AN ABSTRACT OF THE THESIS OF

SEN WANG for the degree of MASTER OF SCIENCE in FOREST MANAGEMENT presented on MAY 2, 1988

Title: MULTITEMPORAL CLASSIFICATION OF VEGETATION IN THE OREGON COASTAL RANGE USING LANDSAT MULTISPECTRAL SCANNER DATA

Abstract approved: Signature redacted for privacy.
Dr. David P. Paine

The vegetation of a 420 square mile area of the Oregon Coastal Mountain Range was mapped using data from the multispectral scanner system aboard Landsat. Advantages of this mapping system include rapid synoptic coverage of the same geographic area at different periods in time at a reduced cost compared to photogrammetric mapping. The main disadvantages are the relatively poor resolution (1.1 acres) and classification accuracy for forest vegetation types.

This project was designed to investigate the use of Principal Components Analysis (PCA) to combine data from two different dates (May and late July) in an attempt to improve classification accuracy. There were two significant results of this study.

First, the overall classification accuracy was 7.7 percent (67.4 to 75.1 percent) higher for the July as

compared to the May overpass when only single dates were used. This may be attributed to the stable phenological condition of July vegetation as compared to more variable condition in May. Spectral reflectance constantly changes over the spring growth period and varies greatly with changes in elevation.

Second, it was found that combining data from the May and July overpasses using PCA resulted in an additional increase in overall classification accuracy by another 7.5 percent (75.1 to 82.6 percent) over the July single date classification.

Multitemporal Classification of Vegetation in the Oregon Coastal Range Using Landsat Multispectral Scanner Data

by

Sen Wang

A THESIS

Submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Completed May 2, 1988

Commencement June 1988

ACKNOWLEDGMENTS

I wish to extend my sincere appreciation to all the people who provided me assistance for my study at Oregon State University. Especially, I thank my major professor, Dr. David P. Paine, who was always so nice and always willing to help. My special thanks also extend to Mr. RJay Murray, who guided me through the complexities of principal components analysis and gave me the opportunity to learn.

I express my appreciation to the international exchange program between Oregon State University and Northeast University of Forestry in Harbin, China. Especially to Oregon State Senator Mae Yih and her two year "Stephen Yih Scholarship", Dean Carl H. Stoltenberg, Dr. James R. Boyle, Dr. Perry Brown, Dr. John F. Bell, Ms. Cindy McCain, President GuoHan Xiu, and my many many professors and friends in China who helped in this program.

I express my thanks to the rest of my committee: Dr. Roger G. Petersen, Dr. William J. Ripple and Dr. Joe B. Stevens.

I am also grateful to many others who assisted me in this study: James Kiser, Art Barstow, Jamie, Marie, Susie, Janet in my department; Mr. Dennis Isaacson in the Environmental Remote Sensing Application Laboratory, Dr. Jon Kimerling, Dept. of Geography, and Dr. Richard Waring, Dept. of Forest Science, Oregon State University; Mr. John Stephenson of the Siuslaw National Forest; Milne Computer Center, Oregon State University for the computing services research grant.

Finally, special thanks to my wife, Hui Zhu, for her patience and support, and to my parents, for their great concern for my education.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Multitemporal Classification of Vegetation in the Oregon Coastal Range Using Landsat Multispectral Scanner Data

INTRODUCTION

As Paine (1981) defined in his book,

> "In the broadest sense, the term **remote sensing** involves techniques used to detect and study objects at a remote distance without physical contact.".

The process of remote sensing may be divided into three parts: data collection, data storage, and data analysis. Data collection requires a sensor such as the human eye, a camera, or a scanner. Data storage for the human eye is the brain, for a camera it is photographic film, and for scanner it is usually a computer tape or disk.

Data analysis may be accomplished by the brain and/or by computers with sophisticated software. All remote sensing systems analyze differences in emitted or reflected electromagnetic energy in one or more discrete ranges of wavelength (spectral bands). The human eye is limited to the visible light range of the spectrum between 0.4 and 0.7 micrometers. Photographic film about doubles this range to 0.3 to 0.9 micrometers. Scanners and other remote sensors greatly expand this range. When the range of sensitivity is separated and recorded

separately for two or more discrete bands, multispectral analysis can be performed. Lillesand and Kiefer (1979) stated that,

> "Probably no combination of two technologies has generated more interest and application over a wider range of disciplines than the merger of remote sensing and space exploration".

The application of remote sensing from space became a worldwide consideration after the first earth resource satellite (Landsat-1) was successfully launched in 1972. Vegetation mapping from space is a particularly interesting area of research and is useful for resource inventory, management, and planning. Much research has been done and many improvements have been made since the launch of Landsat-1. Advantages of satellite imaging are synoptic coverage and repeated (multitemporal) coverage of the same geographic area over time. These attributes enable us to obtain information for resource management in a rapid and economic manner. However, many problems remain of which two are: (1) how can classification accuracy be improved using the relatively coarse resolution data provided by the Landsat Multispectral Scanner (MSS) systems aboard the earth resource satellites, and (2) how can we better utilize multitemporal information that is available due to repeated coverage over time.

This project was undertaken to investigate these two questions. The objective in this project was to improve the classification accuracy of information derived from Landsat-acquired multispectral scanner data for use in forest management and resource planning by using multitemporal information. The plan was to use two classified MSS image segments from different seasons and to use Principal Components Analysis (PCA) as an aid in producing a new and more accurate vegetation map. The increased information resulting from phenological change in forest sites and associated vegetation canopies should decrease the uncertainty inherent in a single date classification. Through the use of PCA, observed responses from several overpasses at different dates (seasons) could be modelled and an optimal combination of temporal-spectral variables could be selected. This study was limited to two dates and the combined PCA results were compared to each of the two data sets that were analyzed separately.

