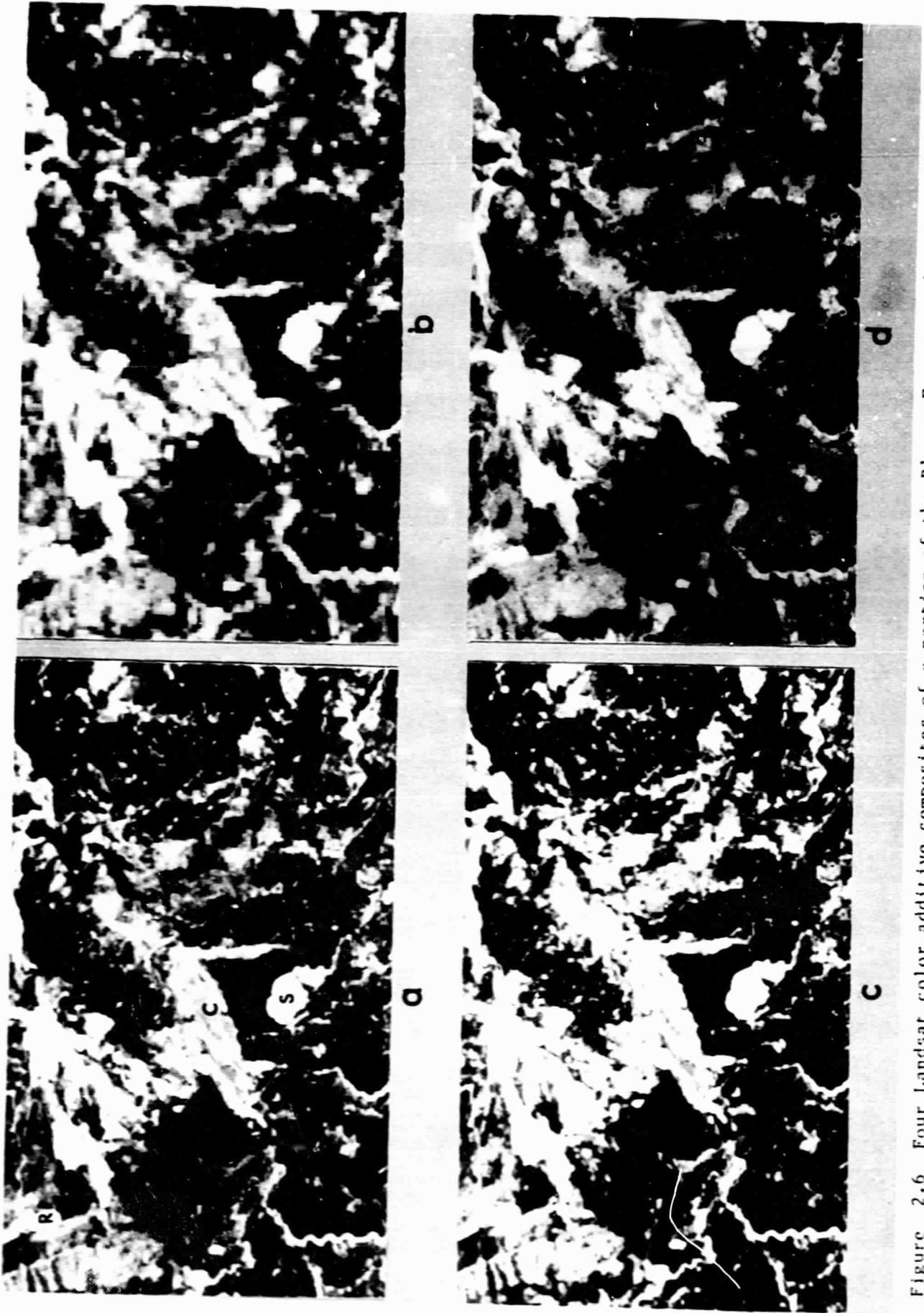
Figure 2.6 Four Landsat color additive composites of a portion of the Plumas Forest study site illustrating the differential reflectance of forest cover types as a result of plant morphological, physiological, and moisture conditions: (A) TM 4, 3, and 2; (B) MSS 4, 2, and 1; (C) TM 5, 3 and 2; and (D) TM 6, 4, and 3. Annotations: MC=mixed conifer, H=hardwood, B=shrub, C=grassland-meadow, S=bare soil, and R=rock.

Figure 2.7   Comparison of TM 4, 3, and 2 (a) and MSS 4, 2, and 1 (b) color composites of the Silver Lake area of our Plumas Forest study site. The accompanying color oblique aerial photograph (c) was acquired approximately one hour after the Landsat-4 overpass. Differences in spatial quality, solar illumination due to relief, and spectral characteristics of the different cover types are evident on these film products. See text for explanation of symbol annotations.

Figure 2.8   Selective photo interpretation key for MSS 4, 2 and 1 image: C
= high density conifer, L = low density conifer, H/C = hardwood
conifer mix, H = dry and wet site hardwood, B = brush, G =
grassland, S = bare soil, and R = rock.   Shown here is our
forestry study site in California's Plumas National Forest. The
largest lake shown here is Bucks Lake, and the sizeable lake in
the center of this image is Silver Lake (see Figure 2.7).

39
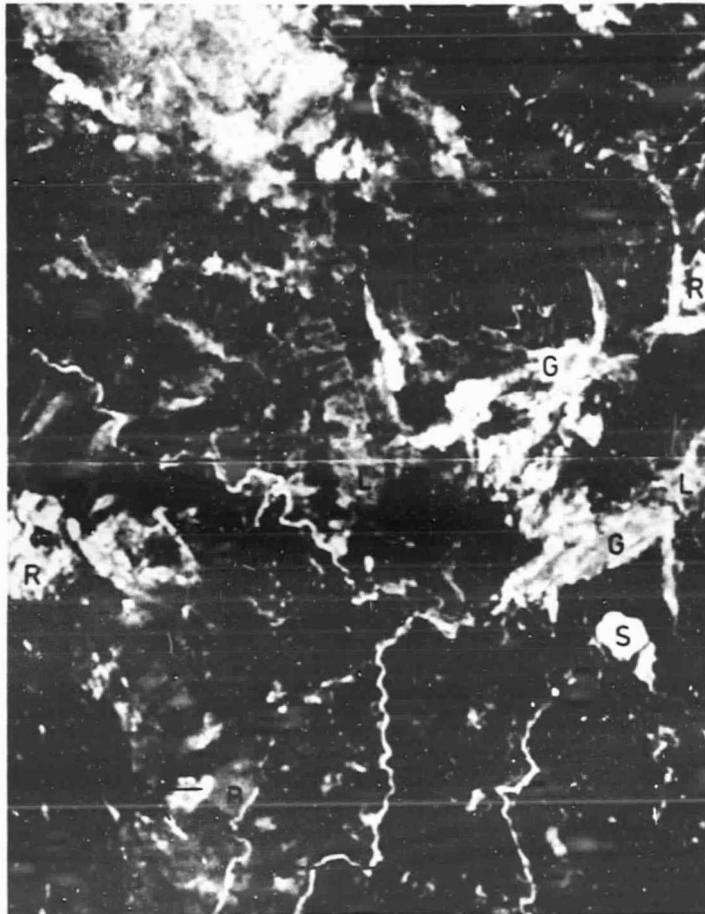
Figure 2.9   Selective photo interpretation key for TM 4, 3 and 2 image: C =
high density conifer,   L = low density conifer,   H/C = hardwood
conifer mix,   H = dry and wet site hardwood,   B =   brush,   G =
grassland,   S = bare soil,   and R = rock.   The area shown here
includes the eastern half of that shown in Figure 2.8.
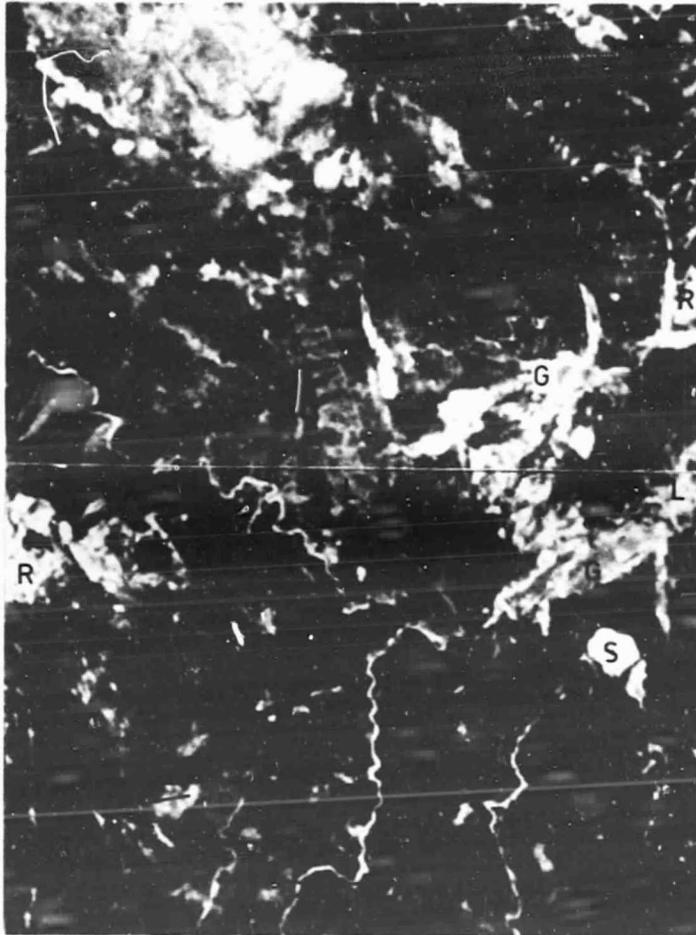
Figure 2.10   Selective   photo interpretation key of the same area   shown   in
Figure 2.9 for TM 5,   3 and 2 image: C = high density conifer,
L = low density conifer,   H/C = hardwood conifer mix,   H = dry
and wet site hardwood, B = brush, G = grassland, S = bare soil,
and R = rock.

41

Figure 2.11   Selective   photo   interpretation key of the same area shown   in
             Figure 2.9 for TM 5, 4 and 3 image: C = high density conifer,
             L = low density conifer, H/C = hardwood conifer mix, H = dry
             and wet site hardwood, B = brush, G = grassland, S = bare soil.
             and R = rock.

42

Table 2.1 Summary statistics for the San Joaquin agricultural study site.

TM BAND		$\bar{x}$	s	CV	Min	Max	Skewness (γ_1)	Kurtosis (γ_2)
1	DN:	54.0	4.3	8.0	44	149	2.8	21.0
	L:	3.4	0.2		2.7	9.4		
2	DN:	22.3	2.7	12.1	15	74	2.2	13.1
	L:	2.7	0.2		1.8	9.1		
3	DN:	21.8	4.3	19.7	12	90	1.9	8.1
	L:	2.1	0.4		1.1	8.5		
4	DN:	37.3	9.8	26.3	5	105	0.8	1.3
	L:	3.4	0.8		0.4	9.6		
5	DN:	45.0	9.6	21.3	1	255	0.2	5.3
	L:	0.5	0.1		-0.02*	3.2		
7	DN:	22.0	7.4	33.6	0	255	2.7	44.2
	L:	0.1	0.03		-0.02*	1.7		
6	DN:	92.0	1.2	1.3	77	102	-0.2	4.2

*Noise

DN = Digital number, counts
L = Spectral radiance, $mWcm^{-2}-sr^{-1}-\mu m^{-1}$
N = 255,328 pixels

Table 2.2 Correlation matrix for the seven TM bands for the San Joaquin study site.

TM BAND	1	2	3	4	5	7	6
1	1.00	0.89	0.91	-0.01	0.54	0.68	-0.15
2		1.00	0.91	0.15	0.62	0.69	-0.17
3			1.00	-0.07	0.66	0.79	-0.15
4				1.00	0.27	-0.08	0.02
5					1.00	0.84	0.03
7						1.00	-0.07
6							1.00

n = 255,328 pixels

43

Table 2.3 TM and MSS spectral statistics for the San Joaquin agricultural study site. Values tabulated are Digital Numbers (DN) extracted from the CCT-PT and CCT-AM tapes.

	SUGAR BEETS		ALFALFA		MIXED PASTURE		ORCHARD		VINEYARD		GRAIN STUBBLE		BARE SOIL		NATIVE VEGETATION	
	$\bar{x}$	CV	$\bar{x}$	CV	$\bar{x}$	CV	$\bar{x}$	CV	$\bar{x}$	CV	$\bar{x}$	CV	$\bar{x}$	CV	$\bar{x}$	CV
TM1	50.4	2.5	50.1	2.3	53.8	2.8	51.4	2.6	52.5	3.1	65.2	2.6	56.4	2.5	50.6	3.2
TM2	20.3	3.9	20.5	3.6	22.5	4.1	20.2	3.8	21.3	3.8	29.9	4.3	23.4	4.0	18.9	3.9
MSS 1	13.8	7.4	14.4	7.5	15.1	7.8	14.1	8.3	14.7	6.8	21.4	7.1	16.1	7.5	13.4	9.2
TM3	16.9	4.0	17.1	3.7	22.4	5.4	19.4	4.4	21.5	5.4	36.3	4.4	24.9	3.8	10.0	6.1
MSS 2	8.8	12.0	9.3	11.5	12.0	8.5	10.4	8.6	11.8	8.3	21.6	6.7	14.2	9.1	9.5	11.4
TM4	60.4	3.9	58.8	3.8	46.2	5.6	30.4	3.5	33.3	5.1	38.6	4.9	28.2	7.1	25.4	11.1
MSS 3	37.2	7.3	38.4	7.0	30.4	1.8	20.1	5.9	22.4	8.9	30.3	5.6	20.8	8.6	15.7	14.7
MSS 4	34.4	7.5	35.4	2.9	28.0	1.8	18.5	5.6	18.9	1.6	22.8	6.8	16.2	8.8	13.4	18.7
TM5	28.0	5.0	44.8	2.8	53.9	4.8	37.5	4.6	46.4	6.7	65.8	4.2	43.8	4.5	33.7	13.2
TM7	9.1	11.1	15.2	9.1	21.5	8.2	17.5	10.6	23.8	7.8	43.7	5.4	25.7	6.4	16.2	18.2
TM6	91.8	0.5	92.7	0.7	93.4	0.6	92.1	0.6	92.7	0.7	91.8	0.6	90.9	1.0	92.0	0.7

Table 2.4 Correlation matrix calculated from a subset of TM data acquired
on 12 August 1983 over the Sacramento Valley, California. A
portion of this subset is visually displayed in Figure 2.4.

TM Band	1	2	3	4	5	7	6
1	1.00	0.97	0.92	-0.13	0.85	0.84	0.71
2		1.00	0.95	-0.00	0.88	0.80	0.67
3			1.00	-0.03	0.89	0.79	0.70
4				1.00	-0.02	-0.36	-0.49
5					1.00	0.89	0.72
7						1.00	0.81
6							1.00

Table 2.5 Summary statistics for the Plumas National Forest study site.

TM BAND		$\bar{x}$	s	CV	Min	Max	Skewness (γ_1)	Kurtosis (γ_2)
1	DN:	66.1	8.1	12.3	41	255	3.1	25.0
	L:	4.1	0.5		2.5	16.1		
2	DN:	23.7	5.2	21.9	11	166	2.5	16.8
	L:	2.8	0.5		1.3	20.6		
3	DN:	23.2	8.2	35.3	7	214	2.4	10.5
	L:	2.2	0.8		0.7	20.2		
4	DN:	64.1	16.1	25.1	0	172	0.2	1.8
	L:	5.8	1.4		-0.06*	15.8		
5	DN:	46.7	22.5	48.2	0	255	1.1	1.5
	L:	0.6	0.2		-0.03*	3.2		
7	DN:	18.0	11.7	65.0	0	164	1.5	3.1
	L:	0.1	0.6		-0.02*	1.1		
6	DN:	190.1	15.1	7.9	139	255	1.1	1.6

*Noise

DN = Digital number, counts
L = Spectral radiance, $\text{mWcm}^{-2}\text{-sr}^{-1}\text{-}\mu\text{m}^{-1}$
N = 1,400,000 pixels

Table 2.6 Correlation matrix for the seven TM bands calculated from a subset of TM data acquired on 12 August 1983, over the Meadow Valley area, Plumas National Forest. A portion of this subset is visually displayed in Figure 2.6.

TM BAND	1	2	3	4	5	7	6
1	1.00	0.94	0.96	0.15	0.74	0.84	0.30
2		1.00	0.96	0.24	0.80	0.87	0.77
3			1.00	0.18	0.81	0.90	0.81
4				1.00	0.43	0.18	0.06
5					1.00	0.93	0.85
7						1.00	0.97
6							1.00

n = 40,000 pixels

Table 2.7 TM and MSS spectral statistics for individual forest cover types. Values tabulated represent average digital numbers ($\bar{x}$) and coefficients of variation (CV) calculated from the mean and standard deviation of each cover type.

SPECTRAL BAND	TRUE FIR		MIXED CONIFER		HARDWOOD		SHRUB		MEADOW		BARE SOIL		ROCK		WATER	
	$\bar{x}$	CV	$\bar{x}$	CV	$\bar{x}$	CV	$\bar{x}$	CV	$\bar{x}$	CV	$\bar{x}$	CV	$\bar{x}$	CV	$\bar{x}$	CV
TM1	63.5	9.3	61.6	3.1	60.0	1.7	65.8	5.1	75.1	10.5	87.6	3.1	99.3	11.4	58.3	3.9
TM2	22.2	14.4	20.0	4.8	20.3	2.8	25.7	8.2	30.5	14.6	38.5	5.7	41.8	14.3	16.4	8.4
MSS1	15.0	13.3	13.4	9.6	14.3	4.0	16.8	10.0	20.7	13.5	25.3	7.7	29.0	10.3	11.4	8.7
TM3	20.7	19.8	17.5	9.2	17.3	6.7	23.8	14.4	33.0	24.8	53.9	9.8	47.7	8.1	12.4	15.0
MSS2	13.1	20.9	10.3	15.5	11.0	20.0	14.8	15.7	21.8	24.1	36.9	11.8	33.5	12.5	6.4	20.9
TM4	57.2	14.9	56.2	6.8	73.3	11.0	103.7	11.1	82.4	15.7	67.9	6.3	58.6	6.4	9.5	49.7
MSS4	35.6	12.2	34.9	5.4	45.7	9.8	67.0	10.6	54.7	14.4	46.7	8.0	38.1	9.3	3.3	45.9
TM5	31.2	52.1	27.0	21.5	33.7	29.8	55.7	16.3	87.8	19.4	138.2	7.2	90.6	13.0	3.4	198.7
TM7	11.8	69.6	8.6	33.7	9.0	38.5	17.5	28.9	34.5	27.9	71.1	11.4	54.4	17.7	1.7	184.2
TM6	168.3	7.1	179.4	3.2	175.3	2.6	188.7	3.2	210.5	3.4	252.9	1.7	200.9	2.4	173.4	1.1

Table 2.8 Forest resource categories used to develop selective image interpretation keys for one multispectral scanner and three thematic mapper images for the Plumas Forest study site, California.

High Density Conifer - Mixed and pure conifer stands. Stands dense to moderately open: dense stands most often present on moist north and west slopes, and more open stands present on the other aspects. Rough image texture.

Low De. -ity Conifer - Open stands of mixed and pure conifer stands. Brush or herbaceous understory may be present. Open stands may be the result of poor site quality (serpentinitic soils) or selective logging practices. Rough to smooth image texture.

Hardwood/Conifer - Mixed hardwood and conifer tree stands. Stands dense to moderately dense. Stands most often present on drier sites (south slopes) or may develop ten to twenty years after selective logging. Rough image texture.

Hardwood - Pure and mixed evergreen and deciduous hardwood stands. Stands may be on dry or wet sites: dry sites present on south aspects; wet sites located in water courses or interspersed throughout dense stands of conifer at the higher elevations. Based on area covered, this class is very small. Moderately rough image texture.

Brush - Dense stands of evergreen or deciduous shrubs. Usually present in large contiguous stands on dry sites. Evidence of vegetation modification may be present. Smooth image texture.

Grassland - Vegetative cover predominantly annual or perennial grasslands. Sites may include dry or wet habitats or improved pasture lands. Smooth image texture.

Bare Ground - Non vegetated soil. Includes recently disturbed sites, logging roads (unpaved), and extremely poor habitat. Not to be confused with the rock category (below). Smooth image texture.

Rock - Areas of bare rock which are devoid of vegetation to include rock outcrops, talus slopes, and gravel bars. Smooth to moderately smooth image texture.

Table 2.9 Summary of Landsat-4 image specifications for one MSS image and three TM images for which photo interpretation tests were performed on the Plumas Forest study site.

SENSOR	IMAGE BANDS	COLOR CODE	BAND MEAN	STANDARD DEVIATION	BAND DISPLAY RANGE	IMAGE SCALE
MSS	4	RED	40.93	10.08	10-71	
	2	GREEN	15.58	14.63	2-31	1:120,000
	1	BLUE	13.42	15.87	7-26	
TM	4	RED	64.07	15.93	16-111	
	3	GREEN	23.14	8.21	7-47	1: 75,000
	2	BLUE	23.69	5.22	11-39	
TM	5	RED	46.30	21.65	0-111	
	3	GREEN	23.14	8.21	7-47	1: 75,000
	2	BLUE	23.69	5.22	11-39	
TM	5	RED	46.30	21.65	0-111	
	4	GREEN	64.07	15.93	16-111	1: 75,000
	3	BLUE	23.14	8.21	7-47	

Table 2.10 Correlation matrix of the seven thematic mapper bands for data extracted for the Plumas Forestry study site (DeGloria, 1984).

BAND	1	2	3	4	5	7	6
1	1.00	0.94	0.96	0.15	0.74	0.84	0.30
2		1.00	0.96	0.24	0.80	0.87	0.77
3			1.00	0.18	0.81	0.90	0.81
4				1.00	0.43	0.18	0.06
5					1.00	0.93	0.85
7						1.00	0.97
6							1.00

Table 2.11 Example of an error matrix constructed for the TM 4, 3, and 2 image, and of the manner in which percent correct, percent commission error, and Kappa are calculated.

| | GROUND TRUTH* | | | | | | | | | PERCENT COMMISSION |
P.I. RESULTS	1	2	3	4	5	6	7	8	TOTAL	ERROR
1 Hi Conifer	51	0	6	3	0	0	0	0	60	15
2 Lo Conifer	3	30	7	2	13	0	0	0	55	45
3 Hard/Con	3	0	34	12	6	0	0	0	55	38
4 Hardwood	1	0	7	18	2	2	0	0	30	40
5 Brush	2	5	5	23	30	1	0	1	67	55
6 Grassland	0	21	0	2	9	36	5	14	87	59
7 Bare Soil	0	3	0	0	0	20	36	24	83	57
8 Rock	0	1	1	0	0	1	19	21	43	51
TOTAL	60	60	60	60	60	60	60	60	480	
PERCENT CORRECT	85	50	57	30	50	60	60	35		

TOTAL CORRECT = 256   PERCENT CORRECT = 53.3
TOTAL ERROR   = 224   PERCENT ERROR   = 46.7

ESTIMATED KAPPA = 0.671   ESTIMATED ST.DIV. = 0.026
95% CONFIDENCE LIMIT 0.416 TO 0.517
99% CONFIDENCE LIMIT 0.400 TO 0.533

$$\text{Percent Correct} = \frac{51+30+...+36+21}{480} \times 100 = 53.3\%$$

$$\text{Percent Commission Error} \atop \text{HI CONIFER} = \frac{0+6+3+0+0+0+0}{60} \times 100 = 15.0\%$$

$$\text{Kappa} = \frac{\frac{51+30+...+36+21}{480} - \frac{(60\times60)+(60\times55)+...+(60\times83)+(60\times43)}{480\times480}}{1 - \frac{(60\times60)+(60\times65)+...+(60\times83)+(60\times43)}{480\times480}} = .671$$

---

* Based on both field data and interpretation of large-scale color and color infrared aerial photographs.

Table 2.12 Summary of image rankings from the Plumas Forest study site Landsat-4 image interpretation tests.

a. Based on <u>eight</u> resource categories: (1) high density conifer, (2) low density conifer, (3) hardwood/conifer, (4) hardwood, (5) brush, (6) grassland, (7) bare soil, and (8) rock.

RANK	IMAGE TYPE	KAPPA AND SIGNIFICANCE*	PERCENT CORRECT -RANGE-	PERCENT CORRECT -MEAN-	PERCENT ERROR -MEAN-
1	TM 5,4&3	.502	41.3 - 73.8	56.5	43.5
2	MSS 4,2&1	.486	46.3 - 70.0	55.0	45.0
3	TM 4,3&2	.467	30.0 - 71.3	53.3	46.7
4	TM 5,3&2	.388	23.8 - 63.0	46.5	53.5

==
* The lines to the right of the ranked Kappa values indicate those image types that are not significantly different at the .95 confidence level.

Table 2.12 (concluded)

b.   Based on individual resource categories.

RANK	CATEGORY	%C/%CE*	RANK	CATEGORY	%C/%CE*
	HIGH-DENSITY CONIFER			LOW-DENSITY CONIFER	
1	TM 4,3&2	85/15	1	MSS 4,2&1	90/39
2	TM 5,3&2	85/33	2	TM 5,4&3	77/44
3	TM 5,4&3	62/10	3	TM 4,3&2	50/45
4	MSS 4,2&1	80/38	4	TM 5,3&2	50/71
	HARDWOOD/ CONIFER			HARDWOOD	
1	TM 4,3&2	57/38	1	MSS 4,2&1	40/41
2	TM 5,4&3	52/34	2	TM 5,4&3	37/39
3	MSS 4,2&1	47/63	3	TM 4,3&2	30/40
4	TM 5,3&2	23/62	4	TM 5,3&2	2C/40
	BRUSH			GRASSLAND	
1	TM 5,4&3	80/49	1	TM 5,3&2	67/49
2	TM 4,3&2	50/55	2	TM 5,4&3	63/53
3	MSS 4,2&1	20/40	3	MSS 4,2&1	43/47
4	TM 5,3&2	22/72	4	TM 4,3&2	60/59
	BARE SOIL			ROCK	
1	MSS 4,2&1	78/50	1	MSS 4,2&1	42/29
2	TM 5,3&2	67/53	2	TM 5,3&2	38/36
3	TM 5,4&3	62/53	3	TM 4,3&2	35/51
4	TM 4,3&2	60/57	4	TM 5,4&3	20/37

==
* %C/%CE = Percent Correct/Percent Comission error for individual resource
categories.   The lines to the right of the ranked values indicate those
image types that are not significantly different at the 95 percent
confidence level based on the Kappa statistic.

Table 2.13 Error matrices from the Plumas Forest study site for Image 1: MSS 4, 2 and 1.