A combination of two Landsat overpasses is the simplest combination of multitemporal data; yet, this combination is an excellent example for investigating the more complicated problem of information extraction from multilayered Geographic Information System (GIS) data. Furthermore, the logic and decisions required to

resolve the assignment of two attributes concerning the same object, often with the attributes in conflict or very uncertain, could be the basis for a computer-based expert system that could provide satisfactory assignment.

LITERATURE REVIEW

Characteristics of Landsat

Landsats 1, 2, and 3 were launched into circular earth orbits at a nominal altitude of 919 km (570 miles). All satellites had an orbital inclination of 99 degrees which is a nearly polar orbit. They orbited the earth once every 103 minutes (14 orbits per day). The sun-synchronous orbit caused the satellites to cross the equator at approximately the same local time (9:30 to 10:30 a.m.). Because each successive orbit shifted westward about 2875 km (1786 miles) at the equator, after 14 orbits (i.e. one day) the 15th orbit shifted westward from orbit 1 by 159 km (99 miles) at the equator. Consequently, 18 days (about 20 times a year) were needed to cover the 2875 km gap at the equator (i.e., the temporal resolution was 18 days). Because the scanning swath of each path was 185 km (115 miles) and the westward shift was 159 km, there was about 26 km (16 miles) of sidelap (14 percent) at the equator and about 85 percent at 80 degrees of latitude (Jensen, 1986, Paine, 1981).

For Landsats 4 and 5 the crossing time at the equator was changed from 9:30 to 11:00 a.m. This change resulted in higher sun angles to reduce the amount of

shadow in the imagery. Generally this change helps vegetation mapping. However, it may reduce the relief effect and increases the possibility of more cloud cover. Another change is that the altitude of Landsats 4 and 5 was changed from 919 km (571 miles) to 705 km (438 miles). This change reduced the 18 day temporal resolution to 16 days, but introduced more relief displacement over mountainous terrain (Jensen, 1986).

Landsat has orbited three kinds of sensor systems: multispectral scanner (MSS), return beam vidicon camera (RBV), and the thematic mapper (TM). For this study the emphasis is on MSS data. The MSS system was placed on each of the five Landsat satellites. The instantaneous-field-of-view (IFOV) of each MSS detector is square with a ground resolution element of about 79 m by 79 m (67,143 ft^2). There are four energy ranges sensed by the MSS system. The spectral range for each band is: band 4 from 0.5 to 0.6 micrometers (green), band 5 from 0.6 to 0.7 micrometers(red), band 6 from 0.7 to 0.8 micrometers (near infrared), and band 7 from 0.8 to 1.1 micrometers (near infrared). MSS bands 4, 5, 6, and 7 were re-numbered as bands 1, 2, 3, and 4 on Landsats 4 and 5 (Jensen, 1986). The two visible bands (4 and 5) are good for identifying cultural features such as urban areas and gravel pits. Because band 5 can better penetrate the

atmosphere, it often provides a higher contrast image. Infrared bands 6 and 7 are useful for delineating water bodies, vegetation differences, and soil condition because of high absorption by water (Lillesand and kiefer, 1979).

Because the scanning mirror oscillates ± 2.89 degrees, the scanner has an 11.56 degrees field of view which results in a swath width of about 185 km (115 miles) for each orbit. The analog signal from each sensor is transformed to a digital value using an on-board analog-to-digital converter. The data are quantized to 6 bits which are ranging from 0 to 63. These data are then rescaled to 7 bits (0 to 127) for three of the four bands (bands 4, 5, 6) after received at ground receiving stations (Jensen, 1986). The sampling rate is about 100,000 times a second for each mirror sweep in a west-to-east swath and results in a ground distance of 56m between samples. Because of this spacing, the nominal pixel (PICture ELement) size of MSS data is 56m by 79m instead of 79m by 79m and there is an overlap of about 23m on the ground (Lillesand and Kiefer, 1979).

A typical MSS scene after processing consists of approximately 2,340 scan lines and each line contains

about 3240 pixels. This is 7,581,600 pixels per channel and over 30 million observations for all four bands. One MSS scene covers an area about 185 by 178 km with approximately 10 percent endlap. It would require approximately 5,000 aerial photographs at 1:15,000 scale to cover the same area (Jensen, 1986).

Resolution Considerations

Many studies have been undertaken in an effort to map different types and levels of ground vegetation and to increase the classification and mapping accuracy. Various factors are involved and overall resolution is very important. The types of resolution may be defined as spatial, spectral, temporal, and radiometric (American Society of Photogrammetry and Remote Sensing, 1983).

Spatial resolution is a measure of the smallest angular or linear separation between two objects that can be resolved by the sensor (Swain and Davis, 1978). Spatial resolution of aerial photography is commonly evaluated as the number of resolvable linear pairs per millimeter on a photograph. Spatial resolution of other sensor systems is just the measurements of the ground-projected IFOV of the sensor system. Spectral resolution is usually defined as "the dimensions and number of

specific wavelength intervals in the electromagnetic spectrum to which a sensor is sensitive" (Jensen, 1986). Temporal resolution of a sensor system is the time intervals of repeated measurements by a given sensor over the same area. Radiometric resolution can be defined as how sensitive for a given detector to detect differences in signal strength when it records the radiant flux reflected or emitted from the terrain (Jensen, 1986).

It is usually assumed that improvements in resolution increase the probability that phenomena may be remotely sensed more precisely (Everett and Simonett, 1976). The trade-off is that higher resolution usually results in additional data processing capability for human and/or computer-assisted analysis (Jenson, 1986). Different levels of vegetation classification require different levels of resolution, and there probably is an optimal range of resolution for a particular level of vegetation classification. Also, any sensor requires a minimum amount of electromagnetic energy in order to collect the data. Lowering the sensor's sensitivity limitation is another possible approach for improving resolution.

Spectral Reflectance Features

How does the sensor on the platform gain information about an object, area, or phenomenon without being in direct contact with it? The key element involved is electromagnetic energy: visible light, infrared radiation, microwave radiation, or any other form of wave-transmitted energy. A sensor in space simply collects the electromagnetic energy reflected or emitted by the objects being sensed. The electromagnetic energy follows the basic equation of the wave theory:

$$C = f\lambda$$

where C is the velocity of light, f is the wave frequency and λ is the wavelength. Because C is a constant ($3X10^8$ m/sec), frequency and wavelength for any given wavelength are related inversely. Figure 1 on next page shows the electromagnetic spectrum on a logarithmic scale. The "visible" portion of the spectrum is relatively small because the spectral sensitivity of the human eye ranges only from about 0.4 to about 0.7 micrometers. Blue ranges from approximately 0.4 to 0.5 micrometers, green from 0.5 to 0.6 micrometers, and red from 0.6 to 0.7 micrometers. Ultraviolet energy is on the short wavelength side of the visible spectral region. Beyond the visible red region are the infrared wavelengths which contain reflected and thermal infrared

energy. The radar wavelength range (microwave portion of the spectrum) is found at much longer wavelengths (1mm to 1m) (Lillesand and Kiefer, 1979, Paine, 1981).

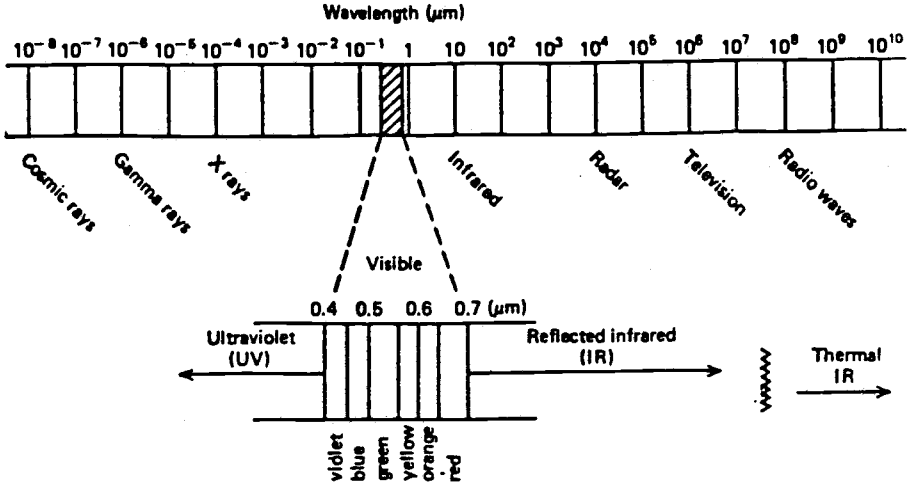

Figure 1. The Electromagnetic Spectrum (Paine, 1981).

The particle theory is important for an understanding of how electromagnetic energy interacts with matter (Lillesand and Kiefer, 1979). The energy of a quantum is given as :

$$E = hf \text{ or } E = hC/\lambda$$

where E = energy of a quantum, Joules(J)

h = Planck's constant, 6.626×10^{-34} J sec

C = velocity of light

f = wave frequency

= wavelength.

The magnitude of numbers in this equation is not important but the relationship is (Paine, 1981). It is apparent that the longer the wavelength, the lower its energy is, because the energy of a quantum is inversely proportional to its wavelength. This indicates that longer wavelength radiation is more difficult to sense than shorter wavelength radiation (Lillesand and Kiefer, 1979, Paine, 1981). For example, Thematic Mapper band 6 (thermal IR) which has the longest wavelength among all the bands in Landsat 4 or 5 required a 120m pixel for a measurable signal instead of the 30m pixel sensed in the other six bands.

The sun is the major source of radiation energy for remote sensing. However, terrestrial objects are also sources of radiation because all matter at temperatures above absolute zero emits electromagnetic radiation (Lillesand and Kiefer, 1979). The Stefan-Boltzmann Law is used to calculate how much energy an object radiates,

$$W = \sigma \cdot T^4,$$

where W = total radiant emittance from the surface of a material, W_m^{-2};

σ = Stefan-Boltzmann constant,
$5.6697X10^{-8}$ W_m-2 K^{-4};
T = absolute temperature(°K) of the emitting material.

This equation adequately describes a blackbody which is a hypothetical, perfect radiator that absorbs and re-emits all energy incident upon it (Lillesand and Kiefer, 1979). The wavelength at which a blackbody radiation curve achieves the maximum is associated to its temperature by Wien's Displacement Law,

$$_m = A/T$$

where $_m$ = wavelength of maximum spectral radiant emittance, and
A = 2898 μm°K,
T = Temperature, °K.

Because the temperature of the earth's surface materials such as soil, water, and vegetation is about 300 °K (27 °C), the wavelength of the maximum spectral radiation from these features is about 9.7 micrometers. Because this radiation relates to terrestrial heat, it is often called "thermal infrared" energy. Reflected IR and emitted IR wavelengths are usually divided at about

3 micrometer wavelength. (Lillesand and Kiefer, 1979).

All radiation caught by remote sensors passes through some part of atmosphere (the energy flow path). The path of incident and reflected sun light to a sensor in space usually passes through the atmosphere twice. However, this is not always the case. For example, only one short atmospheric path length is involved in an airborne thermal sensor (Lillesand and Kiefer, 1979, Paine, 1981).