(a) BASED ON A POOL OF SIX INTERPRETERS

P.I. RESULTS	GROUND TRUTH								TOTAL	PERCENT COMMISSION ERROR
	1	2	3	4	5	6	7	8		
1 Hi Conifer	48	0	24	4	2	0	0	0	78	38
2 Lo Conifer	2	54	4	0	25	2	0	1	88	39
3 Hard/Con	8	0	28	30	9	0	0	0	75	63
4 Hardwood	2	0	3	24	5	5	2	0	41	41
5 Brush	0	1	1	2	12	1	0	3	20	40
6 Grassland	0	5	0	0	5	26	3	10	49	47
7 Bare Soil	0	0	0	0	2	24	47	21	94	50
8 Rock	0	0	0	0	0	2	8	25	35	29
TOTAL	60	60	60	60	60	60	60	60	480	

PERCENT CORRECT 80 90 47 40 20 43 78 42

    TOTAL CORRECT = 264   PERCENT CORRECT = 55.0
    TOTAL ERROR   = 216   PERCENT ERROR   = 45.0

    ESTIMATED KAPPA = 0.486   ESTIMATED ST.DIV. = 0.026
    95% CONFIDENCE LIMIT: 0.436 TO 0.536
    99% CONFIDENCE LIMIT: 0.420 TO 0.552

(b) BASED ON THE "BEST" INTERPRETER

P.I. RESULTS	GROUND TRUTH								TOTAL	PERCENT COMMISSION ERROR
	1	2	3	4	5	6	7	8		
1 Hi Conifer	8	0	5	0	0	0	0	0	13	38
2 Lo Conifer	0	10	1	0	3	0	0	0	14	29
3 Hard/Con	2	0	4	2	1	0	0	0	9	56
4 Hardwood	0	0	0	8	1	0	0	0	9	11
5 Brush	0	0	0	0	4	0	0	0	4	0
6 Grassland	0	0	0	0	0	7	0	1	8	13
7 Bare Soil	0	0	0	0	1	3	10	4	18	44
8 Rock	0	0	0	0	0	0	0	5	5	0
TOTAL	10	10	10	10	10	10	10	10	80	

PERCENT CORRECT 80 100 40 80 40 70 100 50

    TOTAL CORRECT = 56   PERCENT CORRECT = 70.0
    TOTAL ERROR   = 24   PERCENT ERROR   = 30.0

    ESTIMATED KAPPA =0.657   ESTIMATED ST.DIV. = 0.058
    95% CONFIDENCE LIMIT: 0.544 TO 0.770
    99% CONFIDENCE LIMIT: 0.508 TO 0.806

Table 2.14 Error matrices from the Plumas Forest study site for Image 2: TM 4, 3, and 2.

(a) BASED ON A POOL OF SIX INTERPRETERS

				GROUND	TRUTH						PERCENT COMMISSION
P.I. RESULTS	1	2	3	4	5	6	7	8	TOTAL	ERROR
1 Hi Conifer	51	0	6	3	0	0	0	0	60	15
2 Lo Conifer	3	30	7	2	13	0	0	0	55	45
3 Hard/Con	3	0	34	12	6	0	0	0	55	38
4 Hardwood	1	0	7	18	2	2	0	0	30	40
5 Brush	2	5	5	23	30	1	0	1	67	55
6 Grassland	0	21	0	2	9	36	5	14	87	59
7 Bare Soil	0	3	0	0	0	20	36	24	83	57
8 Rock	0	1	1	0	0	1	19	21	43	51
TOTAL	60	60	60	60	60	60	60	60	480	

PERCENT CORRECT 85 50 57 30 50 60 60 35

TOTAL CORRECT = 256   PERCENT CORRECT = 53.3
TOTAL ERROR = 224   PERCENT ERROR = 46.7

ESTIMATED KAPPA = 0.671   ESTIMATED ST.DIV. = 0.026
95% CONFIDENCE LIMIT: 0.416 TO 0.517
99% CONFIDENCE LIMIT: 0.400 TO 0.533

(b) BASED ON THE "BEST" INTERPRETER

				GROUND	TRUTH						PERCENT COMMISSION
P.I. RESULTS	1	2	3	4	5	6	7	8	TOTAL	ERROR
1 Hi Conifer	9	0	0	0	0	0	0	0	9	0
2 Lo Conifer	0	8	0	0	2	0	0	0	10	20
3 Hard/Con	1	0	9	0	2	0	0	0	12	25
4 Hardwood	0	0	1	3	0	0	0	0	4	25
5 Brush	0	0	0	7	6	0	0	0	13	54
6 Grassland	0	1	0	0	0	6	2	0	9	33
7 Bare Soil	0	1	0	0	0	3	7	1	12	42
8 Rock	0	0	0	0	0	1	1	9	11	18
TOTAL	10	10	10	10	10	10	10	10	80	

PERCENT CORRECT 90 80 90 30 60 60 70 90

TOTAL CORRECT = 57   PERCENT CORRECT = 71.3
TOTAL ERROR = 23   PERCENT ERROR = 28.8

ESTIMATED KAPPA = 0.467   ESTIMATED ST.DIV. = 0.058
95% CONFIDENCE LIMIT: 0.559 TO 0.784
99% CONFIDENCE LIMIT: 0.523 TO 0.820

Table 2.15 Error matrices from the Plumas Forest study site for Image 3: TM 5, 3, and 2.

(a) BASED ON A POOL OF SIX INTERPRETERS

		GROUND TRUTH									PERCENT COMMISSION
P.I. RESULTS	1	2	3	4	5	6	7	8	TOTAL	ERROR
1 Hi Conifer	51	0	19	3	1	0	2	0	76	33
2 Lo Conifer	4	30	23	9	35	0	1	0	102	71
3 Hard/Con	1	1	14	14	7	0	0	0	37	62
4 Hardwood	4	0	0	12	2	2	0	0	20	40
5 Brush	0	6	4	22	13	1	0	0	46	72
6 Grassland	0	22	0	0	0	40	7	7	78	49
7 Bare Soil	0	1	0	0	0	14	40	30	85	53
8 Rock	0	0	0	0	0	3	10	23	36	36
TOTAL	60	60	60	60	60	60	60	60	480	

PERCENT CORRECT 85 50 23 20 22 67 67 38

TOTAL CORRECT	= 223	PERCENT CORRECT	= 46.5
TOTAL ERROR	= 257	PERCENT ERROR	= 53.5

ESTIMATED KAPPA = 0.388   ESTIMATED ST.DIV. = 0.026
95% CONFIDENCE LIMIT: 0.338 TO 0.438
99% CONFIDENCE LIMIT: 0.322 TO 0.454

(b) BASED ON THE "BEST" INTERPRETER

		GROUND TRUTH									PERCENT COMMISSION
P.I. RESULTS	1	2	3	4	5	6	7	8	TOTAL	ERROR
1 Hi Conifer	10	0	7	0	1	0	0	0	18	44
2 Lo Conifer	0	9	2	0	6	0	0	0	17	47
3 Hard/Con	0	0	1	2	0	0	0	0	3	67
4 Hardwood	0	0	0	4	0	0	0	0	4	0
5 Brush	0	0	0	4	3	0	0	0	7	57
6 Grassland	0	0	0	0	0	8	0	0	8	0
7 Bare Soil	0	1	0	0	0	2	9	1	13	31
8 Rock	0	0	0	0	0	0	1	9	11	10
TOTAL	10	10	10	10	10	10	10	10	80	

PERCENT CORRECT 100 90 10 40 30 80 90 90

TOTAL CORRECT	= 53	PERCENT CORRECT	= 66.3
TOTAL ERROR	= 27	PERCENT ERROR	= 33.8

ESTIMATED KAPPA = 0.614   ESTIMATED ST.DIV. = 0.059
95% CONFIDENCE LIMIT: 0.499 TO 0.730
99% CONFIDENCE LIMIT: 0.462 TO 0.766

Table 2.16 Error matrices from the Plumas Forest study site for Image 4:
TM 5, 4, and 3.

(a) BASED ON A POOL OF SIX INTERPRETERS

P.I. RESULTS	GROUND TRUTH 1	2	3	4	5	6	7	8	TOTAL	PERCENT COMMISSION ERROR
1 Hi Conifer	37	3	0	0	0	0	1	0	41	10
2 Lo Conifer	12	46	13	2	3	1	5	0	82	44
3 Hard/Con	7	1	31	4	4	0	0	0	47	34
4 Hardwood	2	2	0	22	4	2	2	2	36	39
5 Brush	1	0	15	28	48	2	1	0	95	49
6 Grassland	1	6	1	1	0	38	11	23	81	53
7 Bare Soil	0	1	0	3	1	14	37	23	79	53
8 Rock	0	1	0	0	0	3	3	12	19	37
TOTAL	60	60	60	60	60	60	60	60	480	

PERCENT CORRECT 62 77 52 37 80 63 62 20

TOTAL CORRECT = 271   PERCENT CORRECT = 56.5
TOTAL ERROR   = 209   PERCENT ERROR   = 43.5

ESTIMATED KAPPA = 0.502   ESTIMATED ST.DIV. = 0.026
95% CONFIDENCE LIMIT: 0.~52 TO 0.553
99% CONFIDENCE LIMIT: 0.436 TO 0.568

(b) BASED ON THE "BEST" INTERPRETER

P.I. RESULTS	GROUND TRUTH 1	2	3	4	5	6	7	8	TOTAL	PERCENT COMMISSION ERROR
1 Hi Conifer	9	0	0	0	0	0	0	0	9	0
2 Lo Conifer	0	10	0	0	0	0	0	0	10	0
3 Hard/Con	1	0	10	0	1	0	0	0	12	17
4 Hardwood	0	0	0	2	1	0	0	1	4	50
5 Brush	0	0	0	7	7	0	0	0	14	50
6 Grassland	0	0	0	0	0	7	1	1	9	22
7 Bare Soil	0	0	0	1	1	2	9	3	16	44
8 Rock	0	0	0	0	0	1	0	5	6	17
TOTAL	10	10	10	10	10	10	10	10	80	

PERCENT CORRECT 90 100 100 20 70 70 90 50

TOTAL CORRECT = 59   PERCENT CORRECT = 73.8
TOTAL ERROR   = 21   PERCENT ERROR   = 26.3

ESTIMATED KAPPA = 0.700   ESTIMATED ST.DIV. = 0.056
95% CONFIDENCE LIMIT: 0.591 TO 0.809
99% CONFIDENCE LIMIT: 0.557 TO 0.843

APPENDIX A

GENERAL INFORMATION AND SPECIFIC INSTRUCTIONS
TO BE FOLLOWED IN THE TAKING OF THE
PLUMAS FOREST STUDY SITE IMAGE INTERPRETATION TESTS

OBJECTIVES

Determine the interpretability, for forest resource applications, of each of the following four Landsat-4 image types covering a forest study site in Plumas County, California obtained on 12 August 1983. (Band numbers have been color coded red, green, and blue, respectively.)

Image Type	Approximate Scale
Multispectral Scanner bands 4, 2 and 1	1:120,000
Thematic Mapper bands 4, 3 and 2	1: 75,000
Thematic Mapper bands 5, 3 and 2	1: 75,000
Thematic Mapper bands 5, 4 and 3	1: 75,000

===

MSS SPECTRAL CHARACTERISTICS

Band	Wavelength Range (micrometers)	Description
4	0.81-1.02	near (photographic) reflectance infrared
2	0.60-0.70	red
1	0.50-0.61	green

TM SPECTRAL CHARACTERISTICS

Band	Wavelength Range (micrometers)	Description
5	1.57-1.78	middle reflectance infrared
4	0.78-0.90	near (photographic) reflectance infrared
3	0.62-0.69	red
2	0.53-0.61	green

IMAGE CHARACTERISTICS

IMAGE 1
===============
(MSS4 PHOTO IR)
RED

```
                    RED
                    /\
                   /  \
                  /    \
         MAGENTA /      \ YELLOW
                /        \
               /          \
              /    CYAN    \
             /_____\
          BLUE            GREEN
     (MSS1 GREEN)      (MSS2 RED)
```

IMAGE 2
==============
(TM4 PHOTO IR)
RED
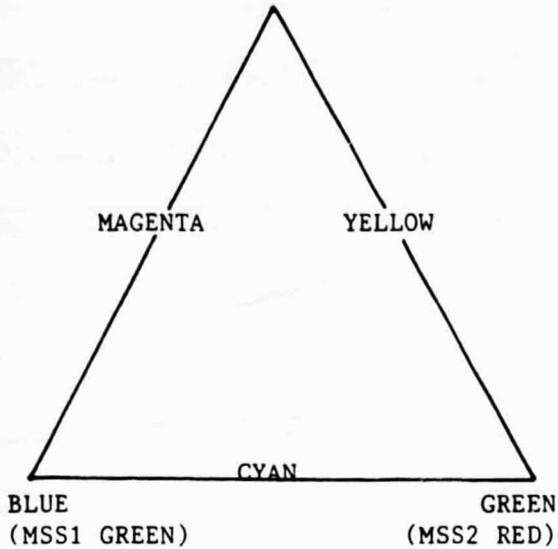
```
                    RED
                    /\
                   /  \
                  /    \
         MAGENTA /      \ YELLOW
                /        \
               /          \
              /    CYAN    \
             /_____\
          BLUE            GREEN
     (TM2 GREEN)       (TM3 RED)
```

IMAGE 3
===============
(TM5 MIDDLE IR)
RED
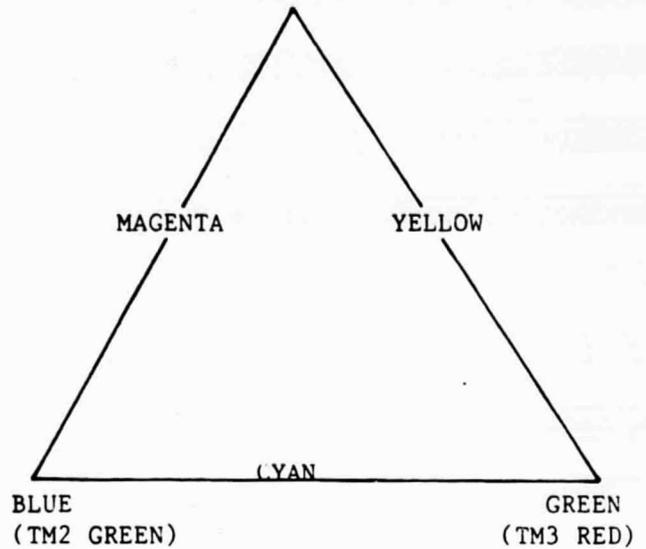
```
                    RED
                    /\
                   /  \
                  /    \
         MAGENTA /      \ YELLOW
                /        \
               /          \
              /    CYAN    \
             /_____\
          BLUE            GREEN
     (TM2 GREEN)       (TM3 RED)
```

IMAGE 4
================
(TM5 MIDDLE IR)
RED

```
                    RED
                    /\
                   /  \
                  /    \
         MAGENTA /      \ YELLOW
                /        \
               /          \
              /    CYAN    \
             /_____\
          BLUE            GREEN
     (TM3 RED)      (TM4 PHOTO IR)
```

58
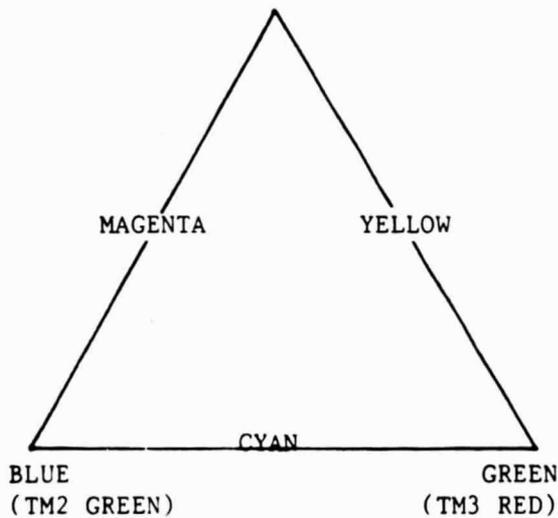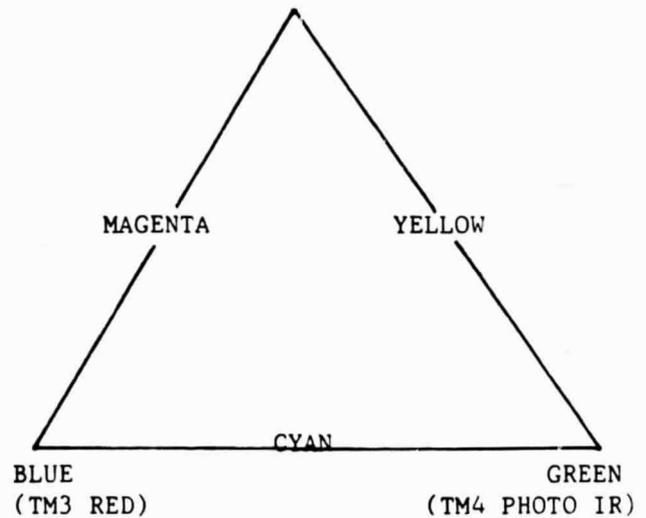
APPROACH

The interpretability of these image types will be judged on the correctness of the photo interpretation labels given to 80 points annotated on each image type. These labels must represent one of the following eight forest resource categories:

	Label	Resource Category		Label	Resource Category
1.	C	High density conifer	5.	B	Brush
2.	L	Low density conifer	6.	G	Grassland
3.	H/C	Hardwood/Conifer	7.	S	Bare soil
4.	H	Hardwood	8.	R	Rock

FILL OUT ALL INFORMATION ON YOUR ANSWER SHEETS AS IN THIS EXAMPLE

Interpreter name *Benson*   ID# _4_

Image Type *MSS421* Image # _1_ Time Start *1230* Time Finish *1320*
==
Class: 1. high density conifer – C 2. low density conifer – L
3. hardwood/conifer – H/C 4. hardwood – H 5. brush – B 6. grassland –G
7. bare soil –S 8. rock – R
==

Point	Class	Point	Class	Point	Class	Point	Class
1	H	21	S	41	S	61	B
2	H	22	C	42	S	62	B
3	H/C	23	C	43	L	63	B
4	G	24	H	44	L	64	B
5	R	25	L	45	HC	65	G
6	L	26	L	46	HC	66	C
7	L	27	H	47	R	67	C
8	C	28	H	48	R	68	C
9	C	29	G	49	S	69	L
10	C	30	G	50	R	70	L
11	C	31	HC	51	C	71	G
12	B	32	HC	52	C	72	G
13	H	33	HC	53	L	73	B
14	H	34	HC	54	L	74	R
15	B	35	L	55	L	75	R
16	G	36	H	56	L	76	S
17	G	37	H	57	H	77	S
18	HC	38	R	58	HC	78	S
19	R	39	S	59	H	79	L
20	R	40	S	60	HC	80	L

Chapter 3

EVALUATION OF THE GEOMETRIC QUALITY
OF THEMATIC MAPPER DATA

A. S. Benson, S. D. DeGloria, and K. J. Dummer
Remote Sensing Research Program

3.0 Introduction

A major aspect of our LIDQA investigation was to evaluate the geometric
qualities of TM photographic and digital products. This chapter summarizes
the results of three studies which examined the geometric properties of the
available data sets in terms of geodetic rectification (Sections 3.1 and
3.2) and of temporal registration (Section 3.3). Section 3.1 examines the
geometric properties of an enlarged interim TM black-and-white photographic
image covering a mixture of agriculture and wildland environments centered
on the northern Sacramento Valley. Section 3.2 examines the registration of
digital TM data to one 7½ minute quadrangle in the northern Sierra Nevada.
And Section 3.3 represents a small study in which two 1024x1024 pixel TM
subscenes from December 1982 and September 1983 were registered to our San
Felipe test site in the Coast Range of central California. In the first two
studies, the map standard again which the imagery was compared was regarded
as being correct.

3.1 Geometric Analysis of Thematic Mapper Film Product

3.1.1 Introduction

The geometric quality of an interim TM film product was evaluated by
measuring the coordinates of known features on both a TM photographic image
product and associated map products, and then relating these paired
observations using a standard linear least squares regression approach. The
major emphasis of our work was to analyze the TM film product from Landsat-4
that would be generally accessible to the user community.

At the time of our analysis, there are three types of Landsat-4 film
products being generated at the EROS Data Center (EDC) and the Goddard Space
Flight Center (GSFC): (1) standard multispectral scanner (MSS) film
products, (2) "interim" Thematic Mapper (TM) analytical film products, and
(3) LAS-Scrounge TM "engineering" film products. The standard MSS film
products are generated at EDC using the CCT-PM digital data with a Laser
Beam Recorder (LBR) to produce master film copies for black-and-white and
color composite reproduction. The "interim" TM analytical film products
were generated at EDC using the return beam vidicon (RBV) image production
system to make the first generation working masters during the Scrounge

environment and prior to the operational film generation under the TIPS environment. The LAS-Scrounge TM "engineering" film products were generated at GSFC for engineering purposes, archiving, and routing with CCT orders for LIDQA investigators.

The interim TM film products were being used for our analysis because they were the only film products available at the time of our investigation. Originally, film products were to be generated by the LAS-Scrounge during the pre-TIPS environment, but as the demand for these products increased, arrangements were made to have the EDC produce the products using the RBV image production system. Using the CCT-PT data, EDC produced the film masters using the LBR and the supporting computer system formerly dedicated to RBV film production. In order to adapt the TM data to this system, the resulting "interim" TM analytical film product represented only a sub-area of a full TM scene produced under operational conditions (Figure 3.1). The interim TM film product represents approximately 72 percent of the area of a full TM scene. The format is 5322 x 5322 pixels, centered on the full scene, with a resulting image of 20.2 x 20.2 cm covering a land surface area of approximately 2,300,590 ha. (5,684,760 ac.)

3.1.2 Approach and Results

A seven-step procedure was used to evaluate the geometric properties of the interim TM film products.

1. Select the image product to be analyzed. A two-times enlargement of interim Band 7 image T0318-007 (WRS Path 044, Row 033) covering the southern Sacramento Valley on 1 February 1983 was selected for this study. This particular scene was selected because: (1) it was currently available, (2) with the exception of small portions of the image, it was cloud-free, (3) Band 7 provided a sharp, moderate contrast image, and (4) the area covered represented a wide range of land uses and elevational zones. The major limitation of using this image was that, due to the low sun elevation at the time of image acquisition (26 degrees), many of the steep canyons in the wildland areas in the western and northern portions of the scene were deeply shadowed. We felt that this limitation was outweighed by the facts that: (1) the image in its two-time enlarged format was available, and (2) all the small water bodies were at the maximum water levels which would allow for better precision in selecting control points.

2. Grid the image into nine equal area blocks. The TM image was gridded to ensure that the control points would be evenly distributed throughout the photo (Figure 3.2).

3. Select control points. The control points represented those natural and cultural features that could be located reliably on both the TM image and on United States Geological Survey 7½' quadrangle maps. In the agricultural areas, these features were predominantly field intersections or irrigation ditches; in the urban areas they were predominantly major road intersections or airfield runways; in the wildland areas they were water bodies, stream courses, and converging points of ridges and canyons. Each

control point was pin-pricked on the image and a corresponding annotation made on the map sheet with a .30mm pen. Initially, 476 control points were selected from 144 map sheets. All of these map sheets fell within the Universal Transverse Mercator (UTM) Zone 10 (Table 3.1).

4. Measure image and map coordinates. The "x" and "y" coordinates of the image control points were measured to the nearest .001 inch using a Talos plane table digitizer. The corresponding UTM east and north map coordinates were scaled off the map sheets to the nearest 10 meters ground distance.

5. Check for image and map coordinate errors. In order to check for digitizing and map scaling errors, a first order regression between map and image coordinates was performed using the program developed by Daniel (1971). For those coordinate pairs for which the residuals were excessive, the digitized and map coordinates were checked, changed where appropriate, or discarded if map or digitizing errors were probable. The final number of control points totaled 448 (Table 3.1).

6. Develop regression between image and map coordinates. Using the odd-numbered control points, first and second order regressions were developed to predict UTM east and north map coordinates from the digitized image coordinates. Four regression models were examined:

$$(1) \quad UTM = a+b(x)\ c(y)$$
$$(2) \quad UTM = a+b(x\ \ c(y)+f(x*y)$$
$$(3) \quad UTM = a+b(x)+c(y)+d(x2)+e(y2)$$
$$(4) \quad UTM = a+b(x)+c(y)+d(x2)+e(y2)+f(x*y)$$

The results of this examination are summarized in Table 3.2. Based on evaluation of the root mean square errors and the range of the residuals, we selected the second order regression (3) to predict the UTM east coordinates, and the first order regression with cross term (2) to predict the UTM north coordinates.

7. Evaluate the geometric properties of the image. Using the regressions selected in 6, the map coordinates of the even-numbered control points (test points) were predicted using the corresponding digitized image coordinates. The residual vectors, summarized by blocks, were plotted (Figure 3.3); and an analysis of variance (ANOVA) was performed on the east and north residuals using the nine blocks as treatments (Table 3.3).

The scales associated with each block were calculated by comparing the image area to the ground area based a sample of control points. The respective areas were calculated using the following relationship:

$$A = \quad Xi[(Yi+1)-(Yi-1)]*(0.5)$$

where: A = area
X = coordinate value for Talos x or UTM east
Y = coordinate value for Talos y or UTM north, and
i = coordinate number

63

In addition, a mean scale was calculated for the whole image using this relationship with the resulting scale of 1:375,610. A nominal scale was determined by comparing the image distance (15.929 inches/line) to the ground distance (28.5 meters/pixel x 5322 pixles/line = 374,885 meters/line) with the resulting scale of 1:374,885. These results have been summarized in Figure 3.4.

After examining the variability of the scales for the nine blocks, we calculated the scales for 104 line segments distributed throughout the image using the Talos and UTM map coordinates. The resulting scales were plotted at the midpoint of the line segments in order to identify any patterns in the scales (Figure 3.5).

3.1.3 Conclusions

Based on an examination of the root mean square error of the residual of 223 test points (Table 3.3) and the distribution of calculated image scales (Figures 3.4 and 3.5), we conclude that there is no systematic image distortion in the TM image that we evaluated. The analysis of variance for the east residuals showed no significant difference between the blocks, and although a significant F value was calculated for the north residuals, no significant difference could be extracted from the data using either the Scheffé least square difference test or the Duncan's new multiple range test.

By examining the distribution of the scales calculated for individual blocks (Figure 3.4) or of those for the 104 line segments (Figure 3.5) it becomes evident that the changes in scale that occur throughout this image are random. The most probable causes of this variability in scale are digitizing and/or mapping errors. There does not appear to be any significant effect on scale as a result of elevation differences. For example, those areas in which the highest ground elevations occur are often associated with the smallest scales such as occur in the northeastern corner of block CC with a mean elevation of 3,000 feet; and some of the largest scales, such as occur in the central portion of block BB which has a mean elevation of 20 feet.

Finally, we would conclude that the user of this image product could expect that the accuracy of mapped points would be within 91 meters ground distance in the easterly direction and within 117 meters in the northerly direction based on the root mean square value for the residuals. This represents .010 mm and .012 mm, respectively, on the interim image product based on the scale calculated from the test coordinates.

3.2 Geometric Analysis of Thematic Mapper Digital Data

3.2.1 Introduction

The analysis of the geometric properties of digital Thematic Mapper data has not been a major emphasis of our research at the Remote Sensing Research Program. We have had to examine this aspect of the TM products, however, in order that we could accurately extract test sites from the magnetic tapes, from which we have produced image products for

interpretation testing (Chapter 2, Section 3). In addition, other research projects being conducted here have required that we look at these properties for the purpose of constructing geographic information systems which include TM digital data. The following paragraphs describe the results of this research which is being conducted in our 100,000 hectare Plumas Forest study site.

3.2.2 Approach

A six step procedure was used to evaluate the geometric properties of the digital Thematic Mapper data.

1. Select TM coverage for study site. The first summer season Landsat-4 scene that covered the forestry study site was acquired on 12 August 1983 (#84039218143, WRS Path 44, Row 32). These data were transmitted to Goddard Space Flight Center (GSFC) via the Tracking and Data Relay Satellite System and the receiving station at White Sands, New Mexico. The TM data were processed by the Thematic Mapper Image Processing System (TIPS) at GSFC as a "P" tape.

2. Extract area of interest. Through the use of the RSRP interactive image processing system, a 1200-by-1200 pixel sub-area which covered the forestry study site was extracted from the "P" tape.