There are two main effects when electromagnetic energy interacts with the atmosphere: scattering and absorption. Atmospheric scattering is usually caused by diffusion of particles in the atmosphere. It is very unpredictable. Absorption is usually caused by water vapor, carbon dioxide, and ozone. These molecules absorb electromagnetic energy in certain wavelengths. Those portions of the spectrum not filtered out are called "atmosphere windows". For example, multispectral scanners use the windows from about 0.3 to 14 micrometers (the thermal infrared windows are at 3 to 5 and 8 to 14 micrometers), and the radar and passive microwave systems operate through the window in the 1 mm to 1m region (Lillesand and Kiefer, 1979, Paine, 1981).

When incident electromagnetic energy hits the earth surface there are three essential energy interaction features: reflected energy, absorbed energy, and transmitted energy. Different earth features will result in various proportions of energy reflected, absorbed, and transmitted. For a given earth feature these proportions will deviate at different wavelengths. These differences help distinguish different features of an image (Lillesand and Kiefer, 1979).

The reflecting surface is either a specular or a diffuse reflector depending on the surface roughness of the object and on the incident wavelength. For example, fine sand appears rough in the visible portion of the spectrum (short wavelength) but it appears smooth in the microwave portion (long wavelength). The portion of incident energy that is reflected by an object is called spectral reflectance which is the ratio of reflected energy to incident energy for a given wavelength. A graph that shows the spectral reflectance of an object as a function of its wavelength is called a spectral reflectance curve or signature (Lillesand and Kiefer, 1979, Paine, 1981).

Vegetation

Understanding observed spectral reflectance

features of vegetation classes is critical for vegetation classification and mapping based on remotely sensed data. Figure 2 illustrates typical spectral reflectance patterns for healthy green grass, dead or senescent grass, and dry soil. There is always a "Peak-and-Valley" shape for live healthy green vegetation (Lillesand and Kiefer, 1979). Valleys in the visible

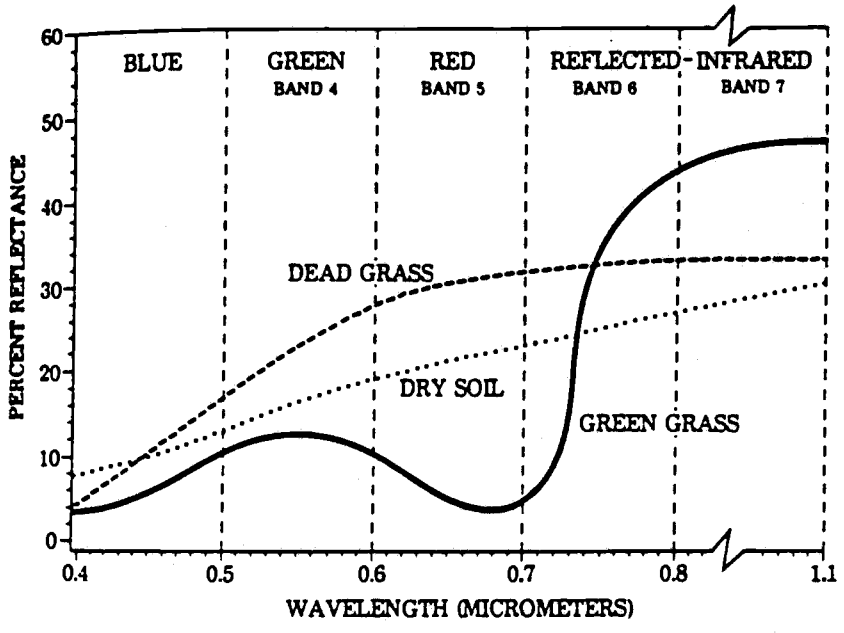

Figure 2. Spectral Signature of Healthy Green Grass, Dead or Senescing Grass, and Dry Soil (Jensen, 1986).

spectral region are mainly caused by leaf pigments, primarily chlorophylls (Gates, et al., 1965). Chlorophylls definitely absorb energy (about 80 to 90

percent) in the visible bands centered between 0.45 and 0.65 micrometers (Jensen, 1986, Lillesand and Kiefer, 1979). Healthy vegetation appears green because of high absorption of energy associated with the blue and red region and the high reflection of energy associated with green region by plant leaves (Lillesand and Kiefer, 1979). If a plant is under certain stress that interferes its normal growth, the concentration of chlorophyll pigment is reduced. This can result in less absorption in the red wavelength region (0.55 to 0.68 micrometers) and a significant shift in the spectral reflectance curve (Waring, 1986). The curve of "dead grass" shows this trend very clearly.

The reflectance of healthy vegetation increases dramatically starting at about 0.7 micrometers where about 40 to 50 percent of the incident near infrared energy is reflected (Jensen, 1986). This reflectance increase in the near infrared band (0.7 to 1.3 micrometers) is caused by scattering at the interfaces of the cell walls (Knipling, 1970). Because the internal structure of plant leaves varys among plant species, many plant species can be separated by reflectance measurements in this range. Also, because many plant stresses actually change the reflectance in this region, vegetation under stress can be detected by sensors

operating in this region (Lillesand and Kiefer, 1979).

Soil

Figure 2 shows that the soil curve has much less peak-and-valley variation comparing to the vegetation curve. Soil moisture content, soil texture, surface roughness, the presence of iron oxide, and organic matter content are some of the causes affecting soil spectral reflectance. Soil moisture content will decrease reflectance. Soil moisture content is also correlated to soil texture. Coarse, sandy soils usually have a low moisture content resulting a relatively high reflectance because of well drained condition while the poorly drained fine soils normally have lower reflectance. However, if there is no water in the soil, there will be a reverse tendency observed: coarse soils will be darker than fine soils (Lillesand and Kiefer, 1979).

Organic matter and iron oxide in the soil can also reduce reflectance. Because these factors are very variable and complicated, the reflectance patterns of a soil are consistent only within particular ranges of conditions, and it is important to be familiar with the conditions at hand (Lillesand and Kiefer, 1979).

Spectral signatures

Spectral reflectance curves are sometimes called spectral signatures (Paine, 1981). Figure 3 shows the spectral signature of coniferous and deciduous trees. The reflected infrared region provides an opportunity to separate conifers from hardwoods. As Paine (1981) mentions, spectral signatures have two valuable functions:

> "(1) they provide a comparison standard for identifying unknown objects, and (2) they are used to identify spectral regions for the differentiation of objects."

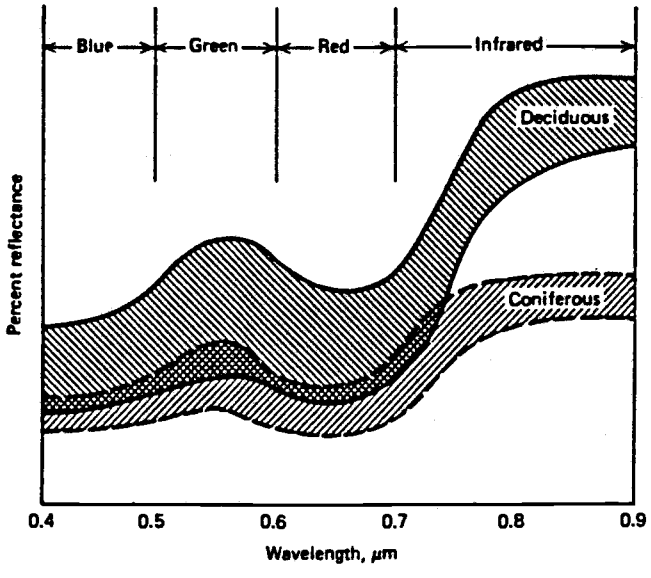

Figure 3. Spectral Signiture of Coniferous and Deciduous Trees (Wolf, 1974).

However, it might be better to call spectral signatures spectral response patterns because there is considerable variation within each wavelength for same type of object under different conditions (Paine, 1981). Spectral reflectance curves measured by a remote sensor may be quantitative but they are not necessarily unique. Temporal effects (seasonal changes), spatial effects (size, shape, and proximity), moisture stress, genetic variation within a species, and soil nutrients can produce different spectral responses for some species. Studying spectral signatures and spectral responses is important in vegetation classification because a few general rules hold; however, these guidelines are not unique and absolute. There is considerable variation under different environmental conditions (Jensen, 1986, Lillesand and Kiefer, 1979).

Spectral class plot

One of the valuable tools available for interpreting clusters of spectral signatures derived from satellite digital data is a two-dimensional plot of spectral data. For Landsat data the most common plot is the spectral response of band 4 (infrared) over band 2 (visible) for known classes and features. Similar plots can be made for TM data using TM bands 3 and 4. Generally, the two bands selected are those two least

correlated; also, the first two principal components from principal components analysis (PCA), to be explained later, can be used (Jensen, 1986, Murray, 1981, 1986).

For both MSS and TM data, including multitemporal combinations, two-dimensional plots provide adequate representation of the range of spectral responses together with similarities and differences. Plotting allows a pictorial representation of the spectral structure of a data set. It greatly helps the analyst in his interpretation (Murray, 1986).

Image Processing and Classification

Statistics

It is useful to look at the fundamental univariate and multivariate statistics of the multispectral data set at the beginning of data processing. Statistics such as the mean, standard deviation, covariance matrix, correlation matrix, and frequencies of spectral band response provide valuable information. Covariance or correlation matrices are helpful for understanding relationships between bands or variables. For example, high correlation between one band and other bands indicates that there is significant redundancy. In this

case one or more bands could be eliminated from the data set to reduce the amount of computation when a large data set is involved such as a full scene of a TM image. A histogram representation is often used in remote sensing image processing, because histograms for each band can provide a clearer view of the quality and specific features of the data (Jensen, 1986).

Image corrections

Some image correction is usually required before analyzing the image. This is usually considered as part of image preprocessing. For example, image rectification is often performed before data classification. There are two common types of errors to be considered in preprocessing: radiometric and geometric. Both kinds of error can be divided further into systematic and nonsystematic, e.g., scan skew is a systematic geometric error, and changing altitude is an nonsystematic error (Jensen, 1986). Corrections for systematic errors and for radiometric response are made when the data are processed into computer-compatible format from the raw telemetry. However, for MSS data the user may choose geometrically corrected, partially corrected, or uncorrected data depending on their application and interest (Holkenbrink, 1978).

For accurate mapping at 1:24,000 scale the user must perform additional image registration. For mapping at this scale, no matter what method is used, the registration technique relies on the use of ground control points (GCPs) located on the image and on a corresponding map in order to empirically determine a mathematical coordinate transformation to correct geometry errors (Murray, 1983). Two basic operations are used in the registration process, spatial interpolation and intensity interpolation (Green, 1983).

There are two general approaches to spatial interpolation, analytic correction and least-squares transformation. The analytic approach uses mathematical models. These models are based on the relative geometric configuration of the scene, the platform, and the sensor. However, because of the complexity of many factors, such as inadequacies of the model, errors in the estimation of model parameters, and unmodeled random distortion, "this approach often does not provide correction at the desired level of accuracy" (Ford, 1985).

The spatial interpolation using a least-squares transformation method has been presented in detail by Ford (1985) and Jensen (1986). Three steps are commonly

used in this procedure. The first step is identifying ground control points in the original image (such as Landsat MSS single-band, grey-level map) and on the reference map (such as a 7.5', 1:24,000 orthophotographic map). Then the least-square coordinate transformation can be computed by using a GCPs data set. For example, a linear regression might be used as follows:

$$X' = a_0 + a_1X + a_2Y$$
$$Y' = b_0 + b_1X + b_2Y$$

where X, Y are the positions in the reference map and X', Y' are the corresponding positions in the original MSS image.