3. Select control points. These points represented those natural and cultural features that could be located reliably on both the digital display of the TM data on the RSRP image processing system and on one $7\frac{1}{2}$ minute orthophoto quadrangle in the study site. A total of 42 points were allocated evenly throughout the quadrangle. The TM point and line pair counts and their respective UTM east and north coordinates were recorded for subsequent analysis.

4. Check for control point errors. In order to check for errors that might have occurred during coordinate extraction from the displayed image and/or the orthophoto quadrangle, a first order regression was used to predict UTM east and north coordinates from TM point and line counts. If the resulting residual was excessive (greater than 30 meters), the corresponding coordinate pair was checked on both the map and on the image display.

5. Develop regressions between point-line coordinate pairs and UTM east-north coordinate pairs. Through the use of 20 of the control points, first and second order regressions were developed to predict UTM east and north map coordinates from the point and line counts. Four regression models were examined:

$$(1) \quad UTM = a+b(P)+c(L)$$
$$(2) \quad UTM = a+b(P)+c(L)+f(P*L)$$
$$(3) \quad UTM = a+b(P)+c(L)+d(P2)+e(L2)$$
$$(4) \quad UTM = a+b(P)+c(L)+d(P2)+e(L2)+f(P*L)$$

where: P = point coordinate
L = line coordinate
P2 = point coordinate squared
L2 = line coordinate squared

65

6.   Evaluate the geometric properties of the digital data.   Through the application of the optimum regression equation, the geometric properties of the digital data were evaluated using the remaining 22 control, or test, points.

3.2.3 Results and Discussion

A summary of the four regression models that were examined is presented in Table 3.4. By examining the four statistical measures of goodness of fit for the regression we concluded that, for an area as small as one $7\frac{1}{2}$ minute quadrangle, a first order regression with no cross terms was superior. This model produced: (1) the lowest residual root mean square, (2) the highest F-value,   (3)   the   narrowest   range   of   residuals,   and   (4)   regression coefficients that were significantly greater than zero. The resulting model that was tested took the following form:

$$\text{UTM EAST} = 652567+28.1675(P)-4.78209(L)$$
$$\text{UTM NORTH} = 436319-4.7076(P)-28.088(L)$$

The statistical results obtained by applying these regression models to the 22 test points are summarized in Table 3.5 and plotted in Figure 3.6. The mean deviation represents the average direction of geometric and plotting errors;   the   root   mean   square (RMS)   deviation   represents   the average magnitude of these errors.   This latter value can be compared with those calculated for the original regression.   Interestingly, the RMS for the test-point eastings was less than that for the regression-point eastings; whereas the RMS for the test-point northings exceeded that for the regression-point northings. While this paradox was unexpected, it was further confirmed by the small range of residuals associated with the east test points when compared to the regression points, and by the larger range associated with the north test points when compared to the regression-point range.   We concluded that the regression did a good job predicting the test points because the calculated mean deviations for east (+0.896 meters) and north (-0.835 meters) were not significantly different than zero, which is, by definition, the expected value of the regression model. Based on these analyses, we conclude that by using a first order regression and sufficient control points, we can predict the map position of a TM pixel within 12 meters east and 14 meters north.   In addition we conclude that, for geographic areas as small as a $7\frac{1}{2}$ quadrangle, the addition of second order or cross term coefficients does not significantly add to the predictive power of the regression.

3.3.0 Analysis of Scene-to-Scene Registration of Thematic Mapper "P" Data

3.3.1 Introduction

The NASA design goal for the full scene-to-scene, or temporal, registration of TM "P" data is 0.3 pixels 90 percent of the time (Beyer, 1983).   The objective of the research documented in this section was to analyze the magnitude of this registration error for a 1024-by-1024 pixel test area extracted from the 31 December 82 and 12 August 83 thematic mapper scenes covering the San Felipe area of California (WRS P44,R34).

3.3.2 Approach

The 1024-by-1024 pixel test area was divided into sixteen blocks to ensure that control points used for subsequent analysis would be distributed evenly throughout the test area. In each block, three to five common control points were selected and their respective point and line coordinates for each date were recorded; a total of 55 control points were selected for the whole test area. The scene-to-scene relationship between the control points was established through the application of first and second order regressions. Four regression models were examined:

$$(1) \quad TM(AUG) = a + b*P + c*L$$

$$(2) \quad TM(AUG) = a + b*P + c*L + f*(P*L)$$

$$(3) \quad TM(AUG) = a + b*P + c*L + d*P2 + e*L2$$

$$(4) \quad TM(AUG) = a + b*P + c*L + d*P2 + e*L2 + f*(P*L)$$

where: P = point coordinate from December subscene,
       L = line coordinate from December subscene,
       P2 = point coordinate squared from December subscene, and
       L2 = line coordinate squared from December subscene.

Through the application of the optimum regression equation, the geometric properties of the digital data were evaluated based on overall regression coefficients and error terms and on pooled values for the 16 blocks.

3.3.3 Results and Discussion

A summary of the four regression models that were developed is presented in Table 3.6. By examining four statistical measures of goodness of fit we concluded that, for overlaying two dates of TM data for an area as small as 1024-by-1024 pixels, a first order regression with no cross terms was superior to the other higher order regressions. This model produced: (1) the lowest residual root mean square for points, lines, and point-lines, (2) the narrowest range of residuals for points and third narrowest range for lines, (3) the highest F-values, and (4) the highest t-values for individual regression coefficients. The resulting model that was analyzed took the following form:

TMPOINT(AUG) = 1.56819 + 0.99943*TMPOINT(DEC) - .00051808*TMLINE(DEC)

TMLINE(AUG) = 1.94778 - .00168156*TMPOINT(DEC) + .999945*TMLINE(DEC)

In an attempt to determine if the residual errors were localized within the test area, the residuals for individual blocks were calculated (Table 3.7), and the the resulting vectors were plotted (Figure 3.7).

Based on the statistics presented in Tables 3.6 we concluded that by using first order regression models, we could overlay two dates of TM data to within 0.778 points and 0.596 lines, with an overall vector fit of 0.980 pixels. Further, by examining the direction of these residual vectors in Figure 3.7, we concluded that the residuals were random with respect to both direction and magnitude. The wide range of the point and line residuals in Table 3.7 confirms the random nature of the distribution of these residuals. It must be emphasized that when reading Table 3.7 mean block residuals and mean block root mean square residuals should be used to evaluate local accuracy of the two-date overlay. These statistics were summarized for the individual blocks based on the regressions determined from all coordinate pairs. Therefore, it was possible to use the direction of the residuals when summing those block values to be squared. Hence, these values are considerably lower than those presented in Table 3.6.

Based on these results, it is clear that we were unable to meet the 0.3 pixel scene-to-scene overlay specifications even on the much smaller subscene. We do not know at this point whether this failure to meet the specifications was due to deficiencies in the Landsat TM sensor or to subsequent image interpretation error.

We did learn, however, that the accurate location of common ground points on the two images was much more difficult than we had anticipated. This was due to two interrelated factors that always exist when attempting to overlay summer and winter images in wildland environments.

(1)   Because of the rugged, heavily dissected terrain of the study area, many ideal control points visible on the August image were obscured by shadows on the December image. Most notable of these were stream and river intersections. In addition to the shadowing, the water levels of these streams were much higher in December than in August.

(2)   In the coastal hills of California the appearance of natural vegetation and some cultural features changes markedly from summer to winter; this was particularly true in 1983 when an earlier and heavier than usual rainfall occurred. The vegetation present in this area is predominantly annual grasslands and mixed deciduous and evergreen hardwoods. In the late summer, the dried grasslands exhibit a high albedo from the visible through middle reflectance infrared range of the electro magnetic spectrum and cannot be distinguished from rock outcrops and bare soil; in the winter, the interface between the nonvegetated areas and the now actively metabolizing grasslands is quite dramatic. In addition, some, but not all, of the hardwood species in this area lose their leaves in the winter. Therefore, a grassland-deciduous hardwood interface that would be easily identified on the August image would be lost on the winter image. And finally, the interface between the decidous hardwood-evergreens and the hardwoods that was quite visible on the December image was not apparent on the August image. The major cultural features used as control points were the many watering ponds present in the study site. Both the sizes and number of the ponds were greater in December due to the heavy rains. Although we took care in trying to locate the control points at the ponds' dams, some locational error was inevitable.

68

Although our sub-scene-to-scene overlay error was greater than design
specifications, we feel sure that the overlay accuracy is more than
sufficient to meet those wildland management information needs that require
such multidate data. It should be emphasized that the conditions
experienced in this study were those of an extreme case in terms of the
changes in sun angle and ground features. If the period between the two
images had been only three months, we feel the overlay error might have been
reduced by half. Similar results have been reported by Brooks, et al.,
1984, Byer, et al., 1984, Walker et al., 1984, and Welch and Usery, 1984,

3.4   References

Beyer, Eric P. 1983. An overview of the Thematic Mapper geometric
    correction system. Landsat-4 Science Investigations Summary, Volume I.
    NASA Conference Publication 2326. Goddard Space Flight Center,
    Greenbelt, Maryland. p.98.

Beyer, E.P., J. Brooks, and V.V. Salomonson. 1984. Geometric correction of
    Landsat 4 and 5 Thematic Mapper data. Eighteenth Intern. Sympos. on
    Remote Sensing of Environment. Paris, France, 1-5 October.

Brooks, J., A. Jai, T. Keller, E. Kimmer, J. Su. 1984. Thematic Mapper
    geometric correction performance evaluation. Proc. Machine Processing
    of Remotely Sense Data Symp. Purdue University. West Layfayette,
    Indiana. pp. 22-28.

Daniel, Cuthbert and Fred S. Wood, 1971. Fitting equations to data. Wiley-
    Interscience, New York.

Durrenberger, Robert W. and Robert B. Johnson, 1976. California - patterns
    on the land. Fifth edition. Mayfield Publishing Company.

Walker, R.E., A.L. Zobrist, N.A. Bryant, B. Gohkman, S.Z. Friedman, and T.L.
    Logan. 1984. An analysis of Landsat-4 geometric properties. IEEE
    Transactions on Geoscience and Remote Sensing, Vol. GE-22, pp. 288-293.

Welch, R. and E.L. Usery. 1984. Cartographic accuracy of Landsat-4 MSS and
    TM image data. IEEE Transactions on Geoscience and Remote Sensing,
    Vol. GE-22, pp. 281-288.

TM SUBAREA

TM FULL SCENE
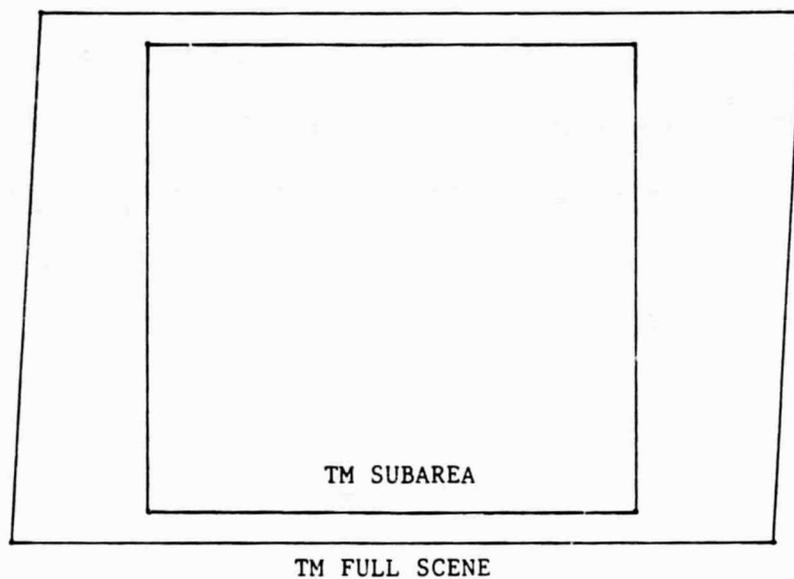
Figure 3.1 Relative size of the interim Thematic Mapper image with respect
to the full frame image.   A two-time enlargement of an image
covering the northern Sacramento Valley, California, was used to
evaluate the geometric properties of the Thematic Mapper photo
products (WRS Path 44 Row 33).

70

```
            A         B         C
        ┌─────────┬─────────┬─────────┐
        │ ++   +  │         │         │
        │    + + +│         │         │
   C    │ + ++  + │ SELECT 40-50 POINTS PER BLOCK    C
        │  ++   + │         │         │
        │   +   + │         │         │
        ├─────────┼─────────┼─────────┤
        │         │         │         │
   B    │         │         │         │    B
        │         │         │         │
        ├─────────┼─────────┼─────────┤
        │         │         │         │
   A    │         │         │         │    A
        │         │         │         │
        └─────────┴─────────┴─────────┘

            A         B         C
```
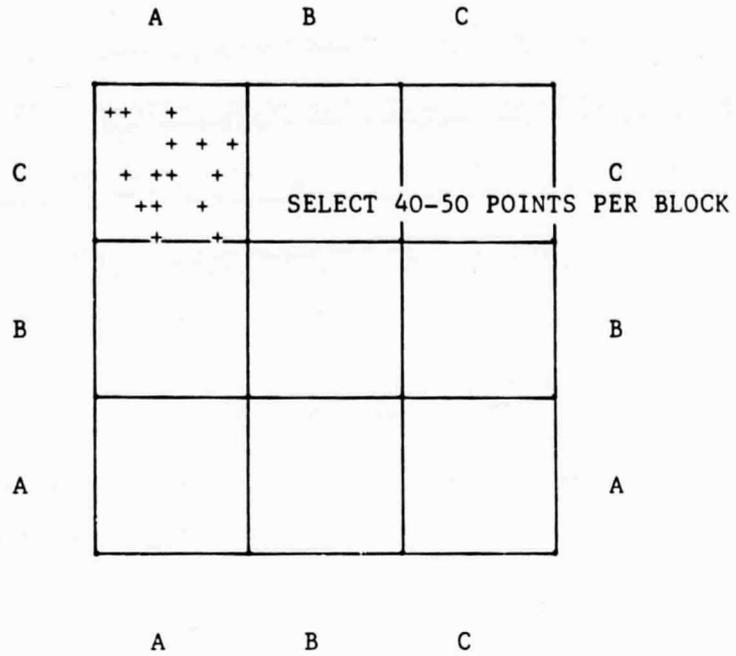
Figure 3.2   Schematic of   the gridded interim TM image.   The 5322-by-5322 pixel   image was gridded into nine equal area blocks to   ensure that the control points would be evenly distributed   throughout the image.

71
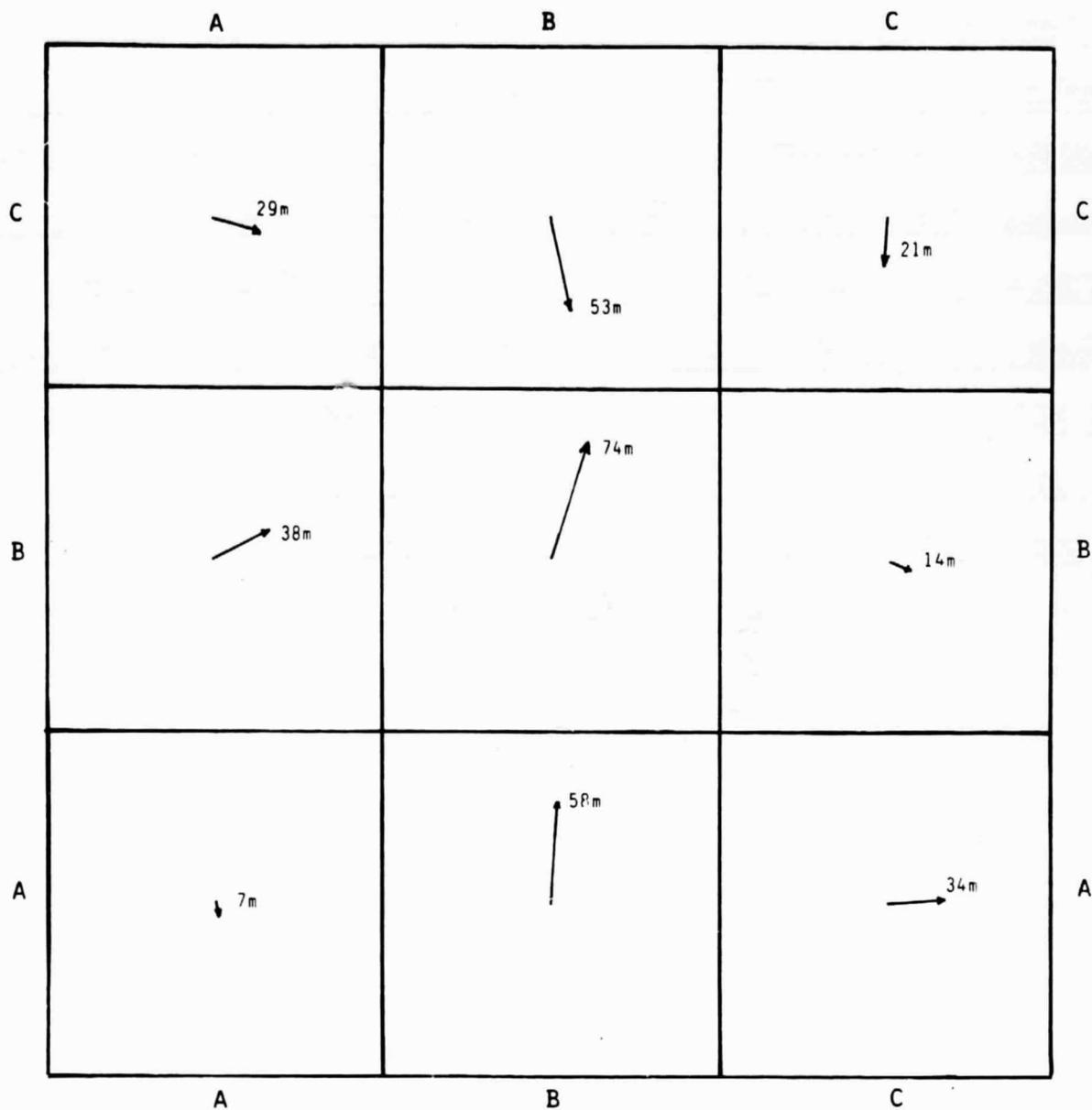
Figure 3.3   Residual vectors, summarized by blocks, of test points located throughout the TM interim image product. These residuals were the result of applying a second order regression and a first order regression with a cross term to Talos x and y tablet digitizer coordinates to predict UTM north and east coordinates, respectively. Notes: (1) the points that were used to develop the two regression equations were <u>not</u> used as test points, and (2) the residual vectors are not drawn to scale with respect to the interim TM image.

SCENE TO318-007, PATH 044,ROW 033, 01 FEBRUARY 83

	A	B	C	
C	(−1.16) $\frac{1}{371,252}$	(−1.74) $\frac{1}{369,088}$	(+0.21) $\frac{1}{376,411}$	C
B	(−0.95) $\frac{1}{372,056}$	(−0.41) $\frac{1}{374,083}$	(−0.18) $\frac{1}{374,924}$	B
A	(−0.15) $\frac{1}{375,048}$	(+0.12) $\frac{1}{376,074}$	(−0.15) $\frac{1}{375,057}$	A

A   B   C

$$\text{CALCULATED MEAN SCALE} = \frac{1}{375,610}$$

$$\text{NOMINAL SCALE} = \frac{15.929 \text{ inches/line}}{28.5 \text{ meters/pixel x } 5322 \text{ pixels/line}}$$

$$= \frac{1}{374,885}$$

Figure 3.4   Calculated scales and percent departures () from the mean scale for the nine blocks on TM Scene TO318-007. These scales, and the overall scene scale, were calculated based on the UTM and digitized points selected for the regression analysis described in the text.   The nominal scale was calculated on the relationship of the digitized photo distance along the "x" axis (15.929 inches) to the expected ground distance (28.5 meters/pixel x 5322 pixels/line).

73
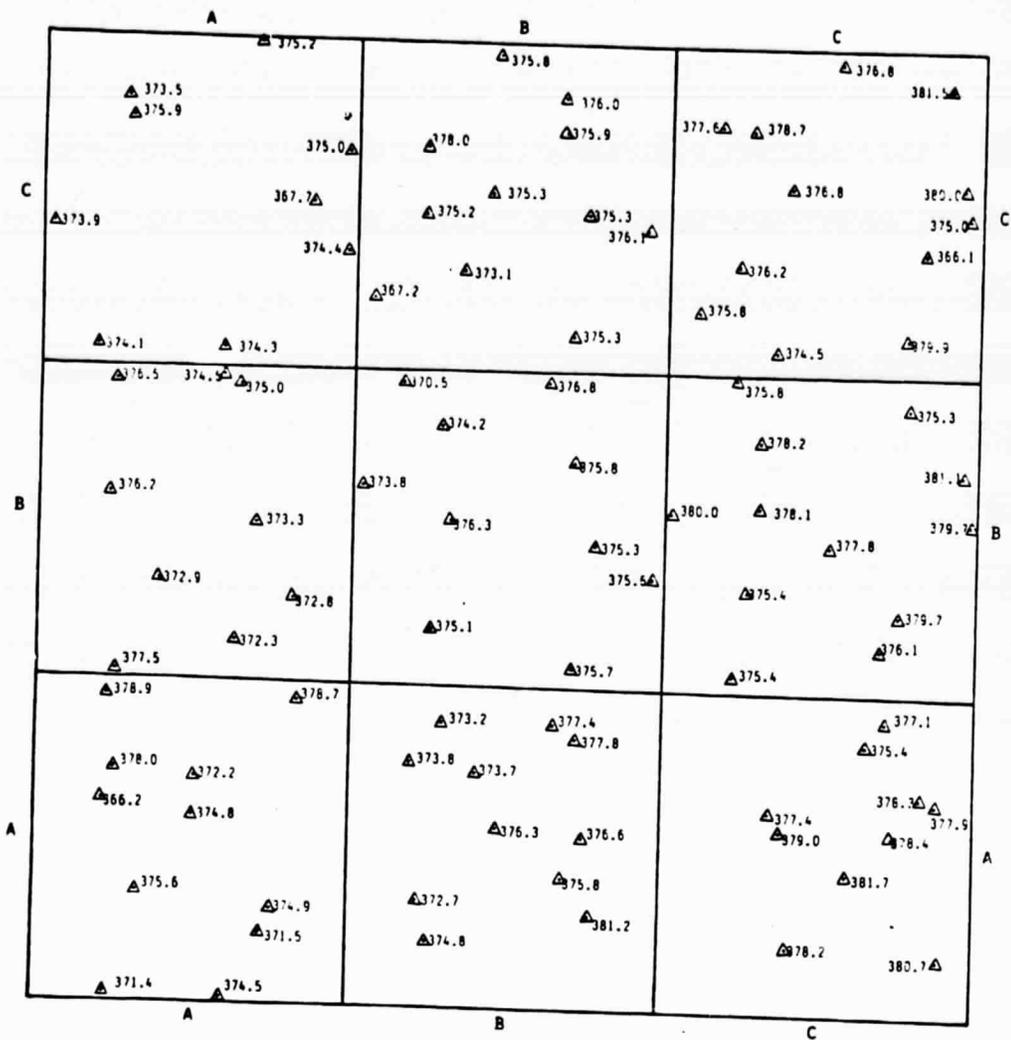
Figure 3.5 Calculated scales (denominators in thousands) for 104 segment midpoints distributed throughout the interim TM image. Because no reasonable isolines could be drawn for these points, we concluded that there was no systematic distortion in the image.
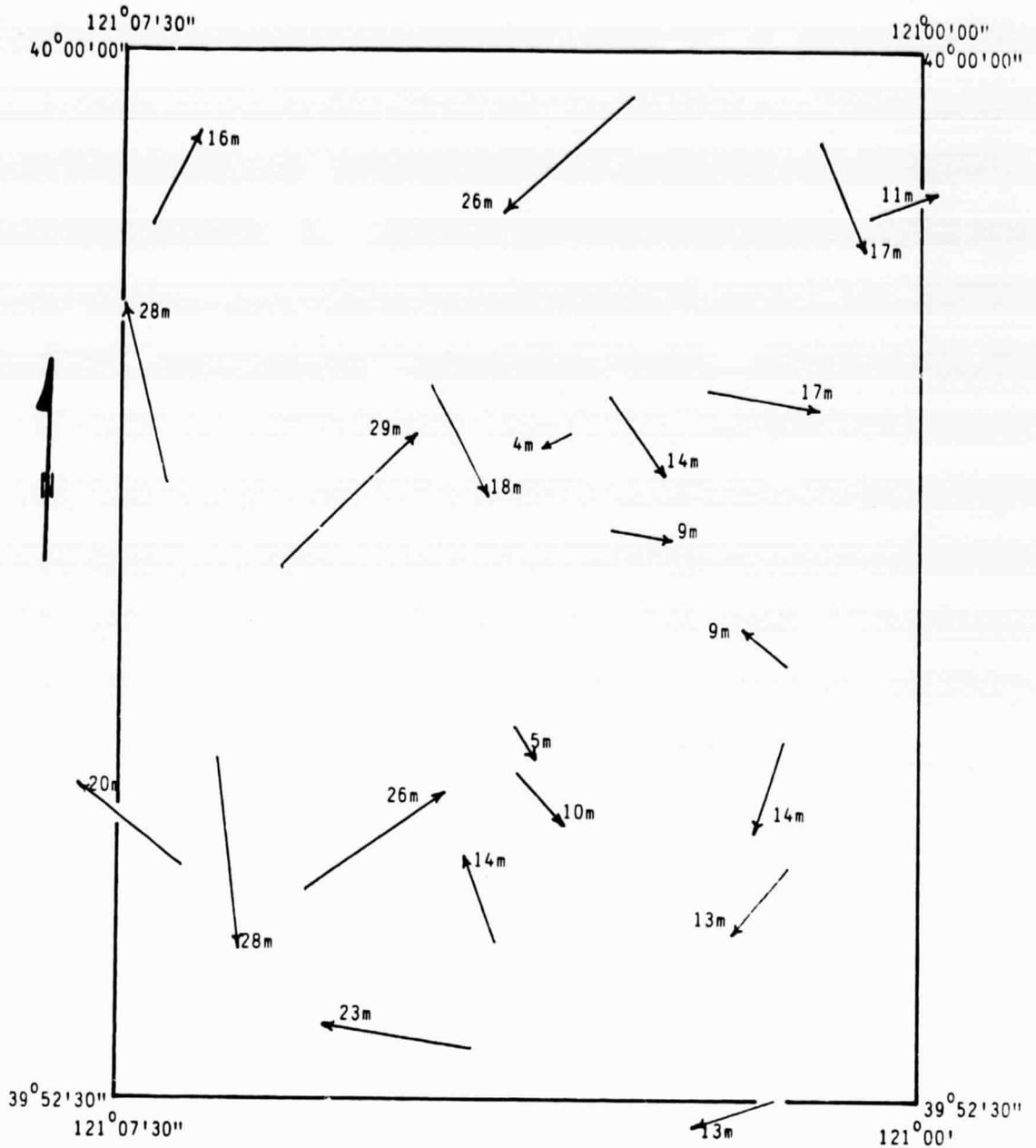
Figure 3.6 Residual vectors for 22 test points. These vectors were calculated by applying first order regressions determined from 20 control points located throughout a 7½ minute quadrangle. Note: the residual vectors are not drawn to scale with respect to the 7½ minute quadrangle.
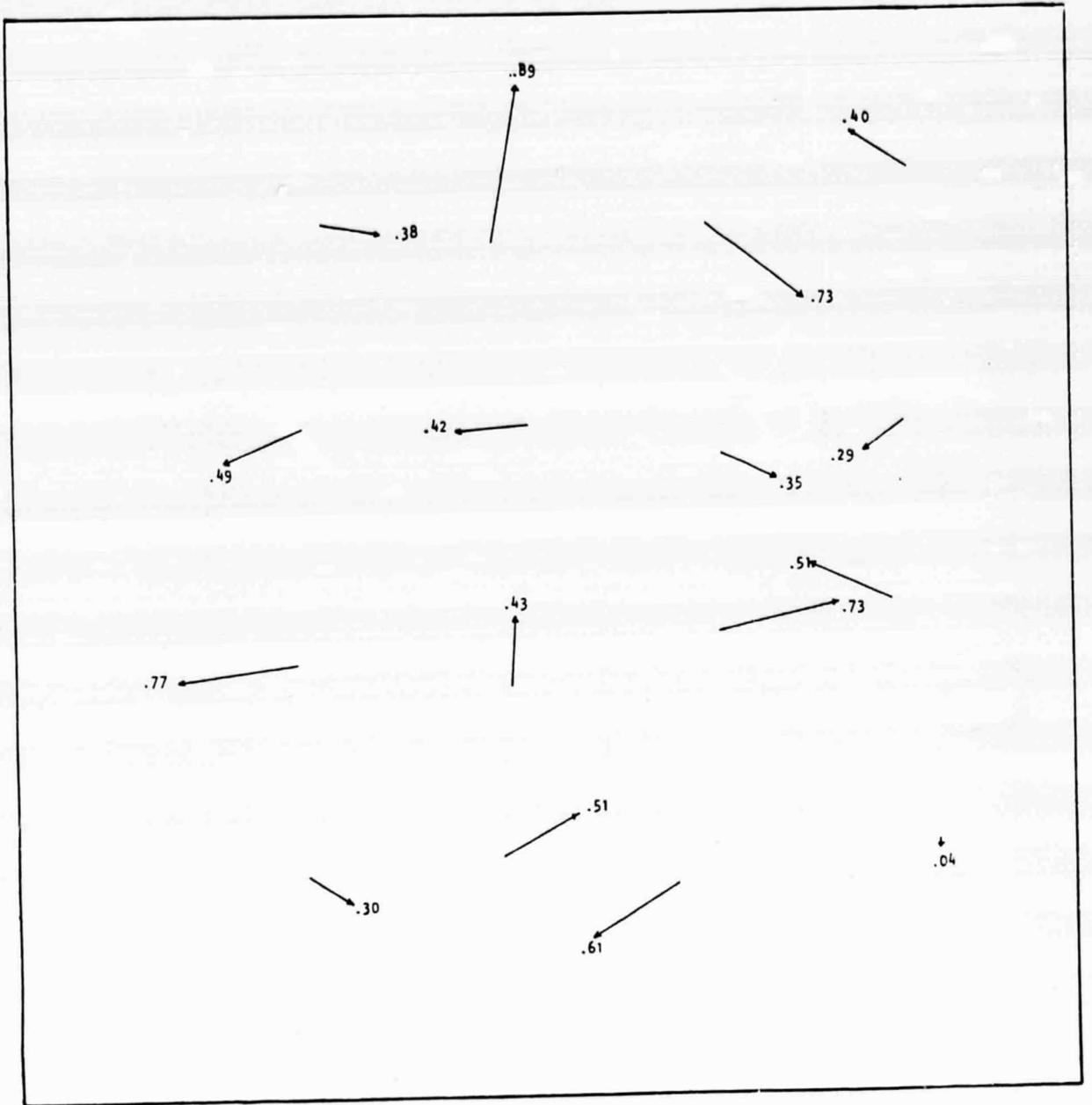
Figure 3.7   Residual   vectors   calculated from a multidate   registration   of
two   TM   1024-by-1024   pixel   subscenes.   These   vectors   were
calculated using first order regressions and summarized for   16
blocks distributed evenly throughout the test area.

Table 3.1 Summary of Control Point Selection for Scene TO318-007 (Path 044, Row 033), 01 February 1983.

BLOCK	CLOUD COVER (%)	ELEVATION RANGE OF POINTS (FEET)	CONTROL POINTS		GEOGRAPHIC DESCRIPTION	PREDOMINANT LAND USE*
			NUMBER OF REGRESSION POINTS	NUMBER OF TEST POINTS		
AA	10	14-2080	26	25	Napa Valley; Coast Range	3,1,4
AB	50	5-384	16	17	Coast Range; Sacramento Valley	1,3,4
AC	5	0-391	20	19	Sacramento Valley; Sierra Nevada foothills	1,3,4
BA	15	43-1960	29	28	Coast Range; Sacramento Valley	3,1
BB	0	20-199	24	24	Sacramento Valley	1
BC	0	20-2000	26	27	Sacramento Valley; Sierra Nevada foothills	3,1,4
CA	0	54-1540	32	30	Coast Range; Sacramento Valley	1,3
CB	0	48-1384	24	26	Sacramento Valley; Sierra Nevada foothills	1,3
CC	0	200-3540	28	27	Sierra Nevada foothills and mountains	2,3
			225	223		

* LAND USE CODE (from: Durrenberger and Johnson, 1976.)

1   Cropland and pasture - irrigated or unirrigated
2   Forest and woodland - grazed or ungrazed
3   Open shrub woodland - grazed
4   Urban

Table 3.2 Summary of root mean square errors for four regressions to predict UTM east and UTM north from 225 digitized observations taken from a two-times enlargement of an interim TM print.

Regression	Residual Root Mean Square (meters)	F-Value	Range of Residuals (meters east)	t-Values* for Coefficients of: x	y	x2	y2	xy
UTM EAST								
a+b(x)+c(y)	120.90	1.621E7	-289 to 284	5567	1012			
a+b(x)+c(y)+f(x*y)	120.76	1.083E7	-287 to 279	1885	447			1.2
a+b(x)+c(y)+d(x2)+e(y2)	76.17	2.043E7	-308 to 241	1694	267	18	4.1	
a+b(x)+c(y)+d(x2)+ e(y2)+f(x*y)	75.46	1.665E7	-312 to 244	1481	248	18	4.0	2.3
UTM NORTH								
a+b(x)+c(y)	156.97	7.990E6	-420 to 369	846	3932			
a+b(x)+c(y)+f(x*y)	117.62	9.487E6	-525 to 283	369	2297			13
a+b(x)+c(y)+d(x2)+e(y2)	155.55	4.068E6	-402 to 422	164	667	1.7	1.8	
a+b(x)+c(y)+d(x2)+ e(y2)+f(x*y)	112.94	6.173E6	-508 to 347	188	839	2.9	3.6	14

===

* If the calculated t-values exceed 1.645, the coefficients are significantly greater than zero. Note that the relative magnitude of the t-values associated with the "x" and "y" coefficients is meaningless. This relationship is a function of how the TM image was oriented with respect to the x-y grid of the Talos digitizing tablet; there is no relationship to the image's orientation with respect to true north.

Table 3.3 Summary of analysis of variance performed on the mean residuals of the test points.

UTM EAST SUMMARY
Regression Equation: UTM East = $a+b(x)+c(y)+d(x2)+e(y2)$

BLOCK	NUMBER OF OBSERVATIONS	RANGE OF RESIDUALS (METERS EAST)	MEAN RESIDUAL (METERS)	ROOT MEAN SQUARE OF RESIDUALS (METERS)
AA	25	-189 to 201	1.007	112
AB	17	-109 to 222	4.249	76
AC	19	-47 to 148	33.659	58
BA	28	-250 to 219	32.840	106
BB	24	-59 to 180	22.848	60
BC	27	-419 to 186	13.227	111
CA	30	-175 to 159	29.522	70
CB	26	-214 to 181	10.082	94
CC	27	-350 to 147	-2.734	101
TOTAL	223		MEAN = 16.079	MEAN = 91

ANALYSIS OF VARIANCE FOR MEAN RESIDUALS

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F-VALUE
TREATMENT	40647.554	8	5080.944	.616
ERROR	1763883.875	214	8242.448	
TOTAL	1804531.430	222		

Table 3.3 (concluded).Summary of analysis of variance performed on the mean residuals of the test points.

UTM NORTH SUMMARY
Regression Equation: UTM North = a+b(x)+c(y)+f(x*y)

BLOCK	NUMBER OF OBSERVATIONS	RANGE OF RESIDUALS (METERS NORTH)	MEAN RESIDUAL (METERS)	ROOT MEAN SQUARE OF RESIDUALS (METERS)
AA	25	-270 to 118	-6.498	87
AB	17	-69 to 177	57.556	89
AC	19	-130 to 187	1.942	86
BA	28	-260 to 365	18.532	140
BB	24	-100 to 239	70.359	104
BC	27	-303 to 333	-3.917	120
CA	30	-468 to 257	-6.415	124
CB	26	-438 to 246	-51.792	156
CC	27	-261 to 166	-20.689	101
TOTAL	223		MEAN = 3.830	MEAN = 117

ANALYSIS OF VARIANCE FOR MEAN RESIDUALS

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F-VALUE
TREATMENT	265522.637	8	33190.330	2.565*
ERROR	2769099.374	214	12939.717	
TOTAL	3034622.011	222		

===
* Although this F-value is significant at .95, no significant differences could be determined using either the Duncan's new multiple range test or the Scheffé test for possible contrasts.

Table 3.4 Summary of root mean square errors for four regressions using Thematic Mapper digital point (P) and line (L) counts to predict UTM east and UTM north. These regressions were based on 20 observations located throughout one 7½ minute quadrangle.

Regression	Residual Root Mean Square (meters)	F-Value	Range of Residuals (meters east)	t-Values* for Coefficients of:				
				P	L	P2	L2	P*L
UTM EAST								
a+b(P)+c(L)**	16.38	3.186E5	-32 to 28	786	182			
a+b(P)+c(L)+?(P*L)	16.81	2.017E5	-3? to 30	261	36			.37
a+b(P)+c(L)+d(P2)+e(L2)	16.93	1.492E5	-28 to 35	77	29	.95	.11	
a+b(P)+c(L)+d(P2)+e(L2)+f(P*L)	17.52	1.114E5	-28 to 35	73	20			
UTM NORTH								
a+b(P)+c(L)**	9.92	1.614E6	-15 to 17	217	1768			
a+b(P)+c(L)+f(P*L)	10.04	1.051E6	-15 to 18	72	352			.78
a+b(P)+c(L)+d(P2)+e(L2)	10.31	7.468E5	-14 to 16	21	281	.61	.59	
a+b(P)+c(L)+d(P2)+e(L2)+f(P*L)	10.28	6.007E5	-13 to 17	21	201	.97	.50	1.04

===
* If the calculated t-value exceeds 1.734, the coefficient is significantly greater than zero at the .95 probability level.

** Regressions selected for testing:

UTM EAST = 652567+28.167(P)-4.78209(L)

UTM NORTH = 436319-4.7076(P)-28.088(L)

Table 3.5   Summary of error analysis of 22 test points using first order regressions to predict UTM east and north from Thematic Mapper digital point (P) and line (L) counts. The regression coefficients were calculated from 20 observations that were distributed, along with the test points, throughout a 7½ minute quadrangle.

REGRESSION EQUATION TESTED	RANGE OF RESIDUALS (METERS)	MEAN DEVIATION (METERS)*	ROOT MEAN SQUARE DEVIATION (METERS)**
UTM EAST = 652567+28.1675(P)−4.78209(L)	−23 to 21	+0.896	12.0
UTM NORTH = 436319−4.7076(P)−28.088(L)	−28 to 27	−0.835	13.7

===

*   The mean deviation represents the average direction of geometric errors. This value can be calculated only from test points, because by definition, the sum of the deviations from the points used to develop the regression coefficients must equal zero.

** The root mean square deviation represents the average magnitude of geometric errors without respect to direction of those errors.

82

Table 3.6 Summary of root mean square errors for four regressions to analyze the scene-to-scene registration for a 1024-by-1024 test area extracted from Thematic Mapper scenes from 31 December 1982 and 12 August 1983. In this study, 55 point and line coordinate pairs were extracted from the December subscene. These pairs were used as independent variables to predict the respective coordinate pairs on the August subscene.

Regression	Residual Root Mean Square (points)	Residual Root Mean Square (pixels)	F-Value	Range of Residuals (points)	t-Values* for Coefficients of:				
					x	y	x2	y2	xy
TM POINT									
a+b(P)+c(L) *	.7784	.9804	2.2*10e6	-2.0 to 1.1	2081	1			
a+b(P)+c(L)+f(P*L)	.7853	.9872	1.4*10e6	-2.0 to 1.5	862	.7			.3
a+b(P)+c(L)+d(P2)+e(L2)	.7836	.9893	1.1*10e6	-2.0 to 1.2	352	.3	1.2	.1	
a+b(P)+c(L)+d(P2)+e(L2)+f(P*L)	.7911	.9988	8.4*10e6	-2.0 to 1.2	319	.4	1.1	.1	.2
	(lines)	(pixels)		(lines)					
TM LINES									
a+b(P)+c(L) *	.5962	.9804	4.1*10e6	-1.3 to 1.0	4	2865			
a+b(P)+c(L)+f(P*L)	.5983	.9872	2.7*10e6	-1.3 to 1.0	2	1028			.1
a+b(P)+c(L)+d(P2)+e(L2)	.6037	.9893	2.0*10e6	-1.3 to 0.9	.5	548	.2	.2	
a+b(P)+c(L)+d(P2)+e(L2)+f(P*L)	.6097	.9988	1.6*10e6	-1.3 to 0.9	.5	454	.2	.2	.1

==

* Regressions selected for testing:

TMPOINT(AUG) = 1.56819 + 0.99943*TMPOINT(DEC) - .00051808*TMLINE(DEC)

TMLINE(AUG) = 1.91778 - .0016815 6*TMPOINT(DEC) + 0.999945*TMLINE(DEC)

83

Table 3.7 Error analysis of the multidate overlay of 55 coordinate pairs, summarized by geographic area (block) within a 1024-by-1024 subscene, using first order regressions to predict point and line coordinate pairs for the August 1983 scene from December 1982 coordinate pairs.

REGRESSION TESTED   TMPOINT(AUG) = 1.56819 + 0.99943*TMPOINT(DEC)
                                          - .00051808*TMLINE(DEC)

RANGE OF BLOCK RESIDUALS   -.764 to .771 points

MEAN BLOCK RESIDUALS   -.022 points

MEAN BLOCK ROOT MEAN SQUARE RESIDUALS   .481 points

REGRESSION TESTED   TMLINE(AUG) = 1.91778 -.0016815б*TMPOINT(DEC)
                                          - 0.999945*TMLINE(DEC)

RANGE OF BLOCK RESIDUALS   -.441 to .880 lines

MEAN BLOCK RESIDUALS   .014 lines

MEAN BLOCK ROOT MEAN SQUARE RESIDUALS   .309 lines

84

Chapter 4

MULTISPECTRAL ANALYSIS OF THEMATIC MAPPER DATA FOR A WILDLAND AREA USING UNSUPERVISED CLUSTERING TECHNIQUES

S.D. DeGloria, K. J. Dummer, and M. Haley
Remote Sensing Research Program

4.1   Introduction

Two popular methods for mapping forest vegetation with digital remote
sensing data are to use either a pattern recognition algorithm, which
employs a maximum likelihood decision rule, or an unsupervised clustering
algorithm, which employs a user-specified intercluster distance and cluster
means criteria.   A third method, and perhaps the most cost-effective one,
combines these first two methods: (1) the clustering algorithm is applied
to a sample subset of pixels (1-10 percent) in order to generate means and
covariance matrices of the existing ground conditions; and (2) these
training statistics are input to a maximum likelihood classifier to produce
the final class map for the entire study area.   Regardless of the method
used,   the final map product represents the spatial distribution of forest
cover types as they relate to spectral classes.   Although there is rarely a
one-to-one correspondence between these spectral classes and an existing
management-based forest classification scheme, the spectral class map can
serve quite adequately as a stratification base for subsequent allocation of
photo and/or ground sample plots.   This is based on the assumption that
there is a high correlation between the class map, which partitions the
spectral variability of the area, and the forest parameters which are to be
inventoried.

Because of the importance of a clustering algorithm's performance as
either a stand-alone classifier or as a precursor to the maximum likelihood
classifier, one needs to evaluate the algorithm's advantages and limitations
when applied to the higher spatial and spectral resolution TM data for
characterizing forest cover types.   The objective of this phase of our
investigation was to determine the degree to which the improved spectral
quality of the TM sensor can discriminate forest cover types using the
unsupervised classification algorithm CLUSTER.

4.2   Experimental Methods

To evaluate the spectral quality of TM digital data for discriminating
specific forest cover types and conditions, an unsupervised clustering
program was employed on a subset of TM bands representing the visible and
reflectance infrared spectral regions.   The Landsat-4 scene used for our
analysis was acquired on 12 August 1983 (#8403921843, WRS Path 44, Row 32).
The data were processed by the Thematic Mapper Image Processing System
(TIPS) at Goddard Space Flight Center as a CCT-PT or "P" tape.   The data on
the "P" tape had been both radiometrically and geometrically corrected.   The

TM data representing the ground area covered by the Meadow Valley 7½' USGS topographic quadrangle were extracted from the CCT representing Quadrant #4 of the full scene; the number of pixels processed was 286,000 covering the 232 square kilometer study area.

A subset of TM spectral bands (TM 2-5) was input to an unsupervised classifier, called CLUSTER, in an attempt to separate the set of multidimensional data points into distinct groups or clusters, which would represent forest cover types important to resource inventory and management practices. The selection of these four bands was based, in part, on the results obtained by Anuta, et al. (1984) where the best classification results using a four-band set of TM data were obtained using a band from each major spectral region. In our study, we selected an additional band in the visible region, the green band (TM 2) in lieu of the thermal band (TM 6).

CLUSTER is an adaptation of the clustering algorithm ISODATA, for Iterative Self-Organizing Data Analysis Technique A. ISODATA was originally developed at the Stanford Research Institute and used in their PROMENADE system (Ball and Hall, 1966). The algorithm was implemented at NASA-Johnson Spacecraft Center (JSC) under the name ISOCLS, developed by E. Kan and A. Holly with documentation by R. Minter (1972). The JSC routine was later modified at the Remote Sensing Research Program, under the name ISOCLAS, in order to run data in batch mode on the Control Data Corporation 7600 computer at the Lawrence Berkeley Laboratory. CLUSTER is a further adaptation of ISOCLAS designed to run on a Data General minicomputer system (Ritter and Kaugars, 1978) in which several improvements were made and some new features were added to take advantage of the smaller computer's interactive image processing capabilities.

The CLUSTER program uses an algorithm which employs as a measure of similarity between data points the absolute (non-Euclidean) distance between a data point and a cluster center. The program uses an interactive process to develop the cluster set, refining clusters at each iteration. The user-specified clustering controls are based primarily on the standard deviation of the normalized data cluster, the distance between clusters, and the number of iterations required to achieve a stable number of clusters (Table 4.1). The setting of the clustering parameters was based on these and other investigators' (Nelson, 1981) experiences with the clustering of MSS data; and the resources available did not permit additional trials to determine the optimum settings of these parameters when clustering TM data. The output products from CLUSTER include cluster assignment matrices, and the mean and standard deviation of each cluster. The program is interactive allowing the results (cluster statistics) to be used as a starting point, or "seed", for subsequent clustering activities on the same or similar data sets. This is a useful and efficient operation when clustering large areas. In the present example, a one percent sample of the TM pixels was used to generate an initial set of statistics using a total of 32 iterations. Through the use of these results as seed statistics, the entire study area was clustered using one additional iteration resulting in 43 spectral clusters.

86

To aid in the labeling of the 43 clusters, a data file was created in which the cluster values for all the pixels in the study area were arrayed and displayed in a map format. In addition, the clusters were ranked, based on each cluster's mean: (1) near reflectance infrared-to-red ratio (TM 4/TM 3); (2) middle reflectance infrared-to-red ratio (TM 5/TM 4); (3) albedo, or four-band Euclidian combination; (4) near reflectance infrared value (TM 4); and (5) middle reflectance infrared value (TM 5) (Table 4.2).

Ground cover classes were assigned to each cluster by spatially locating each cluster in the study area, interpreting vertical, color aerial photography (1:24,000), and checking the correspondence between cluster spectral values and preliminary class assignments with the network of ground plots surveyed during the summer of 1982.

4.3 Results and Discussion

The results of applying the clustering algorithm to this data set indicate that both the improved spectral and spatial resolution of the TM data allow an increased number of unique spectral classes to be defined. After 32 iterations, the standard deviation and distance between clusters had stabilized below and above the respective thresholds set by the investigators, yielding 43 spectral clusters (Figure 4.1). By a comparison of each cluster's spectral properties and locational attributes to forest cover conditions, as interpreted on the aerial photography, ground cover class assignments were generated for each of the clusters (Table 4.3).

Of the 43 clusters generated, 22 of them occurred at one or more of the 130 field plots surveyed during the summer of 1982. These 22 clusters represented 259,244 pixels, or 90.6 percent of the study area subjected to the clustering algorithm. The summary of surface characteristics is shown in Table 4.4. These characteristics were extracted from the field survey sheets for the one or more ground plots corresponding to a cluster. For those clusters which were represented by more than one ground plot, a range of characteristics is provided. By comparing these surface conditions with the ranked spectral values (Table 4.2), the dominant cover class assignments and the visual representations of each cluster (Figure 4.2), one can not only evaluate the spectral quality of the TM digital data for characterizing forest vegetation, but also the adequacy of the clustering algorithm for stratifying, or partitioning, the study area into units, or areas, having similar spectral characteristics correlated to forest parameters of interest.

Detailed comparisons of data products generated from the clustering experiment have yielded the following results:

(1) There is considerable variation in surface characteristics within individual clusters including variation in vegetative cover percentages and in the relief variables of slope gradient (%) and slope azimuth (aspect).

(2) Most surface variability within clusters occurred in areas of low percentage cover of vegetation and light colored soils.

(3) Stands of dissimilar coniferous forest cover types but of similar density were included in the same spectral clusters.

87

(4) Due to the smaller pixel size (28.5m) and the unique spectral characteristics, boundaries between dissimilar cover types and conditions were consistently grouped within four major clusters.

(5) For the most part, major vegetative types (conifer, hardwood, shrub, and meadow) were consistently classified into separate clusters.

The relatively poor performance of the clustering algorithm for separating individual conifer forest cover types can be attributed to the following factors: (1) the default setting of clustering parameters based on MSS experience; (2) the selection of TM bands; and (3) the method used to determine the cluster value at the predicted location of the field plots sampled during 1982. Given the stability of the cluster set after 32 iterations based on the statistics displayed in Figure 4.1, additional iterations would not have significantly improved the quality or number of clusters. Reducing the standard deviation of the normalized cluster (STDMAX) and increasing the minimum distance between clusters (DLMIN) might improve the final cluster set, and should be investigated. This would result in more of the clusters being combined, including many of the conifer clusters which one would want to split even after 32 iterations. This situation could be alleviated by employing two local enhancements to CLUSTER as yet to be tested on digital data of a forest environment. The first includes the clustering parameters INCLUDE, EXCLUDE, and MASK. These options allow further clustering operations on only user-specified clusters while retaining the other clusters in their original state. In this way, the conifer clusters would be INCLUDEd for further clustering, while the bare soil and low density vegetation clusters would be EXCLUDEd from the additional iterations and later combined in the post-processing CHAIN operation. The second enhancement which may improve the clustering of coniferous forest vegetation is the COEFV option in which the coefficients of variation (CV) are used as the split criteria instead of the standard deviations. If this option is selected, a cluster will be eligible for splitting if its CV exceeds STDMAX in any feature. This, again, is an untested enhancement implemented primarily to increase cluster splitting in the "darker" clusters, such as we find in high density conifer stands, and to decrease cluster splitting in the "lighter" clusters, such as the bare soil, low density vegetation, and granitic rock cover types.

The selection of TM spectral bands 2-5 was optimal based on the work of Anuta, et al. (1984) and others; but additional trials using TM 1 (blue) should be conducted based on the positive results of Nelson, et al. (1984) using simulated TM data in northern forests of the United States.

Although we are reasonably confident that the pixels we extracted from the data base corresponded to the location of the field plots sampled during 1982, we must improve the methods used for characterizing the surface conditions and vegetation properties associated with a given cluster. In this study, this characterization was accomplished by detailed interpretation of 1:24,000 scale color photography. This scale was too small to provide the needed detail for the relative small pixel size; further studies would require the acquisition of relatively inexpensive, large-scale (1:1000), small format (35mm or 70mm) vertical color and color infrared photography that has been replicated to provide an adequate sample for each of the clusters. Not only would such a photo set provide the

needed detail for characterization but also would allow us to make detailed photogrammetric estimates, such as tree height, crown closure, stems per acre, and other parameters of interest to forest survey, which could be correlated to the clusters' spectral variability.

In that more clusters were generated than could be efficiently used for stratification purposes, grouping of several "bare soil", shrub, and conifer clusters in a post-processing step would adequately partition the study area into spectrally and ecologically similar units suitable for photo or field plot allocation in a traditional forest inventory. In this regard, the clustering of the four TM bands representing the visible, near reflectance and middle reflectance infrared spectral regions was successful.

4.4 References

Anuta, P.E., L.A. Bartolucci, M.E. Dean, D.F. Lozano, E. Malaret, C.D. McGillem, J.A. Valdés, and C.R. Valenzuela. 1984. Landsat-4 MSS and Thematic Mapper data quality and information content analysis. IEEE Trans. on Geoscience and Remote Sensing, Vol. GE-22, No.3, May. p 222-236.

Ball, G.H. and D.J. Hall. 1966. ISODATA, an iterative method of multivariate analysis of pattern classification. Proc. Int. Communications Conf., Philadelphia, PA. June.

Nelson, R.F. 1981. A comparison of two methods for classifying forestland. Int. J. Remote Sensing 2(1):49-60.

Nelson, R.F., R.S. Latty, and G. Mott. 1984. Classifying northern forests using thematic mapper simulator data. Photogrammetric Eng. Remote Sensing 50(5): 607-617.

Minter, R.T. 1972. ISOCLS - iterative self-organizing clustering program. Computer Program Documentation, CPD#202, Lockheed Electronics Co. Houston, TX.

Ritter, P.R. and A. Kaugars. 1978. CLUSTER Users' Guide, Version 1.0. Remote Sensing Research Program, Department of Forestry and Resource Management, University of California, Berkeley.