The third step is to look at the regression residuals or root mean square error. If any residual or total root mean square error exceeds the threshold, established by the analyst before the GCP calculation, the point which has the largest error is deleted and the regression can be started again until a predetermined goal is reached.

The second operation, intensity interpolation, can be conducted after spatial interpolation. Usually

relocation of a pixel's value from the original image to the corrected image requires a resampling procedure, such as nearest-neighbor, bilinear interpolation, or cubic convolution to decide a new pixel's value in the corrected image. The nearest-neighbor resampling method is often preferred for vegetation classification not only because it is computationally efficient but also because it does not alter the pixel brightness values during resampling (Jensen, 1986).

Multispectral image classification

Image classification is one of the most important steps in the processing of remotely-sensed digital data. Multispectral image classification uses imagery collected in multiple regions of the electromagnetic spectrum. Multispectral classification may be either supervised or unsupervised classification (Jensen, 1986).

In supervised classifications, the analyst usually locates specific sites on the image to be classified that represent homogeneous examples of the interested thematic classes, such as urban, agriculture, or forest (Townshend, 1981). These sites are known by the analyst before the classification and are commonly called training sites (Jensen, 1986). The analyst may identify

and locate them through a combination of field work, aerial photo interpretation, maps, and personal experience (Heaslip, 1975).

In unsupervised classifications, the identities of thematic classes to be classified in the image are not generally known before hand because ground truth is lacking or surface features within the scene are not well defined (Jenson, 1986). The unsupervised classification consists of two steps. The first step is to generate spectral classes (spectral "clusters") according to the class mean, variance and spectral characteristics. The second step is to group spectral classes into the information classes of interest. Unsupervised classification usually requires only a minimal amount of initial input from the analyst. It is computationally efficient and is easier to handle. However, the analyst needs adequate experience and a good understanding of both spectral classification and grouping procedures. The minimum distance classifier is often used in the unsupervised classification (Jensen, 1986, Murray, 1981).

Given the same input statistics, a different classifier will produce comparable results (Hixson, et al., 1980). Variables with the greatest influence on

classification results are the training statistics used. In practical use, the classification algorithm used will be the one most commonly available. Therefore, the analyst should understand the basic principles of the classifier and how to select representative test sites and statistics (Jensen, 1986, Murray, 1981).

Although there have been numerous reports and investigations on selecting an optimal subset of bands for classification, the computational effort required makes this a questionable practice for each study. First, even "redundant" bands contain some additional information. Second, the effort need to select an optimal set of bands is not trivial with only a limited reduction in the overall computing task. Generally speaking the cycle times required to fetch even a limited set of variables far exceeds the time required for the arithmetic operations following memory or disk access.

The classifier used in multispectral classification assigns an unclassified pixel to a known class or a spectrally similar cluster (Jensen, 1986). If it is a supervised classification, the known class is the final information class. It is a spectral class to be grouped into the information class in the unsupervised clas-

sification. The minimum-distance classifier is a computationally simple and commonly used classifier. Its classification accuracy may be comparable to other more computationally intensive classifiers, such as maximum likelihood classifier (Hixson, et al., 1980). The similarity measure used in the minimum-distance classifier is sometimes slightly different, such as,

Similarity measure : $D = \Sigma [(BV - \mu)/\sigma]^2$

Euclidian distance : $D = [\Sigma (BV - \mu)^2]^{1/2}$

where: BV is the brightness value, μ represents mean value, and σ is the standard deviation (Jensen, 1986, Murray, 1981).

In the maximum likelihood classifier, each input pixel is assigned to a class according to the probability which is most probable (or most likely) when compared to all other classes. Maximum likelihood classifier requires normal distribution of the training data statistics for each class in each band (Jensen, 1986). Bayes' decision rule is often used in the maximum likelihood classifier and a priori probabilities are introduced into the classifier (Strahler, 1980). The maximum likelihood classifier is computationally

intensive. It needs more computations than either the minimum distance or parallelepiped classifier and it does not always show superior results (Jensen, 1986).

The parallelepiped classifier is a another widely used classifier. It is established on simple Boolean logic. It is a computationally efficient classifier. However, due to the overlapping problem caused by some parallelepipeds, the unknown pixel might please the rule of more than one class. Both the parallelepiped and minimum distance classifiers are nonparametric, because they do not require the normal distribution in the training data set (Jensen, 1986).

Principal Components Analysis

Principal components analysis (PCA) is a powerful tool for digital image processing. New principal component images transformed from raw data by PCA are often more interpretable than the original data (Jenson, 1986). For example, principal components analysis of spectral class mean reflectance values can provide more interpretable information in multispectral classification (Murray, 1983). PCA can also be used to compress the data set dimensionally without losing a significant amount of information from the original image. For example, the seven thematic mapper bands may

be compressed into just two or three new principal component images to reduce the data dimensionally without losing significant amount of information (Jensen, 1986).

Mathematically, the objective of PCA is to achieve the variances of the succeeding principal components as small as possible. According to this criteria the proportion of the total variance explained by the first component is a maximum. Then the proportion of the remaining variance explained by the second component is a maximum, and so on. Consequently, the proportions of the total variance explained by the last few components is minimized. It is not an easy job, however, to compress multispectral data remotely sensed by satellite because of the dimension and data size. For example, when compressing a 256 by 256 subscene of TM data by the PCA method, we have a data set with 65,536 observations each with 7 band values, and we are required to generate a 7 by 7 covariance matrix from 65,536 observations. This 256 by 256 area is less than 0.2 percent of one TM scene (Jensen, 1986, Johnson and Wichern, 1982).