C-2
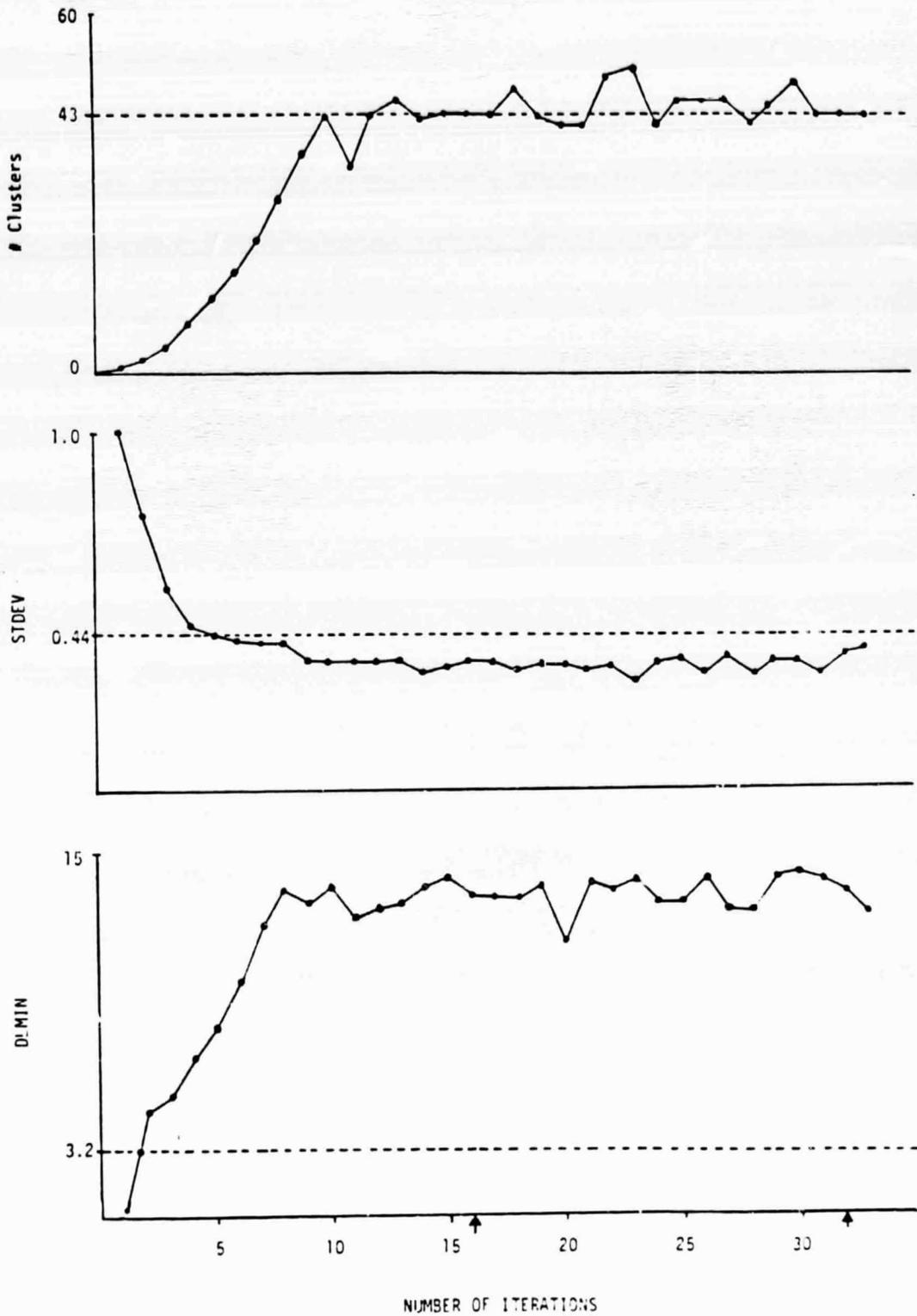
Figure 4.1   Summary statistics, by number of iterations, for clustering four
Thematic Mapper spectral bands acquired by the Landsat-4
spacecraft over a forested environment.
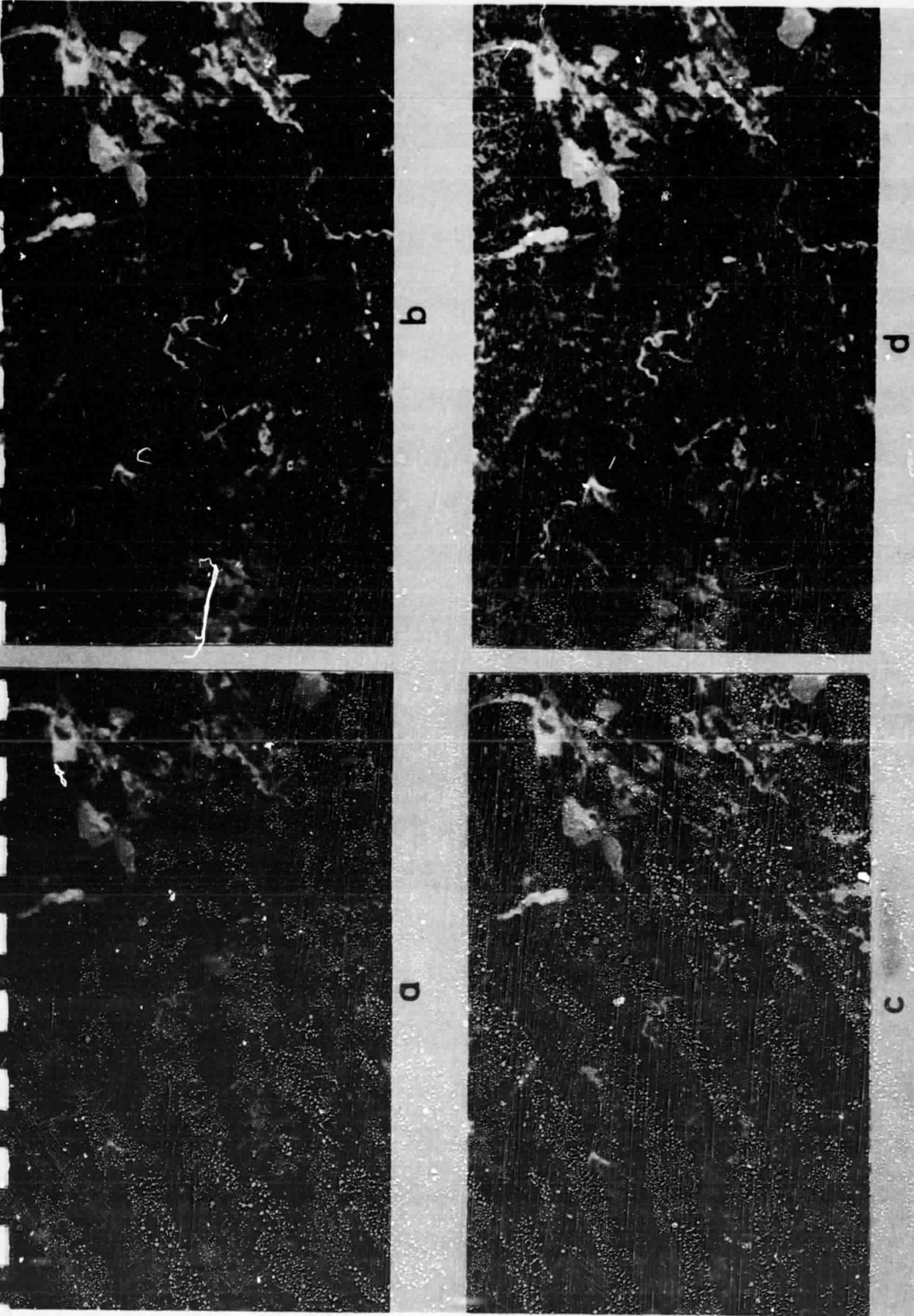
Figure 4.2 Photographic illustrations of the unsupervised clustering results for a portion of the study area. Individual clusters are highlighted by assigning unique color sets which are digitally overlaid on a TM 4, 3, and 2 color composite (A). This procedure allows one to label clusters based not only their tabulated spectral characteristics (Table 4.2 and 4.3) but also on their location within the forest environment. (B) cluster #4 (green) = hardwood/conifer; (C) clusters #5 (green) = shrubs, #19 (blue) = shrubs, #47 (yellow) = high density conifer; (D) clusters #3 (green) = shrubs, #28 (blue) = shrubs, #2 (yellow) = conifer/hardwood.

Table 4.1 Subset of clustering parameters and options used to separate the multidimensional TM data into statistically similar groups, or spectral clusters (Ritter and Kaugars, 1978).

Parameter Setting and Description

INPUT PARAMETERS

NFEAT	4 features; TM Bands 2,3,4 and 5
NORMAL	Normalize data
SEEDST	Use initial subset cluster statistics

CLUSTERING PARAMETERS

DLMIN	3.2; minimum distance between clusters
MAXCLS	60; maximum number of clusters allowed
NMIN	10; delete any cluster with less than 10 data points
STDMAX	0.44; split cluster if normalized standardard deviation exceeds 0.44

CLUSTERING TERMINATORS

ISTOP	15; stop after 15 iterations
PERCENT	5; stop clustering if fewer than 5% of points change cluster in one iteration

OUTPUT PARAMETERS

CHAIN	group clusters whose intercluster distances are less than set limit
INTERIMSTATS	1; number of iterations between printed statistical summary
NOVM	generate a single-band map where each pixel has a cluster value assignment
RANKCLUSTERS	order clusters by a user defined method

Table 4.2 Ranking of the 43 spectral clusters based on five spectral values.

RANK	PHOTOGRAPHIC IR/RED CLUSTER	MEAN LABEL	MIDDLE IR/RED CLUSTER	MEAN	ALBEDO CLUSTER	MEAN	PHOTOGRAPHIC IR CLUSTER	MEAN	MIDDLE IR CLUSTER	MEAN	CLUSTER	POINTS IN CLUSTER
1	51	0.800	51	0.362	51	23.172	51	9.979	51	4.520	1	1239
2	55	1.051	47	1.390	1	47.371	1	33.562	47	22.344	2	34739
3	46	1.098	1	1.416	47	57.761	47	48.675	1	27.000	3	1236
4	15	1.120	55	1.438	58	66.253	37	52.457	4	27.819	4	37600
5	18	1.123	18	1.471	4	75.682	58	52.519	58	31.257	5	4179
6	34	1.212	15	1.583	2	79.403	18	55.126	2	39.291	6	16082
7	49	1.316	4	1.598	7	83.767	31	58.586	8	41.777	7	27898
8	30	1.323	46	1.637	37	86.343	7	59.076	44	48.078	8	23556
9	36	1.328	31	1.688	8	91.165	29	59.982	7	48.370	9	16403
10	42	1.435	52	1.693	29	97.128	2	62.345	9	56.417	10	1644
11	23	1.448	58	1.694	31	97.270	12	63.401	37	56.876	11	4592
12	14	1.466	34	1.772	6	102.421	30	64.477	28	57.051	12	9744
13	52	1.499	54	1.853	9	106.120	6	64.895	31	60.530	13	3286
14	54	1.516	30	1.885	12	108.166	4	65.514	6	61.606	14	1639
15	13	1.604	23	1.938	44	110.042	13	66.021	3	62.896	15	469
16	31	1.633	2	1.953	18	112.639	52	66.132	5	66.568	17	8177
17	29	1.637	49	1.976	17	113.410	14	66.963	17	70.630	18	331
18	53	1.719	37	2.030	52	116.572	34	67.071	12	71.986	19	780
19	10	1.725	10	2.041	11	117.625	53	67.650	18	72.308	20	1962
20	57	1.735	7	2.067	5	122.015	15	68.882	52	74.681	23	804
21	1	1.760	57	2.070	13	122.799	42	70.423	11	77.751	26	1526
22	35	1.863	8	2.096	28	124.822	11	72.160	50	77.988	28	3150
23	37	1.872	13	2.124	26	128.683	23	74.670	26	80.018	29	3727
24	12	1.960	11	2.127	30	129.058	36	75.689	29	82.747	30	669
25	11	1.974	26	2.137	50	129.162	10	76.082	19	83.150	31	2013
26	26	2.288	12	2.226	53	130.894	8	76.180	13	87.441	34	612
27	43	2.296	6	2.250	10	131.584	9	76.269	10	90.014	35	671
28	20	2.319	29	2.258	20	131.876	17	76.720	30	91.866	36	547
29	6	2.370	14	2.281	14	137.023	46	77.680	20	92.209	37	7536
30	7	2.524	9	2.318	34	139.631	35	78.937	43	97.010	42	665
31	17	2.566	44	2.358	23	141.121	20	78.949	15	97.372	43	719
32	58	2.847	17	2.362	3	141.803	49	83.670	34	98.070	44	6909
33	50	2.897	43	2.380	15	142.837	57	84.876	53	98.563	46	495
34	47	3.029	50	2.485	19	142.848	54	85.054	23	99.947	47	20599
35	2	3.099	36	2.486	43	145.290	26	85.670	57	101.234	49	508
36	9	3.134	53	2.505	57	146.639	44	90.163	54	103.996	50	1625
37	19	3.394	5	2.581	35	150.692	50	90.812	14	104.245	51	985
38	5	3.628	28	2.599	54	151.924	5	93.566	35	115.160	52	931
39	4	3.764	42	2.607	42	158.177	43	93.606	46	115.878	53	1974
40	8	3.822	19	2.649	46	161.854	55	95.206	49	125.626	54	487
41	44	4.422	35	2.670	49	169.111	28	105.257	42	127.945	55	272
42	28	4.794	20	2.709	36	174.215	19	106.560	55	130.109	57	676
43	3	5.642	3	2.908	55	195.020	3	122.024	36	141.665	58	32444

93

Table 4.3 Dominant cover class assignments for the 43 clusters, generated using Thematic Mapper data of the Meadow Valley area, as ranked by the mean ratio of the reflectance infrared (TM 4) to the red (TM 3) bands.

Rank (IR/R)	Cluster Number	Dominant Cover Class
1	51	Water
2	55	Mine tailings
3	46	Mine tailings/granitic rock
4	15	Granitic rock
5	18	Granitic rock
6	34	Granitic rock/light bare soil
7	49	Bare soil
8	30	Bare soil
9	36	Bare soil includes low crown cover vegetation
10	42	Bare soil includes low crown cover vegetation
11	23	Bare soil
12	14	Bare soil
13	52	Bare soil
14	54	Bare soil
15	13	Bare soil with some sparse vegetation
16	31	Bare soil with some sparse vegetation
17	29	Conifer, low density; low understory vegetation
18	53	Bare soil/sparse vegetation cover
19	10	Bare soil/sparse vegetation cover
20	57	Bare soil/border pixels with conifer
21	1	Bare soil/border pixels with conifer
22	35	Bare soil/border pixels with conifer
23	37	Conifer, low density: logged areas and peridotite soils
24	12	Conifer, low density: logged areas and peridotite soils
25	11	Conifer, low density with shrub understory
26	26	Meadow
27	43	Meadow
28	20	Meadow
29	6	Conifer, low density
30	7	Conifer/Shrub: border pixels between classes
31	17	Conifer, low density
32	58	Conifer, mature stands
33	50	Shrub, low density
34	47	Conifer, dense stands on north slopes
35	2	Mixed conifer/black oak communities
36	9	Shrub -- manzanita
37	19	Shrub -- willow, alders
38	5	Shrub -- manzanita
39	4	Hardwood/conifer, moist slopes and draws
40	8	Shrub -- manzanita, scrub oak
41	44	Shrub -- manzanita, scrub oak
42	28	Shrub -- dogwood, willow (moist sites)
43	3	Shrub -- willow, dogwood (very moist sites)

Table 4.4 Range of surface characteristics for 22 of the 43 clusters which occurred at one or more of the field plots surveyed during 1982.

SURFACE CHARACTERISTICS

Cluster Number	% Vegetation Cover	% Mature Conifer	% Young Conifer	% Hardwood	% Shrub	% Slope	Aspect	Elevation (in 100 ft)	Number of Ground Plots
31	50	10	0	0	40	40	S	48	1
10	100	60	0	20	20	52	NE	46	1
13	30-90	5-90	0-10	0-5	0-15	5-16	N,SW	38-42	2
36	100	0	0	0	100	2	NW	38	1
42	70	10	0	0	60	0	FLAT	37	1
1	80	50	15	0	15	65	N	38	1
20	100	0	0	0	100	0	FLAT	38	1
12	50	0-50	0-10	0-5	0-45	3-60	N,W,S	38-46	4
37	50-70	20-30	10-50	0	0-10	4-12	NE,NW,S	38-39	3
6	30-100	0-60	0-60	0-5	5-10	11-62	-(N,W)	38-55	15
7	10-100	5-60	0-70	0-5	0-50	0-60	-E	39-53	12
17	50-65	20	20	0-5	10-20	30-33	S,SW	46-49	2
58	60-100	0-70	0-60	0-5	0-70	0-50	-(S,SE,E)	37-49	13
50	0-80	0-50	0	0-5	0-50	0-55	FLAT,SE	39-53	3
47	80-100	0-70	10-30	0-100	0-20	12-58	SW,NW,N,,NE	38-58	7
2	50-100	0-50	0-70	0-5	0-40	0-53	ALL	38-57	14
9	50-90	0-40	0-70	0-5	0-80	5-60	-(W,S)	38-52	9
4	60-100	0-80	0-50	0-5	0-60	2-58	-S	38-57	19
8	50-100	0-60	0-80	0-5	0-100	3-51	S,SE,E,NE	42-56	12
44	50-100	0-75	0-30	5	20-50	35-50	S,NW,SE	48-61	3
28	50-100	0	0-30	0	0-100	17-34	S,SE	54-56	2
26	30	0	0	0	30	0	FLAT	37	1

MULTISPECTRAL CLASSIFICATION OF THEMATIC MAPPER AND
MULTISPECTRAL SCANNER DATA FOR AGRICULTURAL SURVEY USING
BAYESIAN AND CONTEXTUAL TECHNIQUES

S.D. DeGloria
Remote Sensing Research Program
University of California - Berkeley

Ralph Bernstein
IBM-Palo Alto Scientific Center
1530 Page Mill Road
Palo Alto, California

Silvano DiZenzo
IBM-Rome Scientific Center
Via del Giorgione 129
Rome, Italy

5.1 Introduction

The Landsat Earth observation program began in 1972 with the launch of the first in a series of Landsat satellites. These earlier spacecraft contained both a multispectral scanner (MSS) and a return beam vidicon (RBV) camera system. The MSS offered the first opportunity for obtaining repetitive, sun-synchronous, and synoptic multispectral digital data of the Earth from space. Based on several years of basic and applied research, a second generation electro-optical sensor was developed and placed in orbit with the MSS on board the Landsat-4 and Landsat-5 spacecrafts. This new sensor, called the Thematic Mapper (TM), includes many significant improvements over the MSS sensor, the most significant being improved spectral, spatial, and radiometric resolution. Table 5.1 shows the changes that have occurred in the Landsat systems since 1972; Table 5.2 shows the spectral bands and actual bandwidths for both the MSS and TM sensors on the Landsat-4 spacecraft.

These improvements in the spectral, spatial, and radiometric characteristics of the TM sensor provide several opportunities to perform experiments for determining the degree to which these changes may improve classification accuracies over those expected from the MSS sensor.

Historically, the Landsat system has provided spectral data in digital and photographic formats for use in many mapping endeavors. In agricultural surveys, Landsat data are used in such a way as to obtain maximum benefit of both the multispectral and multitemporal nature of the data. Many of the problems encountered in using the MSS data from the early Landsats are related to pixel size unsuitable for discriminating small fields resulting in poor classification results. These poor results stemmed from the high proportion of "mixed pixels", or pixels which represent the composite

spectral signature from the field of interest, adjacent fields, and cultural features. In addition, limited spectral coverage and broad spectral bands where radiant energy was encoded into one of 64 levels (6-bit) for each band contributed to this poor classification performance. In order to increase the number of cover conditions (or "themes") which could be discriminated by a spaceborne sensor and to hopefully increase classification accuracy for each of these cover type conditions, the Thematic Mapper sensor was developed where: (1) the pixel size was reduced to minimize this boundary, or mixed, pixel problem; (2) the number of spectral bands was increased and the band width decreased; and (3) the data were quantized to 256 levels (8-bit per band.)

To determine the degree of improvement in correctly classifying selected agricultural crops in California using the increased spectral, spatial, and radiometric resolution of the Thematic Mapper sensor, multispectral classifications of TM and MSS data using both a standard Bayesian maximum likelihood algorithm and a contextual algorithm of the relaxation type were performed on a single-date Landsat-4 data set acquired on 12 August 1983 over Yolo County, California. Because spectral variability within a field increases with a decrease in pixel size as more within-field agronomic variations are discriminated by the sensor, the use of contextual classification algorithms in concert with standard maximum likelihood classifiers should provide improved classification accuracy. These improved accuracies should be realized when the spatial distribution of the classified pixels is used to further determine the class to which an adjacent, or set of pixels, belongs.

Using both a standard Bayesian maximum likelihood algorithm and a discrete relaxation classification technique, we evaluated the following:

(1) the degree of improvement in classification accuracy as a result of increased spatial resolution while keeping spectral regions constant (green, red, and near reflectance infrared);

(2) the degree of improvement in classification accuracy as a result of increased spectral range and resolution while keeping the spatial resolution constant (57.0m);

(3) the degree of improvement in classification accuracy as a result of the combined effects of increased spectral and spatial resolution by the TM sensor; and

(4) the degree of improvement in classification accuracy as a result of the classification method used (Bayesian, contextual) for each of the above conditions.

5.3 Experimental Methods

Several system parameters and analysis methods are available for control in evaluating classification accuracies. These include the number of bands, location and width of spectral classification algorithms, and pixel size. By suitably varying the controllable parameters, some insight can be gained into the effect of the sensor on classification accuracies. Different algorithms are used in order to determine their relative performance and provide some degree of checking results. Clearly, the data

97

set to be selected is of great importance. We have chosen one for which there was complete ground truth, and for which there existed simultaneous MSS and TM coverage. Thus, the atmospheric, temporal, scene, and solar illumination conditions were identical for the MSS and TM data.

Table 5.3 shows the experimental conditions we used. Note that an attempt was made to conduct evaluations with a minimum number of variables, so that we could quantitatively evaluate the effect of each variable. For example, by selecting the three TM bands most similar to the four MSS bands for a classification run, we were able to assess the effect of the different spatial resolution of the two systems on the classification accuracy.

Study Site

All of the Landsat and ground data that were collected and analyzed for this study came from agricultural fields near the town of Winters in Yolo County, California. In this area, cropping practices were diverse, and field sizes were large enough so that a relatively large number of "pure" TM and MSS pixels could be extracted from a given field for both training and testing purposes.

Spectral Data Set

The spectral data set used in this experiment was acquired on 12 August 1983 by the Landsat-4 spacecraft. Both the TM and MSS sensor were operating at the time of acquisition, and atmospheric conditions over the study site were excellent. The data were processed by the National Aeronautics and Space Administration (NASA), Goddard Space Flight Center (GFC), and received by these investigators as "P" tapes. Both MSS (CCT-PM) and TM (CCT-PT) data in this format have been geometrically and radiometrically corrected, which results in a 57.0m MSS pixel and a 28.5m TM pixel.

Ground Data Set

In order to document the ground cover and conditions prevalent at the time of the Landsat-4 overpass, ground survey data were collected over a large area by a number of individuals. In the agricultural areas, 234 fields were visited, and data recorded pertaining to cover type, canopy height, percent ground cover, growth stage, row direction, degree of weediness, and relative leaf and soil moisture conditions. Of the 17 cover types identified, five major crops were selected for evaluation: corn, rice, alfalfa, sugar beets, and tomatoes. A sixth category , called "background", was used in our classification test and included all other crop or cover conditions prevalent at the time of the overpass. A sufficient number of fields were surveyed for each of these cover types so that independent sets of training and testing fields could be generated.

Aerial Photographic Data

In addition to the Landsat spectral data and the detailed field surveys, large-scale (1:4000), small format (35mm) color and color infrared oblique aerial photography was also collected coincident with the overpass. The aerial photography was acquired to document within-field variations in canopy cover that the ground survey teams were unable to document. In addition, the aerial photography served to broaden the data base to include

98

a number of fields not visited on the ground.   Though cover types and cover conditions could be easily interpreted from the aerial photography, only those fields that had been documented by the ground survey teams were used to develop the training and test field statistics for the classification algorithm.

Data Processing

A 1024x1024 TM pixel area and a 512x512 MSS pixel area were extracted from the digital tapes received from NASA.   These subscene areas contained the set of fields used in our analysis. The MSS data were not geometrically registered to the TM data, but kept as a separate data file as extracted from the CCT-PM.   To simulate the 57.0m MSS pixel, the pixel from every other point and line of the TM data were sampled.

Training fields were selected from the list of available fields in the study area, located cn the available maps and ground survey sheets, and extracted from the previously created data files.   The fields were first located in the TM set using the color monitor display of the IBM 7350 image processing system.   Reference was made to the field maps, data sheets, and the oblique aerial photography in order to confirm the cover type and location of each field. The interior set of pixels of each field to be used for training and testing was outlined interactively using the trackball and cursor on the 7350 system.   The extraction of the corresponding MSS pixels was performed visually by locating the same fields on the color display and using the cursor to delimit essentially the same interior portion of each field.   The digital count values for each TM and MSS spectral band were extracted for all training fields, and summary statistics, including the means and covariance matrices, were calculated.   Six multivariate normal density functions were defined using the training field statistics for each crop type, sensor, pixel size, and spectral coverage test.

An independent set of test fields was selected from the list of fields visited on the ground. The sets of pixels representing the interior regions of each field were used to conduct the classification test outlined above.

Once the appropriate statistical functions were defined using the training field data, the classification of the test fields was performed using a Bayesian maximum likelihood algorithm, viz., equal a-priori probabilities with a threshold of 0.30.   Each pixel in each test field was assigned to one of six classes (corn, rice, alfalfa, sugar beets, tomatoes, or background) using the maximum likelihood decision rule with a threshold setting of 0.30 for all classification tests.

To improve the pixel classification rates by considering the spatial dependence, or context, of each classified pixel, a discrete relaxation algorithm using higher order constraints was applied to each of the data sets where every pixel had been assigned a class code.   The relaxation algorithm calculates a new probabilistic class assignment based on the old, or Bayesian-generated, class assignment and a quantitative representation of the spatial dependence among the classified pixels.

For the present experiment, the following set of spatial parameters, or constraints, was used to control the process of changing pixel assignments using the discrete relaxation algorithm.   The first parameter was set to

evaluate whether or not a pixel should be kept as a candidate for the "background" class. We set this parameter to "4" meaning that at least four neighboring pixels should also be candidates for the "background" class; if this condition is not satisfied, the pixel in question is no longer considered to a be a candidate for this class. The second parameter was set to evaluate whether or not a pixel in one of the remaining five crop classes should be kept as a candidate for that given class, "k". We set this parameter to "2", meaning that in order to keep a pixel as a candidate to class k (where k≠0, and k is not the "background" class), at least two neighboring pixels should still be candidates for class k. If this condition was not satisfied, the pixel was no longer considered to be a candidate for class k. The third parameter was set to establish the number of iterations to be used in order to assign a pixel to one of the six classes of interest; the number used in this experiment was nine. For zero iterations, the discrete relaxation method reduced to simple multispectral thresholding where the probability that a pixel belongs in a given class is either "0" or "1". Probabilistic relaxation, which uses floating point probabilities, is more computationally expensive, requires more disk storage, and cannot be implemented in the hardware of special purpose image processors which do not have the capability for floating point arithmetic. Previous work by a number of investigators, however, has found that the use of probabilistic relaxation operators as a post processing step in pixel classification yields significant improvements in classification accuracy (Davis, et al., 1983). We have applied a more computationally simplistic relaxation algorithm in this study where image processing systems which employ fixed point arithmetic and Boolean functions implemented in hardware can be used to advantage.

To evaluate both the Landsat-4 sensor systems and the performance of the two algorithms, error matrices were generated for each classification test listed in Table 5.3. Each matrix was constructed by tabulating both the true number of pixels in each cover class, as documented by field survey and image processing, and the number of pixels in each cover class as calculated by either the Bayesian or contextual algorithm. The matrix is an effective method for showing the number of pixels correctly and incorrectly identified. Statistical evaluation of data in error matrices included calculation of omission and commission errors, overall classification accuracy, classification accuracy by crop type, and inter-matrix agreement using the nonparametric KHAT statistic (Congalton, et al., 1981).

5.4 Results

Individual error matrices for each of the eight classification tests are shown in Tables 5.4 to 5.11; the classification accuracy and the errors of commission for each of the six cover types are shown in Table 5.12; and the statistical ranking of each classification test is shown in Table 5.13 where the calculated KHAT value and level of significance are tabulated along with the overall classification accuracy. These results are graphically represented in Figure 5.1 where total percent correct interpretation and commission errors are plotted for each crop type and classification method tested.

Overall classification accuracy was highest (77.6%) for the TM sensor using all seven bands at full spatial resolution (28.5m pixels) and employing the Bayesian maximum likelihood algorithm; and the lowest (57.7%)

using all four bands of MSS (57.0m pixels) and employing the discrete relaxation algorithm. Additional research is being conducted using probabilistic relaxation methods in order to evaluate the effect of higher order constraints on classification performance. Several significant results can be interpreted from these tables which effectively summarize the results of this experiment.

(1) The TM data, whether at full or degraded spatial resolution, achieved higher classification accuracies than the MSS data using comparable spectral bands and either classification approach.

(2) Use of all TM spectral bands significantly improved classification results over the use of only those TM bands corresponding to the principal MSS spectral bands.

(3) Classification accuracies obtained by using single-date TM data meet or exceed accuracies obtained using multi-date MSS data for some of the crops studied (Thomas, et al., 1984).

(4) The Bayesian maximum likelihood algorithm achieved higher classification accuracies than the relaxation algorithm in all tests conducted due to the greater power of the Bayesian method which uses floating point probabilities.

(5) Use of an iterative process in the relaxation algorithm reduced classification error from that achieved by using no iterations (thresholding).

(6) The use of a contextual classification algorithm which does not use floating point arithmetic, and hence is easily implemented on low to moderately priced image processing systems, provides classification accuracies competitive with maximum likelihood algorithms for several of the crop types evaluated in this experiment.

Based on the results obtained in this series of experiments, several areas of further research and development can be identified. These include determining: (1) the effect of varying the values of the higher order relaxation constraints on classification performance, (2) the effect of removing the thermal band on classification accuracy, and (3) the degree of classification improvement by using a probabilistic relaxation algorithm and TM data for the same study area.

5.5 Acknowledgements

5.6 References

Congalton, R.G., R.A. Mead, R.C. Odewald, and J. Heinen. 1981. Analysis of forest classification accuracy. Cooperative Res. Rept., U.S. Forest Service and Virginia Polytechnic Institute and State University, #13-1134. Blacksburg, Virginia 24061.

Davis, L.S., C.Y. Wang, and H.C. Xie. 1983. An experiment in multispectral, multitemporal crop classification using relaxation techniques. Computer Vision, Graphics, and Image Processing 23, 227-235.

Thomas, R.W., S.L. Wall, C.E. Brown, and L.H. Beck. 1984. California Landsat-aided crop estimation. Final Report, U.S.D.A. Cooperative Agreement. Space Sciences Laboratory, University of California (in press).
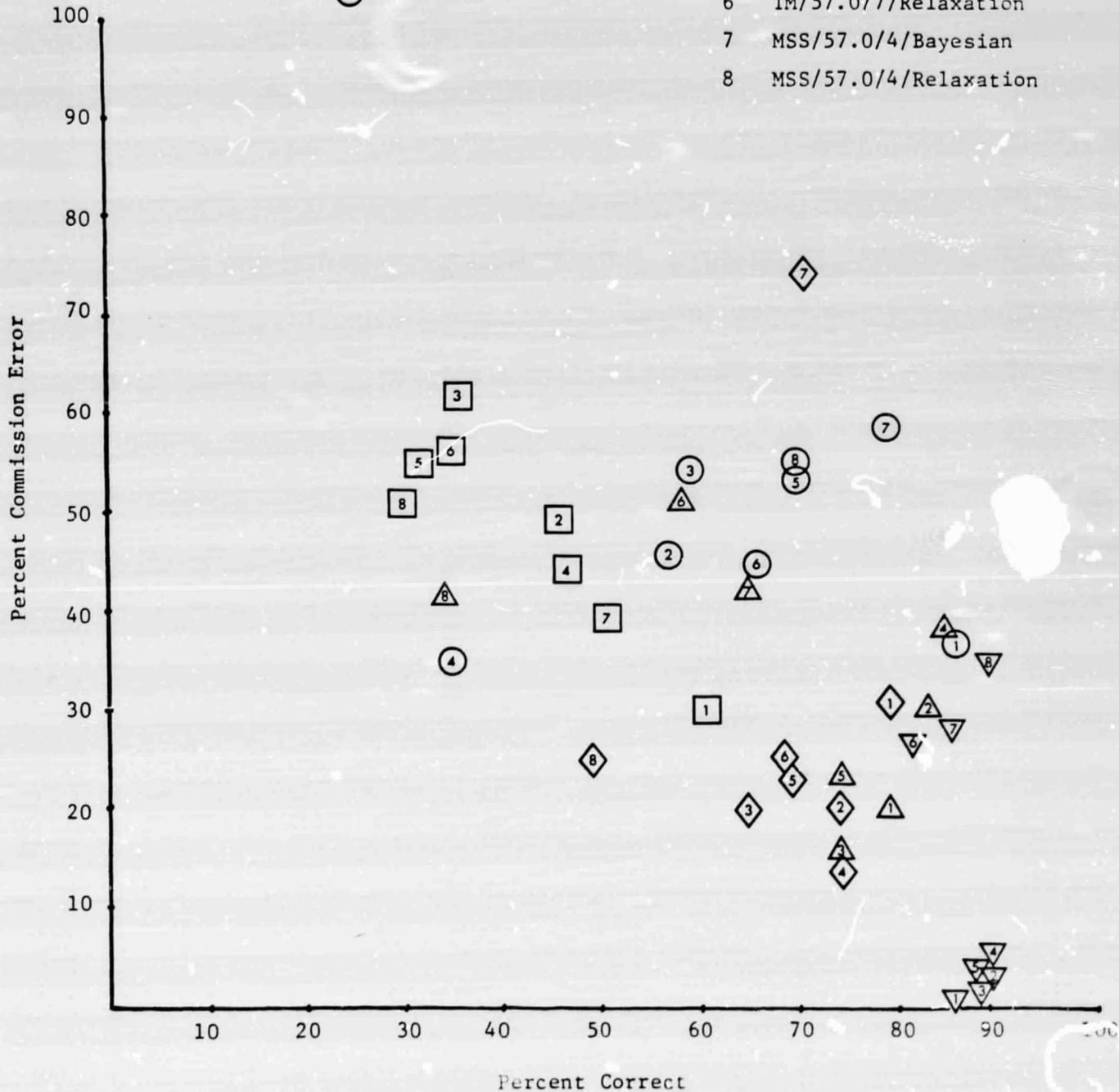
Figure 5.1 Graph of percent correct by crop type vs. percent commission error for each classification test. Each symbol represents a crop type and the number inside the symbol represents the classification test performed.

103

Table 5.1 Changes in the characteristics of the Landsat multispectral sensors.

Landsat	Year	Sensor	Resolution	Bands	Bits/Sample	Bits/Scene
1	1972	MSS	79 m	4	6	1.8 x E08
2	1975	MSS	79 m	4	6	1.8 x E08
3	1978	MSS	79 m	5	6	1.9 x E08
4	1982	TM	30 m	7	8	18.4 x E08
		MSS	81 m	4	6	1.8 x E08
5	1984	TM	30 m	7	8	18.4 x E08
		MSS	81 m	4	6	1.4 x E08

===

Table 5.2. Spectral bands and actual bandwidths of both the TM and MSS sensors on Landsat-4.

SENSOR	BAND	SPECTRAL REGION	BANDWIDTH
MSS	1	Green	0.50-0.61μm
	2	Red	0.60-0.70μm
	3	Near Reflectance Infrared	0.70-0.81μm
	4	Near Reflectance Infrared	0.81-1.02μm
TM	1	Blue	0.45-0.52μm
	2	Green	0.53-0.61μm
	3	Red	0.62-0.69μm
	4	Near Reflectance Infrared	0.78-0.90μm
	5	Middle Reflectance Infrared	1.57-1.78μm
	6	Far Reflectance Infrared	2.10-2.35μm
	7	Thermal Infrared	10.42-11.66μm

===

Table 5.3 Experimental conditions for the classification accuracy assessment.

	SPECTRAL BANDS		SPATIAL RESOLUTION		CLASSIFICATION ALGORITHM
	TM	MSS	TM	MSS	
Spatial	2-4	1-4	28.5m	57.0m	Bayesian and Contextual
Spectral	1-7	1-4	57.0m	57.0m	Bayesian and Contextual
Spatial/Spectral	1-7	1-4	28.5m	57.0m	Bayesian and Contextual

Table 5.4 Classification accuracy error matrix for the TM sensor, 28.5m pixels, seven spectral bands, and the Bayesian maximum likelihood algorithm.

| | | | Ground Cover Class (field verified) | | | | | |
	B	C	R	A	SB	T	TOTAL	% COMMISSION
B	3625	270	484	99	118	111	4707	23.0
C	264	1202	28	1	3	0	1498	19.8
R	1	46	3229	0	0	0	3276	1.4
A	164	1	0	923	60	178	1326	30.4
SB	273	0	0	304	1369	25	1971	30.5
T	720	2	24	177	194	1905	3022	37.0
TOTAL	5047	1521	3765	1504	1744	2219	15800	
% OMISSION:	28.2	21.0	14.2	38.6	21.5	14.2		

Landsat Classification

n = 15,800 test pixels

Overall Accuracy = $\frac{12253}{15800}(100)$ = 77.6%

Table 5.5 Classification accuracy error matrix for the TM sensor, 28.5m pixels, three spectral bands, and the Bayesian maximum likelihood algorithm.

Ground Cover Class (field verified)

Landsat Classification	B	C	R	A	SB	T	TOTAL	% COMMISSION
B	4003	174	289	187	98	480	5231	23.5
C	423	1258	94	11	3	0	1789	29.7
R	51	69	3362	0	0	0	3482	3.4
A	42	1	2	684	162	466	1367	49.2
SB	35	1	0	275	1288	12	1611	20.0
T	493	18	18	337	193	1261	2320	45.6
TOTAL	5047	1521	3765	1504	1744	2219	15800	
% OMISSION:	20.7	17.3	10.7	53.9	26.2	43.2		

n = 15,800 test pixels

Overall Accuracy = $\frac{11866}{15800}(100) = 75.1\%$

107

Table 5.6 Classification accuracy error matrix for the TM sensor, 28.5m pixels, three spectral bands, and the Discrete Relaxation algorithm.

Landsat Classification

| | Ground Cover Class (field verified) | | | | | | TOTAL | % COMMISSION |
	B	C	R	A	SB	T		
B	3869	332	369	423	317	73	5383	28.1
C	190	1122	45	17	0	0	1374	18.3
R	22	54	3312	0	0	0	3388	2.2
A	29	0	0	542	39	810	1420	61.8
SB	4	0	0	251	1125	24	1404	19.9
T	933	13	39	271	263	1312	2831	53.7
TOTAL	5047	1521	3765	1504	1744	2219	15800	
% OMISSION:	23.3	26.2	12.0	64.0	35.5	40.9		

n = 15,800 test pixels

$$\text{Overall Accuracy} = \frac{11282}{15800}(100) = 71.4\%$$

Table 3.7 Classification accuracy error matrix for the TM sensor, 57.0m pixels, seven spectral bands, and the Bayesian maximum likelihood algorithm.

Landsat Classification

		Ground Cover Class (field verified)							% COMMISSION
	B	C	R	A	SB	T	TOTAL		
B	922	34	89	107	43	323	1518		20.1
C	216	417	22	8	6	4	673		38.0
R	24	32	875	0	0	0	931		6.0
A	3	1	0	181	34	102	321		43.6
SB	0	0	0	56	374	10	440		15.0
T	28	8	2	36	50	233	357		34.7
TOTAL	1193	492	988	388	507	672	4240		
% OMISSION:	22.7	15.2	11.4	53.4	26.2	65.3			

n = 4240 test pixels

$$\text{Overall Accuracy} = \frac{3002}{4240}(100) = 70.8\%$$

109

Table 5.8 Classification accuracy error matrix for the TM sensor, 28.5m pixels, seven spectral bands, and the Discrete Relaxation algorithm.

Ground Cover Class (field verified)

Landsat Classification	B	C	R	A	SB	T	TOTAL	% COMMISSION
B	3557	309	387	402	262	121	5038	29.4
C	266	1126	44	30	0	0	1466	23.2
R	38	70	3297	0	0	0	3405	3.2
A	32	0	1	474	40	496	1043	54.6
SB	23	2	0	282	1204	42	1553	22.5
T	1131	14	36	316	238	1560	3295	52.7
TOTAL	5047	1521	3765	1504	1744	2219	15800	
% OMISSION:	29.5	26.0	12.4	68.5	31.0	29.7		

n = 15,800 test pixels

Overall Accuracy $= \frac{11218}{15800}(100) = 71.0\%$

Table 5.9 Classification accuracy error matrix for the TM sensor, 57.0m pixels, seven spectral bands, and the Discrete Relaxation algorithm.

Ground Cover Class (field verified)

	B	C	R	A	SB	T	TOTAL	% COMMISSION
B	704	45	64	56	64	63	996	29.3
C	105	287	86	52	30	22	582	50.7
R	147	147	812	0	0	0	1106	26.6
A	6	7	2	137	45	114	311	56.0
SB	0	1	1	81	344	31	458	24.9
T	230	6	24	61	24	442	787	43.8
TOTAL	1192	493	989	387	507	672	4240	
% OMISSION:	40.9	41.8	17.9	64.6	32.2	34.2		

Landsat Classification

n = 4240 test pixels

Overall Accuracy $= \frac{2726}{4240}(100) = 64.3$

111

Table 5.10 Classification accuracy error matrix for the MSS sensor, 57.0m pixels, four spectral bands, and the Bayesian maximum likelihood algorithm.

Ground Cover Class (field verified)

	B	C	R	A	SB	T	TOTAL	% COMMISSION
B	124	10	7	10	15	13	179	30.7
C	102	324	92	17	14	13	562	42.4
R	153	134	769	3	1	1	1061	27.5
A	4	6	2	229	54	78	373	38.6
SB	10	15	6	141	373	1	546	74.2
T	402	6	15	51	66	399	939	57.5
TOTAL	795	495	891	451	523	505	3660	
% OMISSION:	84.4	34.6	13.7	49.2	28.7	21.0		

Landsat Classification

n = 3660 test pixels

Overall Accuracy = $\frac{2218}{3660}(100)$ = 60.6%

112

Table 5.11 Classification accuracy error matrix for the MSS sensor, 57.0m pixels, four spectral bands, and the Discrete Relaxation algorithm.

| | Ground Cover Class (field verified) | | | | | | TOTAL | % COMMISSION |
	B	C	R	A	SB	T		
B	410	53	16	128	120	24	751	45.4
C	27	169	62	6	12	17	293	42.3
R	153	255	791	14	7	0	1220	35.2
A	4	3	1	133	30	102	273	51.3
SB	2	3	0	72	257	10	344	25.3
T	199	13	21	97	97	352	779	54.8
TOTAL	795	496	891	450	523	505	3660	
% OMISSION:	48.4	65.9	11.2	70.4	50.9	30.3		

Landsat Classification

n = 3660 test pixels

Overall Accuracy = $\frac{2112}{3660}(100)$ = 57.7%

113

Table 5.12 Percent correct (COR) and percent commission (COM) errors by ground cover class and total percent correct for the eight classification tests.

CLASSIFICATION TEST

GROUND COVER CLASS		TM 28.5/7 Bayesian	TM 28.5/3 Bayesian	TM 28.5/3 Relaxation	TM 57/7 Bayesian	TM 28.5/7 Relaxation	TM 57/7 Relaxation	MSS 57/4 Bayesian	MSS 57/4 Relaxation
Corn	COR:	79.0	82.7	73.8	84.8	74.0	58.2	65.4	34.1
	COM:	19.8	29.7	18.3	38.0	23.2	50.7	42.4	42.3
Rice	COR:	85.8	89.3	88.0	88.6	87.6	82.1	86.3	88.8
	COM:	1.4	3.4	2.2	6.0	3.2	26.6	27.5	35.2
Alfalfa	COR:	61.4	46.1	36.0	46.6	31.5	35.4	50.8	29.6
	COM:	30.4	49.2	61.8	43.6	54.6	56.0	38.6	51.3
Sugarbeets	COR:	78.5	73.8	64.5	73.8	69.0	67.8	71.3	49.1
	COM:	30.5	20.0	19.9	15.0	22.5	24.9	74.2	25.3
Tomatoes	COR:	85.8	56.8	59.1	34.7	70.3	65.8	79.0	69.7
	COM:	37.0	45.6	53.7	64.7	52.7	43.8	57.5	54.8
Background	COR:	71.8	79.3	76.7	77.3	70.5	59.1	15.6	51.6
	COM:	23.0	23.5	28.1	20.1	29.4	29.3	30.7	45.5
TOTAL % CORRECT:		77.6	75.1	71.4	70.8	71.0	64.3	60.6	57.7
N (PIXELS):		15,800	15,800	15,800	4,240	15,800	4,240	3,660	3,660

Table 5.13 Statistical ranking, level of significance, and overall accuracy of each classification test.

Rank		$\hat{K}$ and Significance*	Percent Correct Mean
1	TM, 28.5m pixels, 7 spectral bands, Bayesian	.7187┤	77.6
2	TM, 28.5m pixels, 3 spectral bands, Bayesian	.6851┤	75.1
3	TM, 28.5m pixels, 3 spectral bands, Relaxation	.6372┤	71.4
4	TM, 57.0m pixels, 7 spectral bands, Bayesian	.6337	70.8
5	TM, 28.5m pixels, 7 spectral bands, Relaxation	.6336┤	71.0
6	TM, 57.0m pixels, 7 spectral bands, Relaxation	.5581┤	64.3
7	MSS, 57.0m pixels, 4 spectral bands, Bayesian	.5246┤	60.6
8	MSS, 57.0m pixels, 4 spectral bands, Relaxation	.4788┤	57.7

* The lines to the right of the ranked Kappa values indicate those classification tests which are not significantly different at the 95 percent confidence level.

115

Chapter 6

VISUAL COMPARISON OF SIMULATED SPOT IMAGERY WITH
LANDSAT-4 THEMATIC MAPPER AND MULTISPECTRAL SCANNER
IMAGERY FOR THE INTERPRETATION OF FOREST RESOURCES

R.N. Colwell and S.D. DeGloria
Remote Sensing Research Program

6.0   Introduction

Based on the desire to enhance spectral data acquisitions from space, France initiated a project in 1977 to develop a space-based terrestrial observation satellite system which will become a component of the French National Space Program when the satellite is launched in 1985.   This satellite system, called SPOT (Systeme Probatoire d'Observation de la Terre), is being developed by the French Centre National d'Etudes Spatiales (CNES) in cooperation with Belgium and Sweden. The design objectives of the satellite system are twofold: (1) provide image products, in CCT or film format, with high spatial resolution using nadir or off-nadir viewing ($\pm 27^\circ$ across-track); and (2)   through the use of images acquired in adjacent orbits to produce stereoscopic image pairs of selected areas for assisting resource survey operations.   The major attributes of this Earth observation satellite are shown in Table 6.1.

The SPOT satellite will carry two high resolution visible (HRV) range sensors employing multispectral linear array technology. These sensors will record reflected electromagnetic energy in either a multispectral or panchromatic mode using the nadir, off-nadir, or stereoscopic format.   The linear arrays in the focal plane of the instrument are electronically sampled for across-track coverage, and the satellite orbital movement provides along-track coverage. Multispectral imaging in the visible and near reflectance infrared region and panchromatic imaging in the visible region will be accomplished using three narrow spectral bands (100nm) and one broad band (220nm), respectively.   The major characteristics of the sensors are summarized in Table 6.2.

During 1983, simulated SPOT data were acquired over selected sites on a global basis to: (1) enable potential users to evaluate SPOT data within their application field, (2) provide the value-added industry an opportunity for product development, and (3) serve as an aid to SPOT IMAGE Corporation in perceiving and responding to the needs of the user community in the United States.   The United States SPOT Simulation Campaign acquired multispectral and panchromatic data over 61 sites nationwide using the Daedalus AADS1268 Multispectral Scanner mounted on-board a Learjet 25-C.   In addition to the multispectral scanner, a Zeiss RMK A 15/23 metric camera with a 153 mm (6in) focal length lens acquired stereoscopic photography (60 percent forward overlap) of each site using Kodak Aerochrome Infrared film, Type 2443.

One of the anticipated uses of SPOT data, in either digital or film format, is to improve the efficiency of renewable resource inventories and assessments.   In the field of forestry, multispectral data from aircraft and

spacecraft are used for three interrelated activities: mapping, inventory, and monitoring. The purpose of forest mapping is to delimit and quantify the spatial distribution of available resources. In terms of forest inventory, SPOT data can possibly be used to advantage in the stratification phase in which resource categories of interest are grouped for increased sampling efficiency. Forest monitoring attempts to estimate changes in the resources over a selected period of time to include changes in selected vegetative growth, ingrowth of secondary species of commercial and non-commercial importance, plant mortality, tree crop harvest, and loss or movement of vegetative components between forest stands over long periods of time (Wensel, 1977).

The objective of the SPOT simulation study at the University of California - Berkeley was to evaluate the utility of SPOT high resolution, multispectral film products for: (1) the stratification of forest cover types based on species composition, percent crown closure, and tree or stand size class based on average tree crown diameter; (2) the delineation of soil mapping units based on vegetation type and density, parent-material, topographic relief variables, and drainage patterns; and (3) the enhancement of land management data bases or information systems to support the production of such maps as timber site suitability, transportation routes and accessibilty, and harvest planning. The results of this study related specifically to the evaluation of simulated SPOT image products has been reported by DeGloria (1984).

Because of the successful acquisition of both simulated SPOT and Landsat-4 data over our Plumas National Forest study site (see Figure 1.1 and Section 2.3.2, respectively, for the location and a description of the study site) in the summer of 1983 the Principal Investigator of this LIDQA project (Professor R.N. Colwell) requested additional funds from the Landsat Science Office in order to investigate the relative quality of these image products for the interpretation of forest resources. Because the request was favorably endorsed, this chapter presents the results of our recently conducted evaluation.

6.1 Experimental Methods

Data Set Description

The simulated SPOT data were acquired as a "P" site for our forestry study area on 25 June 1983 (Table 6.3). Of the data products received, only the high resolution (10m IFOV pixels), multispectral film product was qualitatively evaluated using manual interpretation techniques (Figure 6.1). This image was generated by the SPOT IMAGE Corporation by combining the 3-band multispectral data with the broad band panchromatic data of higher spatial resolution.

The Landsat-4 imagery were acquired on 12 August 1983 and utilized by these investigators as film products generated from the CCT-PT data. Four color composite images were generated using both the TM and MSS data: TM 4, 3, and 2; TM 5, 3, and 2; TM 5, 4, and 3; and MSS 4, 2, and 1. Each band was color coded on the RSRP's interactive image display using the red, green, and blue color guns, respectively, and then recorded on Ektachrome film using the Matrix color graphics camera.

117

In addition to these data, aerial oblique and ground photographs were acquired at selected sites previous to and coincident with the overpass. Historical field data (August 1982) collected from 130 plots were consulted for developing a quantitative understanding of the vegetation and soil conditions in the area imaged by the sensors on the aircraft.

Interpretation Methods

Using the image characteristics of tone, texture, pattern and association, all of the images were interpreted to determine the extent to which the following four major forest cover conditions could be discriminated:

A = non-commercial, mature forest stands having large sawlog size timber, low crown closure percentage, and sparse understory.

B = young forest stands having small sawlog size timber, low crown closure percentage, and dense shrub understory.

C = mature forest stands, having large sawlog size timber, medium crown closure percentage, and sparse shrub understory.

D = mature forest stands, having large sawlog size timber, high crown closure percentage, and variable shrub and shade-tolerant tree understory.

Direct comparison between the SPOT and Landsat-4 images was made during the fall of 1984. This comparison was conducted not to determine which Earth observation provides the higher quality data, but rather to determine those attributes of each system which will improve the renewable resource inventory and monitoring process.

6.2 Results and Discussion

The results presented here include those previously published by DeGloria (1984) with additional comments related to the Landsat-4 image products shown in Figure 6.1.

The high contrast ratio between the dark tree crowns (and shadows) and the bare soil at site "A" result in the individual trees being detected by a sensor with 10 meter ground pixel size (panchromatic). Because of the larger pixel sizes (28.5m and 57.0m), the Landsat imagery does not permit the detection of these individual tree crowns. At site "B", young growth conifers with a high proportion of hardwood tree and shrub species results in brighter red tones and smoother texture where cover is the highest. As the stand matures and shrub cover diminishes, more open stands with discernible tree crowns are evident at site "C" resulting in lighter tone and rougher image texture. Mature forest stands in this region are two- and three-tiered with an understory of shade tolerant tree species and a lower density of shrub cover. The image characteristics are usually a deep red tone with a smoother texture as the crown closure percentage approaches 80 percent or greater. Again, these variations in the age class structure of the stand as related to average crown diameter are not readily apparent in the Landsat-4 image products. Variations in stand densities over large areas, however, are detectable on the Landsat products.

Through a visual comparison of the simulated SPOT and Landsat-4 images with the Zeiss aerial photography and to the field plot data collected in 1982, the following significant results were obtained:

(1) Individual forest stands having different stand densities and size classes can be discriminated using the SPOT film product. The relatively high contrast ratio between large tree crowns and bright soil or shrub background enables the discrimination of mature trees in low crown closure environments.

(2) Plotting on maps and navigating on primary, secondary, and tertiary road networks is possible using the enhanced SPOT film product. This is more so than on either the TM or MSS composites. This capability would make the SPOT image a valuable source of information for updating transportation networks on existing maps or for generating new maps in areas scheduled for timber harvest operations.

(3) Major soil associations and sites suitable for reforestation practices can be identified on the SPOT imagery when the association between distinct vegetation patterns and landforms can be made.

Even without the benefit of quantitative interpretation tests, the quality of both the simulated SPOT and Landsat-4 imagery is excellent in terms of the consistent spectral representation of forest vegetation features when compared to standard aerochrome infrared photo products. In addition, the improved spatial resolution of the SPOT imagery allows the discrimination of features which are currently extracted from high altitude aerial photography and not from Landsat imagery, such as transportation networks and the stratification of large homogeneous vegetation units. Both the stereoscopic and off-nadir imaging capability will enhance the assessments of forest resources. In that changes to the forest resource base are less dynamic than in an agricultural environment, use of the stereoscopic viewing and mapping capability of the SPOT system will be used to a greater extent than the off-nadir viewing capability. Stereoscopic viewing and interpretation of enhanced film products will enhance forest vegetation and soil mapping activities, because the relief variables (elevation, slope, aspect) contribute significantly to the distribution of unique plant and soil properties.

6.3 Conclusions

The SPOT and Landsat TM data in film format can be used to satisfy a number of renewable resource assessments in forested environments. These include using the film product as a mapping base, a medium for stratification in support of timber inventory, and a source for updating existing data bases of forest information systems by digitized entry of manually-interpreted and delineated polygons or attribute data. Because of the high quality of the film products, these authors envision increased use of manual interpretation techniques of SPOT and Landsat data products for many forest research and and operational activities, and that enhanced film products from the SPOT-provided CCT's will be a strong component of the value-added industry.

119

For the global mapping and monitoring of forest resources and forest land management practices, the SPOT data will be a valuable addition to the suite of Earth observation systems currently in operation or in the planning phases. Technically, these systems have subtle, but unique, differences so as to improve their marketability to the global user community. We will see many comparisons of these systems to set one apart from another with the intent to gain some advantage in the marketplace. Though this competition will be beneficial in terms of improved data products and pricing of these products, we must not lose sight of our end objective to use the best attributes of each system to improve our understanding of the Earth's global resources for the benefit of all mankind.

6.4 Acknowledgements

6.5 References

CNES. 1984. SPOT -- Satellite-based remote sensing system. Centre National d'Etudes Spatiales, Toulouse. Brochure, March.

DeGloria, S.D. 1984. The interpretation of forest resources in California on simulated SPOT imagery. SPOT Simulation Handbook. Am. Soc. Photogrammetry, 9p.