Algebraically, principal components are expressed as a set of linear combinations of the n random variables X_1, X_2, ... , X_n. Geometrically, this set of

linear combinations represents the selection of a new coordinate system. This new coordinate system is generated by rotating the original system with X_1, X_2, ... , X_n as the coordinate axes. The new axes show the directions with maximum variability. These new axes give a simpler way to characterize the covariance structure because they are mutually orthogonal. It is noticeable that all mathematical operations of PCA are only related on the covariance matrix (or the correlation matrix). Development of PCA does not require a multivariate normal assumption (Johnson and Wichern, 1982).

As Johnson and Wichern (1982) discussed in their text book, suppose we have the original random vector $\underline{X}=[X_1, X_2, \ldots, X_n]$ (For example, the four bands of MSS Landsat data with n = 4). COV_x is the covariance matrix of vector $\underline{X}$, $\underline{e}$ is an eigenvector and λ is an eigenvalue of covariance matrix COV_x,

where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$

Consider the linear combinations $\underline{Y} = \underline{L}'\underline{X}$:

$$Y_1 = l_1'X = l_{11}X_1 + l_{21}X_2 + \ldots + l_{n1}X_n$$

$$Y_2 = l_2'X = l_{12}X1 + l_{22}X2 + \ldots + l_{n2}X_n$$

.

.

.

$$Y_n = l_n'X = l_{1n}X_1 + l_{2n}X_2 + \ldots + l_{nn}X_n$$

where $Var(Y_i) = l_i'(COV_X)l_i$ and $COV_(Y_i,Y_k) = l_i'(COV_X)l_k$

$(i = 1, 2, \ldots, n, \quad k = 1, 2, \ldots, n)$

According to PCA requirements, the linear combination $l_i'X$ that is the ith principal component will maximize $var(l_i'X)$ subject to $l_i'l_i = 1$ (i.e. standardized) and $cov(l_i'X, l_k'X) = 0$ for $k<i$ (i.e. orthogonal) (Johnson and Wichern, 1982). The eigenvalues and eigenvectors of COV_X will satisfy the requirements. Using the eigenvalue-eigenvector pairs (λ_1, e_1), $(\lambda_2,e_2), \ldots, (\lambda_n, e_n)$ where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$, the i^{th} principal component is (Johnson and Wichern, 1982):

$$Y_i = e_i'\underline{X} = e_{1i}X1 + e_{2i}X2 + \ldots + e_{pi}X_n$$
$$i = 1, \ldots, n$$

with these choices,

$$Var(Y_i) = e_i'(COV_X)e_i = \lambda_i$$

$$COV(Y_i, Y_k) = e_i'(COV_x)e_k = 0 \quad i = k$$

and also the total population variance $= \sum_{i=1}^{n} Var(X_i) = \sum_{i=1}^{n} Var(Y_i) = \lambda_1 + \lambda_2 + \ldots + \lambda_n$

The percent of total variance explained by each of the principal components, $\%_i$, is calculated using the equation (Jensen, 1986):

$$\%_i = \lambda_i / \sum_{i=1}^{n} \lambda_i \qquad (i = 1, \ldots, n).$$

Within each component the magnitude of e_{ki} ($e_i' = [e_{1i}, \ldots, e_{ki}, \ldots, e_{ni}]$) evaluate the importance of the k^{th} variable to the i^{th} principal component. Also e_{ki} is proportional to the correlation coefficient between Y_i and X_k (Johnson and Wichern, 1982). It may be seen from the following:

$$Yi, Xk = e_{ki}(\lambda_i / \sigma_{kk})^{1/2} \quad i, k = 1, 2, \ldots, n.$$

As mentioned above, PCA is used to explain the variance-covariance structure through a few linear combinations of original variables. The general objectives are data reduction (or compression) and facilitating data interpretation. One more point should be mentioned; that is, principal components analysis is functioned as "more of a means to an end rather than an end in themselves because they frequently serve as

intermediate steps in much larger investigations" (Johnson and Wichern, 1982).

PCA can be based on a covariance matrix that is not centered on the variable (column) means (Noy-Meir, 1973). By using a non-centered matrix, the principal components are referenced to a zero point and for spectral classes that zero point is a spectral class with no reflectance. This reference point is useful for interpreting the first principal component as a weighted measure of overall scene brightness, more so than the usual variance-centered covariance matrix (Murray, 1986).

Accuracy Assessment

If we want to use the classification results derived from a remote sensing image such as Landsat MSS data, some method or measurements are needed to evaluate classification accuracy. This usually requires the analyst to collect ground truth data which can then be used to compare with the derived classification map. Consequently, there are two classification maps: (1) remote-sensing derived map, and (2) ground truth map. The ground truth map can be obtained from ground visiting or quite often from the interpretation of aerial photography (Jensen, 1986). For example, the in-

terpretation of the 1:24,000 photography could be used as basic ground truth information for MSS or TM data. Also, the same scale and near perfect registration are usually required in the accuracy assessment (Jensen, 1986). However, the ground truth map usually is not error-free. There may be both interpretation and registration errors present, but the implied assumption is that errors in the ground truth map are minor in comparison with errors in the classification map.

The accuracy of classification results is usually expressed by calculating the percentage of correctly classified areas as compared with ground truth data. This accuracy measure is derived from sampled classified data. It is quite often given in the form of an error matrix, sometimes called a confusion matrix or a contingency table. Table 5 on page 72 is an example of an error matrix (Lillesand and Kiefer, 1979, Story and Congalton, 1986).

Either random or systematic sampling can be used to generate the error matrix. Sometimes systematic sampling is used because of the ease of application. Less often, stratified systematic sampling in MSS data is used to improve the sampling efficiency. Because of registration errors, the size of each sampling plot is at least than

9 pixels in the classified image (at 1:24,000 scale).