Wensel, L.C. 1977. Wildland Resource Sampling. Dept. For. Res. Mgt. Univ. of California, Berkeley.
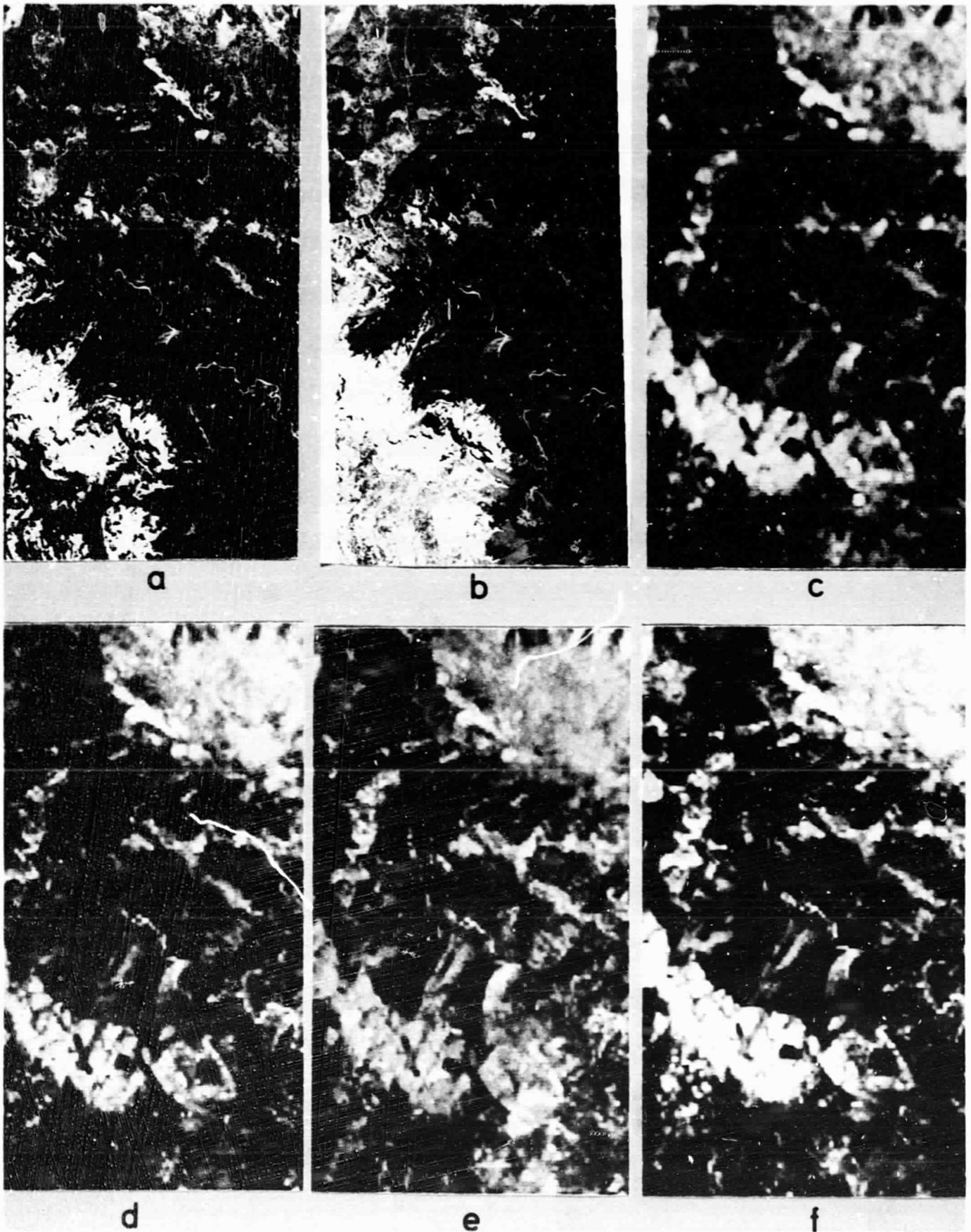
Figure 6.1   Portion of study area imaged by the Zeiss metric camera (a);
the combined high resolution panchromatic and lower resolution
simulated SPOT multispectral scanner image (b);   Landsat 4 MSS
4, 2, and 1 (c);   Landsat-4 TM 4, 3, and 2 (d); TM 5, 4, and 3
(e); and TM 5, 3, and 2 (f).

Table 6.1 Major attributes of the SPOT Satellite System (CNES, 1984).

Systeme Probatoire d'Observation de la Terre (SPOT)

- Pointable, high resolution visible (HRV) imaging sensors (across-track $\pm 27^{\circ}$)
- Stereoscopic Imaging
- Multispectral (visible, near-infrared), 20m Ground IFOV pixels
- Panchromatic, 10m Ground IFOV pixels
- 60km swath width, each sensor
- 117km swath width, both sensors (3km overlap)
- 98.7 orbital inclination
- 832km orbital altitude
- 26-day repeat cycle
- 10:30 acquisition, local solar time
- Direct X-band telemetry (50M bits/s), or tape recorders

Table 6.2 Major characteristics of the two high resolution visible (HRV)
imaging sensors on the SPOT satellite (CNES, 1984).

- 2-High Resolution Visible (HRV) range instruments:

    - Multispectral: 0.50 - 0.59um
                      0.61 - 0.68um
                      0.79 - 0.89um

    - Panchromatic: 0.51 - 0.73um

    - 8-bit data quantization format (256 grey levels)

- Multispectral Linear Array (MLA) technology:

    - Linear arrays in focal plane are electronically sampled for across-track coverage

    - Satellite orbit movement provides along-track coverage

- Detectors:

    - 3000/spectral band, 3ms sampling rate, 20m pixels, 60km swath (3000 pixels/line)
    - 6000/spectral band, 1.5ms sampling rate, 10m pixels, 60km swath (6000 pixels/line)
    - Charge Coupled Devices (CCD), Fairchild
    - Dichroic prism for spectral separation

Table 6.3 Description of the "P" Site data for the 1983 U.S. SPOT Simulation
Campaign, Site #74.

Simulated SPOT Data Set Description

- Location: Plumas County at Bucks Lake
- Acquisition period: 6/25/83 (18:54-18:57 GMT)
- Ground area coverage: 5 x 12km
- Panchromatic data sampled to 10m IFOV pixels (CCT)
- False color composite print (10m + 20m) at 1:24,000 scale
- Panchromatic Black & White print at 1:24,000 scale
- Black & White transparencies of each spectral band
- 9 x 9 in. color infrared photographs acquired simultaneously
- Flight log, topographic maps, and auxiliary information

Chapter 7

FIELD BOUNDARY DEFINITION USING
THEMATIC MAPPER BLACK-AND-WHITE FILM PRODUCTS

Jay Baggett

Statewide Planning Branch
Department of Water Resources
The Resources Agency
State of California

7.1 Introduction

The Department of Water Resources (DWR) is actively participating in a
cooperative project which is developing an operational Landsat-based
agricultural resource inventory system in California.   In addition to the
DWR, the participants in this "California Cooperative" are the U.S.
Department of Agriculture, University of California - Berkeley, National
Aeronautics and Space Administration, and California Department of Food and
Agriculture.   The responsibility of the DWR is to collect, analyze, and
update field survey data including field boundaries of cover crops of
interest for subsequent integration into the Landsat spectral data
classification procedures.   The DWR's work to date in the California
Cooperative Project indicates that the USDA's California June Enumerative
Survey sample segments do not adequately capture the crop type diversity of
California's agriculture.   An assessment of this variability is required if
an adequate statistical training set is to be developed for subsequent use
in supervised or unsupervised image classification algorithms.

One method recently developed by the DWR staff to account for this
variability is to collect detailed ground observations along an irregular
transect which is so located as to traverse most crops and agronomic
practices found in a given study area.   Analysis of transect data collected
for phenological and Landsat date selection purposes has shown that the
transect method may prove to be superior to the established methods of using
randomly located fields for calculating the mean and covariance matrices for
the crop types of interest.   Current activities have focused on designing a
test of this transect method for a portion of the Sacramento Valley.

A major problem encountered in such a transect method is that the
location and boundaries of individual fields must be accurately portrayed at
the time of the inventory.   The field surveyor will have copies of USGS
topographic quadrangle maps;   DWR land use maps;   county road maps;   and,
occasionally, 35mm prints of portions of recently acquired (0-3 years) high
altitude photography.   Field boundaries, however, change over time and
accurate verification of them in the field is time consuming and expensive.
One method that could be used to rectify this problem would be to have
concurrent aerial photography acquired, but this would be costly, and
difficult to acquire and archive in a consistent fashion.

An alternative approach to the field boundary problem, and one which we at DWR feel is operationally feasible, is to use Thematic Mapper imagery to capitalize on its relatively high spatial resolution, and repetitive and synoptic coverage. Black-and-white photo products would be used to further reduce costs. This chapter describes my preliminary evaluation of using TM Bands 2, 3, 4, and 7 in positive film transparency format as a substitute for using concurrently acquired high altitude aerial photography in support of a Sacramento Valley-wide ground data collection effort.

7.2 Methods and Materials

Seven agricultural areas were selected to capture the range in agronomic practices as noted during a visual interpretation of the Landsat-4 scene acquired on 12 August 1983 (Path 44, Row 33). The selected sites and dominant cover conditions prevalent at the time of the overpass are listed in Table 7.1. Evaluation of the film products entailed of reproducing each TM spectral band in black-and-white format for each site using 35mm slide film. The ground area covered by each slide corresponded to a 15' block (in units of latitude and longitude). Copies of the most recent DWR land use maps were photoreduced for each test site from a scale of 1:24,000 to a scale of 1:63,360 and 1:50,000. Through the use of a projection table with a clear glass top and a mirror adjusted at a 45 degree angle, the TM copy slides were rear-projected onto the respective land use maps. Scale and orientation adjustments were made by positioning the projector until the TM slide and map products were registered. Once in registration, qualitative evaluations were made to determine which TM bands best defined field boundaries.

7.3 Results and Discussion

Figure 7.1 illustrates the quality of the individual TM bands evaluated (TM 2, 3, 4, and 7) for field boundary discrimination in the Dunnigan Hills/El Dorado Bend area of the Sacramento Valley, California. The results of the evaluation of TM film products for the seven test sites listed in Table 7.1 are as follows:

TM Band 2 (Green, 0.52-0.62μm)
Overall, this band was the most pleasing to the eye, having a "pan-chromatic" quality. It was generally medium-to-low contrast with a grayish cast over all scenes. The low contrast gray was particularly noticeable in the orchard and vineyard areas where these individual fields seemed to blend together to form one "super-field". (This phenomenon raises some interesting options for image-driven stratification, but this was not part of our current task.) Very careful viewing of these areas did result in distinguishing some fields, although many were pockets of non-orchard/vineyard crops. In the field crop areas, Band 2 was excellent; even check dams or dikes in the rice paddies were faintly visible. In the dryland areas, the image was adequate allowing some discrimination of native grasses, grain stubble, and bare soil.

TM Band 3 (Red, 0.63-0.69μm)
This was a high contrast image that showed some evidence of striping on the slides. Compared to Band 7, another high contrast band, it was slightly muted. The high contrast virtually "whited-out" the dryland grain and native grassland environments making this band an interesting selection for

some type of stratification, but useless for boundary definition in these areas. In the field crop areas, the high contrast increased definition between fields with dissimilar brightness, but adjacent fields that were similar (as seen on Band 2) appeared to merge into single fields. This band was very good in the orchards and vineyards; many individual fields were visible.

TM Band 4 (Near-reflectance Infrared, 0.76-0.90µm)

Brightness on cropped fields was almost a "negative" of those seen on the other three bands. The image was low contrast and many of the slides were quite dark. The most significant contribution of this band was the sharp definition of freeways and urban and suburban street patterns; similarly distinct were waterways. Interior boundaries (e.g. checkdams in rice fields, service roads in large fields with the same crop, and homesteads) were easily visible. Boundaries between adjacent fields with different brightness were not significantly improved over Band 2, and the dark nature of my copy slides caused the field crop area to appear less distinguishable than when using Band 2. This image was almost useless for the orchard and vineyard areas which appeared quite muted. In the dryland grain areas, this band was superior to all others, especially for distinguishing bare fields. In this type of environment, where fallow soils are almost a sure precursor to planting nonirrigated grains, this band would make a valuable contribution.

TM Band 7 (Far-reflectance infrared, 2.08-2.35µm)

This band had the highest contrast between light and dark fields when compared with the other bands. The bright areas such as grain stubble, bare soils, and native grasslands, were totally washed out. Adjacent dark fields tended to merge together. These factors made the band totally useless in the dryland areas, and only marginally useful in the field crop areas. In the orchard and vineyard areas, however, this band was the best, with many individual "fields" being visible. Because canopy understory management practices vary considerably and because DWR maps do not show boundaries between adjacent fields of the same crop type, it was difficult to determine if the "fields" were real or apparent. Such fields were more easily discriminated on Band 7 than on Band 3, and striping was not present.

Based on this evaluation, I recommend the use of (1) TM Band 2 for discriminating field boundaries because of its adequacy in the dryland, its superiority in the field crop areas, and its overall utility, and (2) TM Band 7 for its superior performance in the orchard/vineyard areas. At some future date, however, re-evaluation of Band 4 and an evaluation of other bands and color composites should be considered.

In an operational environment, the ground survey should take place beginning in mid-June to mid-July. The field person would have maps, relatively recent U-2 imagery, tabular listings of expected general locations of particular crops, and perhaps a "flag-map" showing locations of "spring-active" fields derived from early season Landsat data. It is important that we capture double-cropping and full season variability, as our field visit is timed for maximum summer activity and evidence of spring activity is not always immediately obvious in the field; and such flagging may prove beneficial and eliminate the need to make two field visits.

Appropriate bands from a TM scene concurrent with the ground data collection period would be ordered, and the analyst would make the appropriate 35mm slides. These would be projected onto the maps to correct locations and boundaries of sampled fields after the ground survey had been completed.

Other options exist if very fast turn-around (two weeks after acquisition) can be assumed. Inexpensive black-and-white 35mm prints can be taken into the field so that boundary and location questions can be settled on a real-time basis.

The primary reason for using black-and-white film products in this investigation was cost. The estimated cost for Landsat data acquisition for the agricultural portion of the Sacramento Valley is $800 (8 scenes x 2 bands/scene x $50/band). This is a very minor cost. Color Landsat coverage, however, would range from $1120 ($140/scene) to $3560, depending on whether or not we paid the color composite generation charge ($305/scene). This color composite cost still represents a minor part of the total inventory budget, and considering the added information available to the analyst and the number of times that composites are used to clarify results from digital analysis, the added cost may be justified in an operational inventory.

Because the DWR would like to use this technique for next summer's (1985) work, we continue to be supportive and most interested in an operational Thematic Mapper system.
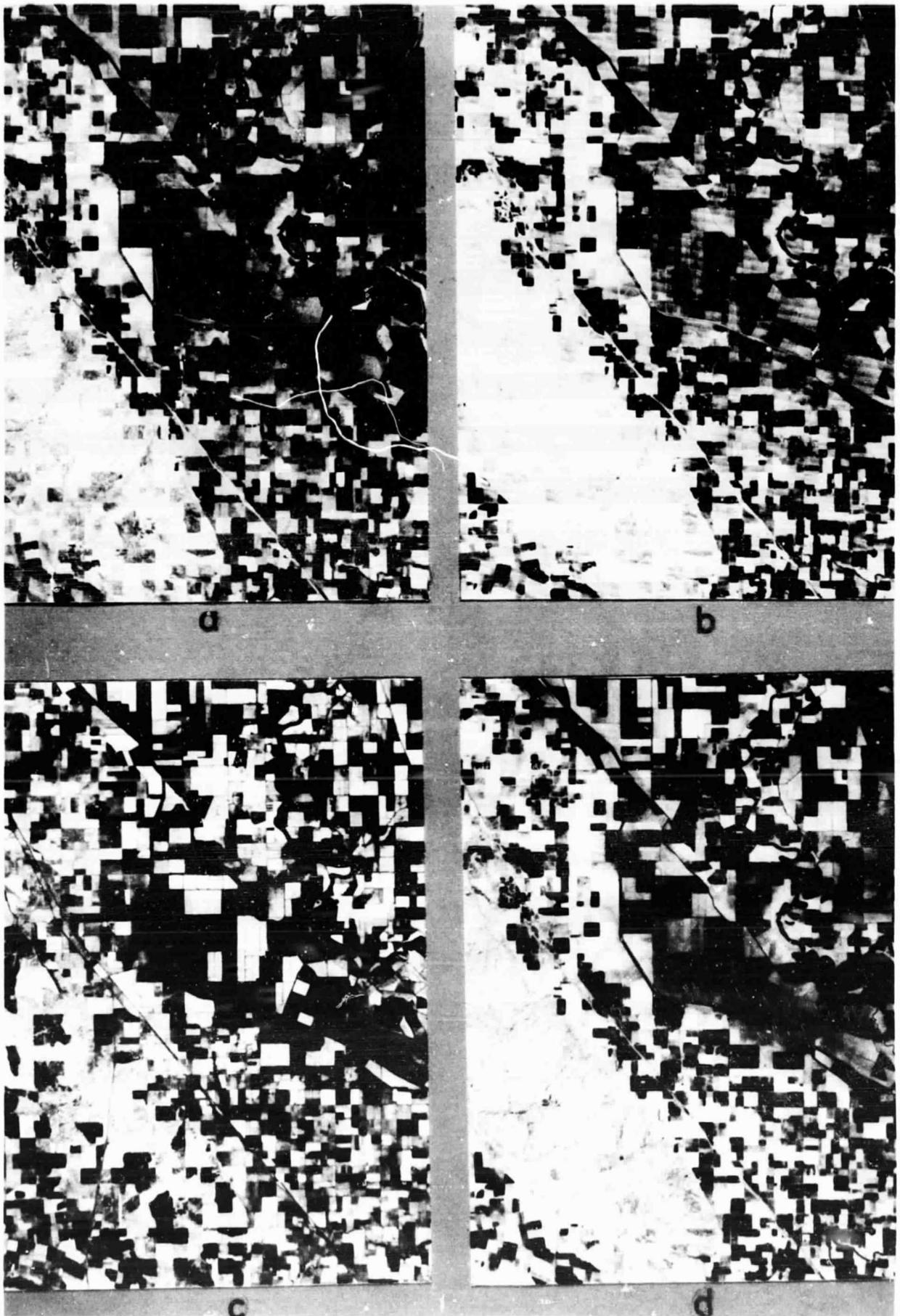
Figure 7.1 The four TM spectral bands used for field boundary discrimination in the Central Valley of California: a – TM 2, b – TM 3, c – TM 4, and d – TM 7.

Table 7.1. Sites used to evaluate the use of black-and-white Thematic Mapper film products for updating field boundary locations in support of Landsat-aided crop surveys in California.

SITE

COVER CONDITIONS

1. Marysville/Yuba City

urban, deciduous orchards, rice, field crops

2. Capay Valley

small fields (primarily deciduous orchards and nonirrigated grains), native grasslands-nonirrigated grains

3. Dunnigan Hills/El Dorado Bend

dryland grain, large rice fields, field crops

4. Winters/Woodland

field crops, urban, orchard, dryland grain; concurrent transect data

5. Sacramento/Galt

pasture, wetlands, field crops, native grasses, expanding urbanization

6. East of Stockton (Waterloo/Lockford)

vineyards, deciduous orchards; wide range of soil brightness

7. Napa Valley

vineyards, native grasses, urban; Coastal Valley site

Chapter 8

USE OF LANDSAT THEMATIC MAPPER DATA IN
COASTAL WATERSHED LAND USE PLANNING

Thomas G. Dickert
Robert B. Olshansky

Center for Environmental Design Research

8.1 Introduction

This work has grown out of research conducted in conjunction with the California Coastal Commission with support from the NOAA Sea Grant Program and the California Resources Agency (reported in Dickert and Tuttle, 1981; Dickert and Tuttle, 1985; Dickert and Olshansky, 1985). The purpose of the research has been to understand the effects of upland land disturbance on sedimentation in a coastal wetland, and to develop a land use planning approach that recognizes the link between watershed land uses and wetland sedimentation. Our research focused on the watershed of Elkhorn Slough, a large coastal wetland located at central Monterey Bay, California (see Figure 8.1). Field measurements confirmed that vegetation removal and row crop cultivation have contributed significant amounts of sediment to the stream system and have caused actively-growing alluvial fans to develop at the margins of the wetland.

The field evidence was used as the rationale for development of a land use planning system for the 180 square-kilometer watershed, as part of the North Monterey County Local Coastal Program (LCP). The planning system of two components: land disturbance targets developed for 37 sub-units of the watershed, and an erosion susceptibility map portraying the relative erodibility of lands within the watershed. The land disturbance targets are based on the concept of hydrologic land use intensity, which assumes that both bare ground and impervious surfaces in the watershed have altered the natural hydrologic processes and may contribute to wetland degradation. Thus, the land use plan sets targets for the total amount of land disturbance (including both the percentage area of bare ground and impervious surface) to be permitted in each sub-watershed.

The research included the development of a geographic data base, as an aid to analysis and graphic output. The base consists of 71,976 quarter-hectare (50m x 50m) grid cells. Each cell contains the following variables: bedrock geology, soil type, soil erosion k-factor, slope category, overland flow distance, and land use category. The data were compiled from existing geologic reports, soil surveys, topographic maps, and aerial photography.

The mapping of land uses, measurements of land disturbance from aerial photographs, and other data analysis were all completed prior to the acquisition of Landsat MSS and TM data. Part of our research consisted of land disturbance measurements from a random sample of parcels, stratified by sub-watershed and land use category. For each parcel we measured percent impervious surface coverage and percent exposed soil. In this rural

watershed, bare ground is much more common than impervious surfaces. Measurements were made on 1:1200 scale color infrared aerial photos. The photographs were taken April 10, 1980, and thus were representative of bare ground conditions following the rainy season. Measurements were field-checked and adjusted where necessary; for row crops, the reported bare ground values represent estimates of the ratios of crop cover to inter-row distance, and only represent actively cropped fields. Therefore, this method probably underestimates exposed soil in cultivated parcels, since it does not account for temporarily fallow fields (Bean, 1981). Table 8.1 shows the results. This work revealed that each nominal land use category is extremely heterogeneous in disturbance characteristics; it also showed that dwelling unit density does not necessarily correlate with the amount of land disturbance present.

More recently, we have been analyzing Landsat MSS and TM imagery to test its utility as an alternative to aerial photography, for detecting land disturbance in coastal watersheds. Should this effort succeed, Landsat imagery could be used to periodically monitor land disturbance not only in Elkhorn Slough watershed, but in coastal watersheds throughout California and other coastal states. A low cost means of monitoring is needed in order to assure ongoing compliance of proposed projects with the approved local coastal plans.

8.2 Methods

Our Landsat research thus far has focused on comparing the spectral values for the MSS and TM data with the land use data interpreted from the low altitude color infrared imagery. The first goal was to correlate spectral data to our land use categories. We used Landsat MSS data from May 1980 and 9 November 1981, and TM data from 31 December. The scenes were resampled to 50-meter pixels and were registered to the UTM projection of our data base. Thus, the analysis of the MSS data was at a greater resolution than warranted by the 80-meter data, whereas the TM analysis did not use the full capabilities of the 30-meter resolution of the original data. The residual root mean square registration errors were approximately 40 meters for the May 1980 MSS scene, 52 meters for the November 1981 MSS scene, and 21 meters for the December 1982 TM scene. The May MSS and December TM scenes were of high quality, but the November MSS scene was approximately 10 to 20 percent cloud obscured.

Two-band ratios were calculated for each scene and averaged for each land use category. For the MSS scenes we calculated twice the ratio of bands MSS 7/MSS 5, and used the corresponding ratio of TM 4/TM 3 for the TM scene. These ratios were chosen because they are generally felt to be the best indicators of vegetation intensity; they could therefore be used as inverse measures of land disturbance. The means of these ratios for each general land use category are shown in Table 8.2. Unfortunately, direct comparison of the three scenes was difficult for several reasons: seasonal variation in vegetation and sun angle, cloud cover in the November scene, and nonequivalence of MSS 7/MSS 5 and TM 4/TM 3 ratios.

8.3 Results and Discussion

The results from analysis of the TM scene was shown to be the most useful for our work. The late December date of the TM scene is a time of minimal vegetation and heavy rains, and is therefore the season for which we most desired land disturbance measurements. In addition, because of the superior resolution of the TM scene, it can be more accurately to the detailed land use map. In Figure 8.1 the two-band ratios for the TM data are plotted against the mean values of land disturbance. There is no clear correlation between the two.

Table 8.3 shows the complete breakdown of two-band ratios by land use category. The standard deviations are quite large, as they were for the photo-interpreted land disturbance data. In all three scenes low ratios are seen for the least vegetated land use categories: commercial, industrial, highway, railroad, bare ground, sand dunes, high density residential, mudflats, and open water. However, there are no clear patterns among the major vegetation categories in the watershed: brush, pasture, oak woodland, row crops, strawberries, and orchards (summarized in Table 8.2). Nor is there any systematic relationship either with residential density or with the disturbance values previously measured for each density category (see Figure 8.2). Some seasonal changes can be detected between the May and December scenes. Pasture is relatively more vegetated in December, whereas brush, oak woodland, and strawberry fields are relatively less vegetated.

The greatest difficulty in this research was the inability of the Landsat and the TM data to clearly detect the great amounts of exposed ground in cultivated areas. One possible explanation is that the pixel sizes of 80 meters and 30 meters are too large to detect bare soil between rows of crops; they instead detect a mean spectral value that is heavily influenced by spectral reflectance of the leafy portion of the crops. This is further confused by the mix of fallow and cropped fields in each land use category in each scene. In general, this analysis has been obscured by the noise of the tremendous variations in disturbance characteristics within each nominal land use category.

The best way to resolve this question would be to resample specific parcels or polygons containing a single land use from the aerial photos and directly compare them to the same parcels on the TM scene, masked by the digitized parcel boundaries. Such an analysis could be done at both the full 30-meter resolution of the TM data, as well as at a resampled 50-meter resolution. In this way, we could directly measure the ability of the TM data to detect land disturbance, without using nominal land use categories as an intermediate step.

8.4 Conclusions

Landsat MSS and TM data were analyzed to test their utility in detecting land disturbance in a coastal wetland watershed. They were found to be able to detect large expanses of bare ground and urbanization. Use of the TM data to detect more subtle, smaller scale disturbance is less conclusive. Further research is needed for a more definitive answer to this question. Spatial resolution appears to be a key variable in the feasibility of using Landsat as a watershed monitoring tool. The TM data is therefore a step in the right direction.

8.5 References

Bean, M. 1981. A site disturbance index for a California coastal watershed: Elkhorn Slough watershed case study. Unpublished. Department of Landscape Architecture, University of California, Berkeley.

Dickert, T.G., and Olshansky, R.O. 1985. Evaluating erosion susceptibility for land use planning in coastal watersheds. Coastal Zone Management Journal. Submitted.

Dickert, T.G. and Tuttle, A.E. 1985. Cumulative impact assessment in environmental planning: a coastal wetland watershed example. Environmental Impact Assessment Review. 4(3): in press.

Dickert, T.G. and Tuttle, A.E. 1981. Elkhorn Slough watershed: linking the cumulative impacts of watershed development to coastal wetlands. University of California, Institute of Urban and Regional Development, Berkeley.
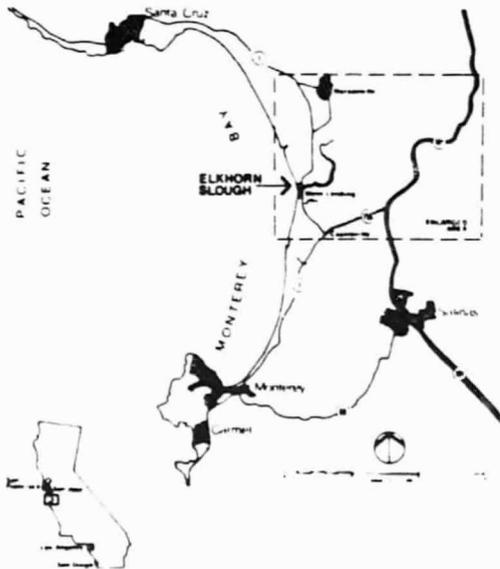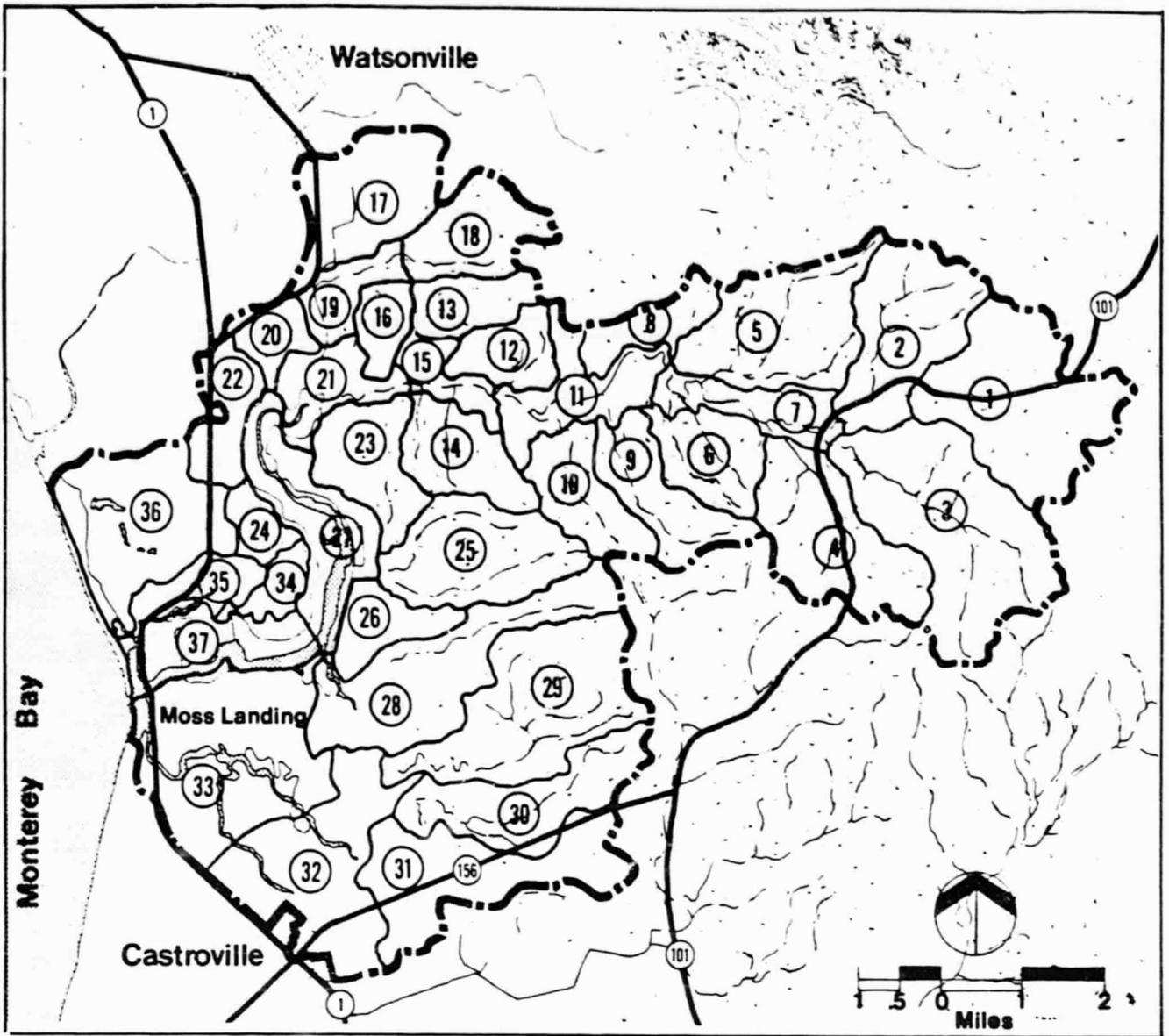
Figure 8.1 Location of the Elkhorn Slough study site along California's central coast and its associated watershed and subwatershed boundaries.

133

```
         5.00    |
                 |
                 |
                 |                        ✿
                 |
Mean             |         ✿    ✿    ✿         ✿   ✿              ✿✿
2(TM4)/TM3       |        ✿✿   ✿  ✿2        ✿  ✿              ✿
                 |         ✿  ✿2   3  ✿2✿   ✿✿✿    ✿       ✿
                 |                        ✿
                 |
                 |
                 |
                 |                                                        ✿
                 |
                 |
                 |       ✿        ✿
         1.50    |--------------------------------------------------------
                 0.   Mean land disturbance (percent)                60.00
```
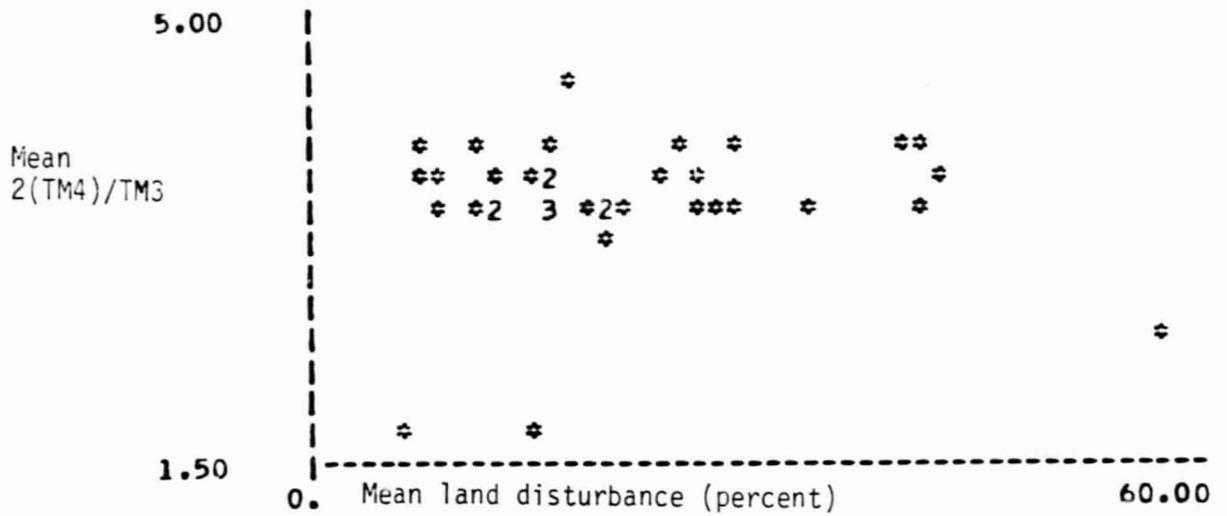
variable	coeff	corr.	t-stat
1	-0.926123e-03	-0.198261e-01	-0.117316

intercept	multiple r	f-stat
3.56000	0.198261e-01	0.137630e-01

Figure 8.2 Mean land disturbance versus the TM 4/TM 3 band ratio for land use categories in the Elkhorn Slough watershed.

134

Table 8.1 Land disturbance values for various land uses and residential
parcels in the Elkhorn Slough watershed.

Land Disturbance Values

Land Use	Impervious Surface mean (std. dev.)	Bare Ground mean (std. dev.)
Brush (coastal scrub)	0.0% (0.0%)	6.0% (8.7%)
Oak woodland	0.3% (0.9%)	1.1% (1.1%)
Eucalyptus	0.0% (0.0%)	3.2% (4.8%)
Pasture, grassland	1.3% (2.1%)	6.5% (6.5%)
Rowcrops (artichokes)	0.2% (0.3%)	63.1% (19.2%)
Strawberries	0.2% (0.3%)	69.5% (13.6%)
Orchards	0.3% (0.6%)	93.2% (2.5%)
Commercial	33.1% (15.9%)	51.4% (28.5%)
Industrial	12.3% (11.1%)	48.2% (42.0%)
Residential Parcels:		
51+ acres	0.6% (0.7%)	9.6% (5.3%)
31-50 acres	0.6% (0.3%)	7.9% (6.0%)
15.5-30 acres	2.9% (3.5%)	13.2% (6.3%)
10-15.4 acres	0.5% (0.2%)	8.1% (4.8%)
5-9.9 acres	1.9% (1.6%)	9.9% (10.6%)
2.4-4.9 acres	9.2% (6.2%)	3.8% (4.1%)
1-2.3 acres	9.6% (9.3%)	14.7% (9.1%)
0.5-0.9 acres	15.7% (5.1%)	19.5% (19.0%)
0.2-0.4 acres	24.8% (7.6%)	0.8% (1.1%)
0.1 acres	30.7% (9.0%)	11.4% (7.1%)
less than 0.1 acres	59.1% (3.8%)	51.4% (12.8%)

Table 8.2   Mean values of two-band ratios by land use category and residential uses.

Land Use	N of pixels	May 1980 2(MSS7)/MSS5	Nov. 1981 2(MSS7)/MSS5	Dec. 1982 2(TM4)/TM3
Brush	5801	3.032	2.502	3.294
Oak woodland	10323	3.124	2.541	3.363
Eucalyptus	2293	3.886	2.821	4.095
Pasture	33028	2.459	2.369	3.880
Row crops	7848	3.212	2.582	3.861
Strawberries	3859	3.134	2.778	3.382
Orchards	301	3.044	2.831	3.756
Commercial	870	1.946	1.823	2.571
Industrial	525	1.525	1.450	2.078
Urbanized	811	2.156	1.915	3.099
Salicornia	1700	1.803	1.644	1.664
Wet grassland	1366	3.094	2.306	2.808
Residential Uses:				
51+ acres	6722	3.034	2.668	3.932
31-50 acres	1357	2.702	2.549	3.821
15.5-30 acres	1899	2.663	2.164	3.539
10-15.4 acres	3104	2.578	1.992	3.069
5-9.9 acres	2722	2.490	2.398	3.806
2.4-4.9 acres	1630	2.548	2.336	3.748
1.0-2.3 acres	3339	2.604	2.402	3.736
0.5-0.9 acres	1275	2.758	2.407	3.637
0.2-0.4 acres	615	2.478	2.211	3.617
0.1 acre and less	453	1.934	1.748	2.691

Table 8.3   Two-band ratios for different land use categories in the Elkhorn
             Slough watershed.

Land Use Type (and parcel size if residential)	N of pixels	Two-band ratios (and std. dev.)		
		May 1980 2(MSS7)/MSS5	Nov.1981 2(MSS7)/MSS5	Dec.1982 2(TM4)/TM3
Brush (non-res.)	4999	3.049 (.806)	2.494 (.643)	3.259 (.697)
(51+ acres)	67	3.826 (.881)	3.318 (.722)	3.783 (.656)
(15.5-30 acres)	80	2.891 (.471)	2.343 (.567)	3.473 (.833)
(5-9.9 acres)	80	2.987 (.580)	2.506 (.383)	3.425 (.644)
(2.4-4.9 acres)	152	2.960 (.613)	2.578 (.530)	3.573 (.823)
(1.0-2.3 acres)	393	2.765 (.495)	2.486 (.349)	3.446 (.720)
(0.5-0.9 acres)	22	2.810 (.358)	2.095 (.331)	3.383 (.541)
Pasture (non-res.)	18367	2.440 (.759)	2.386 (.575)	3.878 (1.072)
(51+ acres)	3921	2.518 (.879)	2.290 (.716)	3.996 (1.137)
(31-50 acres)	1287	2.659 (.814)	2.519 (.700)	3.855 (1.149)
(15.5-30 acres)	1063	2.384 (.518)	2.350 (.487)	4.082 (1.275)
(10-15.4 acres)	1373	2.454 (.613)	2.336 (.517)	3.754 (.957)
(5.0-9.9 acres)	2184	2.373 (.566)	2.355 (.541)	3.881 (1.000)
(2.4-4.9 acres)	1273	2.448 (.532)	2.272 (.391)	3.778 (.989)
(1.0-2.3 acres)	2502	2.509 (.507)	2.380 (.445)	3.827 (.922)
(0.5-0.9 acres)	860	2.601 (.579)	2.396 (.477)	3.683 (.891)
(0.2-0.4 acres)	151	2.409 (.442)	2.320 (.470)	3.844 (.985)
(0.2 or less)	47	2.567 (.950)	2.029 (.410)	3.132 (1.050)
Row crops (non-res.)	3737	3.138 (1.196)	2.759 (.905)	4.513 (1.550)
(51+ acres)	1677	4.031 (1.500)	3.384 (1.048)	4.229 (1.615)
(15.5-30 acres)	828	3.023 (1.379)	1.921 (.508)	2.841 (1.284)
(10-15.4 acres)	1485	2.620 (1.844)	1.624 (.545)	2.411 (.938)
(5.0-9.9 acres)	76	2.473 (.496)	2.110 (.358)	3.361 (.681)
(1.0-2.3 acres)	29	3.259 (.677)	2.630 (.612)	3.971 (1.057)
Strawberries (non-res.)	2843	3.072 (.911)	2.744 (.729)	3.451 (.900)
(51+ acres)	800	3.267 (.858)	2.947 (.833)	3.155 (.886)
(31-50 acres)	70	3.495 (.759)	3.103 (.547)	3.097 (.601)
(10-15.4 acres)	112	3.353 (.806)	2.314 (.623)	3.300 (.482)
Oak woodland (non-res.)	8797	3.138 (.766)	2.538 (.585)	3.350 (.713)
(51+ acres)	257	3.456 (.638)	2.738 (.546)	3.395 (.548)
(10-15.4 acres)	44	2.590 (.742)	2.026 (.619)	2.904 (.895)
(5.0-9.9 acres)	306	3.040 (.568)	2.594 (.403)	3.435 (.588)
(2.4-4.9 acres)	147	2.865 (.534)	2.565 (.431)	3.576 (.657)
(1.0-2.3 acres)	407	2.967 (.621)	2.434 (.569)	3.422 (.717)
(0.5-0.9 acres)	306	2.984 (.499)	2.602 (.458)	3.491 (.626)
(0.2-0.4 acres)	59	2.845 (.452)	2.585 (.361)	3.435 (.638)
Bare ground	73	1.991 (.468)	1.881 (.356)	2.647 (.672)
Commercial	870	1.946 (.599)	1.823 (.429)	2.571 (.801)
Alluvial fans	40	2.591 (1.185)	2.052 (.655)	2.925 (1.217)
Highway	35	1.659 (.300)	1.775 (.360)	2.997 (.676)
Industrial	525	1.525 (.415)	1.450 (.297)	2.078 (.564)
Mudflats	600	.799 (1.139)	1.120 (.625)	1.344 (.380)
Eucalyptus (non-res.)	2075	3.857 (1.044)	2.809 (.839)	4.077 (.926)
(0.5-0.9 acres)	87	3.502 (1.086)	1.903 (.213)	3.791 (.989)
Young Eucalyptus	123	4.671 (.986)	3.662 (.633)	4.577 (.901)
Orchards (non-res.)	207	3.199 (1.160)	2.919 (.761)	3.700 (1.034)
(5.0-9.9 acres)	59	2.799 (.478)	2.563 (.330)	4.045 (.798)
(2.4-4.9 acres)	25	2.657 (.400)	2.704 (.789)	4.038 (1.283)
Open water	1352	1.774 (1.532)	1.589 (.782)	1.879 (.858)
Railroad	158	1.505 (.510)	1.451 (.229)	1.934 (.313)
Salt ponds	459	.565 (.514)	1.264 (.283)	1.687 (.252)
Beach	13	.646 (.552)	.918 (.259)	1.496 (.665)
Schools	115	2.409 (.925)	2.213 (.656)	3.514 (1.300)
Sand dunes	128	2.275 (.655)	1.909 (.384)	2.601 (.460)
Pickleweed (pure stand)	1189	1.672 (.940)	1.559 (.501)	1.578 (.587)
(many tidal channels)	185	.729 (.568)	.954 (.386)	1.277 (.195)
(pure or mixed)	102	2.204 (.891)	1.893 (.655)	2.199 (.814)
(mixed stands)	224	3.202 (.933)	2.551 (.584)	2.179 (.635)
Urban (0.2-0.4 acres)	405	2.451 (.322)	2.116 (.366)	3.555 (.769)
(0.2 or less acres)	406	1.861 (.347)	1.715 (.295)	2.637 (.700)
Wet grassland (wettest)	866	3.093 (1.065)	2.357 (.655)	2.556 (.775)
(intermediate condition)	68	2.686 (.404)	2.253 (.273)	4.674 (1.341)
(driest condition)	260	2.699 (.601)	2.339 (.507)	3.264 (.825)

Chapter 9

ATMOSPHERIC SCATTERING AND DIFFERENTIAL
ILLUMINATION STUDIES USING THEMATIC MAPPER SPECTRAL DATA

Dr. Walter H. Carnahan
Director
Image Processing Laboratory
Indiana State University
Terre Haute, Indiana 47809

9.1 Introduction

Research activities at the Image Processing Laboratory have addressed the estimation of the line spread function of the atmosphere and remote sensor systems, and the classification of vegetation in mountainous terrain using Landsat digital data. Only recently have Thematic Mapper data from the Landsat-4 spacecraft become available to address research topics of interest to the staff of the Laboratory. In June, this investigator sought to acquire a high quality TM scene which not only contained a large, vertically oriented smoke plume but also covered a forested area in mountainous terrain. I was fortunate to have found such an image at the Remote Sensing Research Program, University of California – Berkeley. Because our investigative interests were similar to theirs, the TM data from Path 44, Row 32 acquired on 12 August 1983, were temporarily loaned to our Laboratory for conducting two research projects briefly outlined below. Detailed research reports will be published upon the completion of our analyses.

9.2 Atmospheric Scattering Project:

In 1983 McGillem et al. used blurred edges from remotely sensed data to estimate the line spread function (LSF) of the atmosphere and remote sensor system. Two methods were directly applied to estimate the LSF in the spatial domain. One of them is to expand the LSF into a finite set of nonoverlapping rectangular basis functions; while the other is to utilize the scaled derivatives of smoothed raw data to estimate the LSF. Their study has shown the potential of using similar techniques to evaluate the atmospheric particulate blurring effects when radiation travels through the atmosphere.

The research conducted at the Image Processing Lab, Remote Sensing Lab, Indiana State University, has successfully accomplished its experiments to correlate the particulate concentrations with some parameters of the line spread function derived from processing Thematic Mapper data. It is found that broader equivalent width of a LSF calculated from an edge from the Landsat scene corresponds, in most cases, with higher particulate concentrations in the sampled area.

The TM data acquired from the northeastern portion of California show a forest fire, and it is hoped they will be useful for testing the techniques developed in this research. Our main attempt to use the data provided by the University of California - Berkeley, is to find some possibly useful edges along the smoke plume of the fire, calculate the LSFs for these samples, and then determine whether a correlation exists.

9.3 Differential Illumination Project:

Active research currently performed at ISU consists of utilizing Landsat digital data and Thematic Mapper data in soil and vegetation applications. In mountainous areas, however, these data contain differential illumination. This component causes misclassification of earth resources, and induces spectral variation, thereby reducing the data's potential for providing the maximum amount of information for earth surface features. Our research is concerned with modeling and isolating the differential illumination component in Thematic Mapper data, so that correction procedures can be developed and applied to TM data before soil and vegetation application procedures are implemented. The analysis procedure will consist of modeling TM data with digitized topographic data, and utilizing principal component analysis and Fourier transform techniques in producing Thematic Mapper images which solely contain information on the intrinsic properties of earth surface features.

9.4 References

McGillem, C.D., P.E. Anuta, E. Malaret, and K.B. Ywi. 1983. Estimation of a remote sensing system point spread function from measured imagery. Proceedings Machine Processing Remotely Sensed Data. Purdue Univ. pp 62-68.

Chapter 10

PRESENTATIONS AND PUBLICATIONS RELATED TO THE
UNIVERSITY OF CALIFORNIA - BERKELEY LIDQA INVESTIGATION

S. D. DeGloria
Remote Sensing Research Program

10.0 Introduction

This chapter summarizes the technical presentations, seminars, lectures, invited papers, reports, articles, and other forms of publication generated and distributed to the public during the reporting period of the project. This material is presented in a list format and in chronological order; it represents our best effort in satisfying the second major responsibility of the University as a LIDQA investigator, namely, "to promote the routine and practical use of (Landsat-4) data for the betterment of mankind" (NASA, 1984).

10.1 Presentations

"Characterization of Landsat-4 TM and MSS Image Quality for the Interpretation of California's Agricultural Resources"

    Landsat-4 Early Results Symposium
    NASA-Goddard Space Flight Center
    Greenbelt, Maryland
    22 February 1983

"Characterization of Renewable Resources Using Landsat-4 TM and MSS Imagery"

    American Society of Photogrammetry
    49th Annual Convention
    Washington, DC
    17 March 1983

"Agriculture, Forestry and Soil Survey (Lecture 14)"

    Landsat Sensor Design and Operations Short Course for Users
    Hughes Santa Barbara Research Center and University of California
      Santa Barbara Extension
    Santa Barbara, California
    16-19 August 1983

"Evaluation of Landsat-4 Image Quality for the Interpretation of Forest, Agricultural and Soil Resources"

    Pecora VIII Memorial Symposium
    Sioux Falls, South Dakota
    4-7 October 1983

"TM Data Quality Presentation"

   California Map Society Meeting
   San Francisco, California
   October 1983

"TM Data for Forestry Purposes"

   Remote Sensing Training Course
   Mendocino National Forest
   Willows, California
   8-9 November 1983

"Assessment of TM Image Quality for Forestry Purposes"

   Third Landsat-4 Investigator's Workshop
   NASA-Goddard Space Flight Center
   Greenbelt, Maryland
   6-7 December 1983

"TM Image Product Evaluation"

   California Community College Instructor's Workshop
   Modesto, California
   10 December 1983

"Quality of TM Data for Forestry and Agricultural Purposes"

   Landsat-D' Launch User Symposium
   Santa Barbara, California
   28 February - 1 March 1984

"Visible and Infrared Remote Sensing Techniques"

   Graduate Botany Seminar
   University of California
   Davis, California
   April 1984

"Advanced Image Interpretation Methods (Including an Evaluation of TM
Film and Digital Products)"

   Graduate Forestry and Resource Management Course
   University of California
   Berkeley, California
   Spring Semester, 1984

"Quality of TM Data for Hydrological Applications in a Semi-Arid
Agricultural Environment"

   Faculty Seminar
   Department of Land, Air, and Water Resources
   University of California
   Davis, California
   June 1984

"TM Image Product Discussion and Evaluation"

> Geography 106 Guest Lecture
> University of California
> Davis, California
> June 1984

"Evaluation of Thematic Mapper Data for Mapping Forest, Agricultural, and Soil Resources"

> XXV Congress
> Committee on Space Research (COSPAR)
> Graz, Austria
> 27 June 1984

"Spectral, Spatial, and Geometric Quality of TM Data for Renewable Resources"

> Engineering Faculty Seminar
> University of Naples
> Naples, Italy
> 2 July 1984

"Characterization of Crop and Forest Cover Types Using Landsat, TM, and Simulated SPOT Imagery"

> Faculty and Staff Seminar
> Centro Studi E Applicazioni in Technologie Avanzate (CSATA)
> Bari, Italy
> 3 July 1984

"Interpretability of California's Natural Resources on Simulated SPOT and Landsat-4 Imagery"

> Staff Seminar
> Centre National D'Estudes Spatiales (CNES)
> Toulouse, France
> 5 July 1984

Distinguished Visitor Seminars on Landsat-4 TM and MSS Image Quality presented at the Space Sciences Laboratory, University of California – Berkeley to the following individuals or delegations:

Mr. Camino, Director General, Promotion of Agriculture for Spain

Mr. Nedal, Director, Spanish Agricultural Research Service

Mr. Tasias, Director, Spanish Agricultural Experimental Station at Tarragona

Mr. Mathura D. Rajbhandari, Dean, Institute of Forestry, Hetauda, and Tribhuwan University, Tritureswor, Katmandu, Nepal

Mr. Sit Bo, Department of Forestry, Rangoon, Burma

Dr. Sergio Vetrella, Professor of Engineering, University of Naples, Naples, Italy

Dr. Walter Alvarez and graduate students, Department of Geography and Geophysics, University of California, Berkeley, California

Mr. Adrian Cubberly and Mr. John Tonelli, Science and Technology Corporate Research Center. International Paper, Tuxedo Park, New York.

Cheng Chi, Associate Professor of Geography-Peking University and Vice Director, Technical Training Department, National Remote Sensing Center, Beijing, People's Republic of China (PRC).

Liu X-Tian, President, Commission of Agricultural Regionalization of Shanxi Provence, Taiyuan, PRC.

DuRen, Vice President, Commission of Agriculture Reorganization of Shanxi Province, Taiyuan, PRC.

Gao Oijiang, Vice Director, Institute of Remote Sensing Applications of Shanxi Province, PRC.

Feng Chengji, Engineer, Institute of Remote Sensing Applications of Shanxi Province, PRC.

Li Shusun, Engineer and Graduate Student, University of California, Santa Barbara, California.

10.2 Publications

DeGloria, S.D. and R.N. Colwell. 1983. Characterization of Landsat-4 TM and MSS image quality for the interpretation of California's agricultural resources. Proc. Landsat-4 Early Results Symp., NASA-Goddard Space Flight Center, Greenbelt, MD.

DeGloria, S.D., A.S. Benson, and R.N. Colwell. 1983 Evaluation of Landsat-4 image quality for the interpretation of forest, agricultural, and soil recources. Proc. 8th W.T. Pecora Memorial Symp., Augustana Res. Inst., Augustana College, Sioux Falls, SD. p 147-161.

DeGloria, S.D. 1984. Spectral variability of Landsat-4 Thematic Mapper and Multispectral Scanner data for selected crop and forest cover types. IEEE Trans. Geoscience and Remote Sensing GE-22:303-311.

DeGloria, S.D., A. Benson, K. Dummer, and E. Fakhoury. 1984. Evaluation of Thematic Mapper data for mapping forest, agricultural, and soil resources. J. Adv. Space Research, Pergammon Press (in press).

DeGloria, S.D. and R.N. Colwell. 1984. Characterization of Landsat-4 TM and MSS image qualty for interpretation of agricultural and forest resources. NASA Conference Publ. #2326, Vol II abstr. p 98-102.

10.3 Miscellaneous Efforts

Preparation of material for General Electric Landsat-4 Brochure for Mr. David Wright, General Electric Corporation. December 1983.

Evaluation of Landsat-4 film products for Dr. Nicholas Short, NASA Goddard Space Flight Center, February 1984.

Highlighting of Landsat-4 TM and MSS image examples in:

"Introduction to Remote Sensing" -- a day-long presentation in Edmonton, Canada on 20 February 1984 by R. N. Colwell at the Twelfth Annual Remote Sensing Seminar, attended by 50 representatives concerned with the inventory, monitoring, and management of natural resources. (To be repeated on 19 February 1985.)

Keynote address given by R. N. Colwell at Fall Technical meeting of American Society of Photogrammetry/American Congress on Surveying and Mapping in San Antonio, Texas on 10 September 1984. Published in the proceedings of that meeting in his paper entitled "The Remote Sensing Picture in 1984".

10.4 References

NASA. 1984. Landsat-4 Science Investigations Summary. NASA Cont. Publ. #2326, Vol I,II. Goddard Space Flight Center, Greenbelt, MD.