The number of sample points can be estimated from the binomial probability formulas (Jenson, 1986). The formula for the number of sample points to be selected is:

$$n = 4(p)(q)/E^2$$

Where: n is the sample size,
p is the expected percent accuracy,
q is the expected probability of classification error,
E is the allowable error expressed as a proportion.

In Table 5 on page 72 the total column shows the presumed true number of pixels in each class (reference data) while the total at the bottom of each column indicates the number of pixels in each class found within the sample sites on the classified area. The major diagonal elements exhibit the agreement between the classified and reference data. The ratio of the total number of correct classifications (the sum of the major diagonal elements) to the total number of samples taken is the overall accuracy for the classified area.

The overall accuracy can provide a general estimate of accuracy. It does not, however, provide information on the accuracy of the individual classes. Sometimes there are statistical differences among these accuracies of the individual classes. The off-diagonal elements can provide more information on errors of omission and commission. Errors of omission for each class is the ratio of the total number of pixels assigned to incorrect categories along each row to the total number of true pixels in the category. Errors of commission is the ratio of the total number of pixels assigned to incorrect categories along each column to the total number of pixels assigned to the column category (Jensen, 1986, Story and Congalton, 1986).

Traditionally we usually use the ratio of the number of correctly classified samples of a certain category to the total number of the assumed true samples (ground truth) of that category as accuracy assessment. This percentage actually shows how a reference (ground) sample will be correctly classified. This is actually related to the errors of omission. Story and Congalton (1986) called this as the "producers accuracy" because the producer of the classified image "is interested in how well a specific area can be mapped". They also mentioned that

> "an important, but often overlooked, point is that a misclassification error is not only an omission from the correct category but also a commission into another category".

They defined the percentage of the number of correctly classified samples of a certain category divided by the total number of samples that were classified in that category (i.e., the column total) as "user's accuracy". This is actually related to the commission error. The user accuracy provides the user with the reliability of the map, "or how well the map represents what is really on the ground". It is important to know that both the "producer's" and the "user's" accuracy are needed to have a better evaluation of accuracy because "using only a single value can be extremely misleading" (Story and Congalton, 1986).

One more question about accuracy assessment is how to quantitatively compare two or more different remote sensing data classification results under varied conditions. Two approaches are generally used in making these comparisons: analysis of variance and discrete multivariate analysis (maybe called contingency table analysis). Because analysis of variance uses only the diagonal elements, requires normally distribution and independence of the categories in the error matrix, discrete multivariate analysis techniques are usually preferred (Congalton and Oderwald, 1983). One of the

discrete multivariate analysis methods is to test the overall agreement between two separate error matrices. The measure of agreement, called KHAT (i.e., K), is calculated by:

$$K = \left(N \sum_{i=1}^{r} X_{ii} - \sum_{i=1}^{r} (X_{i+} * X_{+i})\right) / \left(N^2 - \sum_{i=1}^{r} (X_{i+} * X_{+i})\right)$$

where r is the number of rows in the error matrix, X_{ii} is the i^{th} element of the error matrix, X_{+i} and X_{i+} are the marginal totals for row i and column i respectively, and N is the total number of observations (Bishop et al., 1975, Congalton and Oderwald, 1983).

The approximate large sample variance, $\sigma(k)$, can be calculated. The formula can found in Bishop's (1975) book or Hudson and Ramm's (1987) paper.

The test statistic for significant difference in large samples is given by:

$$Z \quad (K_1 - K_2)/(\sigma_1 + \sigma_2)^{1/2}$$

This test uses the normal curve deviate (Z) to determine if the two error matrices are significantly different assuming two KHAT's are independent (Cohen, 1960).

Using this method we can, for example, test different classification algorithms, determine which date of imagery yields the best results, or compare the imagery from different sensors. However, this method is "limited in that only one factor in the classification may vary at a time" (Congalton and Oderwald, 1983).

STUDY AREA

The study area includes over 1087 square kilometers (420 square miles) located in the Coast Range of western Oregon between 44 degrees 15 minutes and 44 degrees 45 minutes north latitude and between 123 degrees 45 minutes and 124 degrees east longitude. The area is covered by eight 7.5-minute topographic quadrangles made by United States Geological Survey (USGS). The names of these topographic quadrangles and correspondent orthophotographic quadrangles are listed below:

Topographic quadrangle	Orthophotographic quadrangle
TOLEDO NORTH	TOLEDO NW
EDDYVILLE	TOLODO NE
TOLEDO SOUTH	TOLEDO SW
ELKCITY	TOLEDO SE
TIDEWATER	TIDEWATER NW
HELLION RAPIDS	TIDEWATER NE
CANNIBAL MOUNTAIN	TIDEWATER SW
FIVE RIVERS	TIDEWATER SE

The Alsea and Yaquina rivers pass through the area on the way to the sea. More than half of the area is within the Siuslaw National Forest with the rest

comprised of private, state, and Bureau of Land Management (BLM) lands.

The Coast Range expands from the middle fork of the Coquille River in Oregon northward into southwestern Washington. The elevation of main ridge summits range from about 450 to 750 meters (1476 to 2461 ft). Scattered peaks are often capped with intrusive igneous rocks. Marys Peak at 1,249 meters (4098 ft) is the highest. Soils over most of the Coast Range are Haplumbrepts. They were derived from basalt. Their color are usually reddish-brown. They are relatively stone free. Surface textures are generally clay loam. Some Haplohumults on basalt parent materials can be found in the southern portion of the range (Franklin and Dyrness, 1973).

Climate over the Coast Range is uniformly wet and mild. There are variations, however, in this wide range because of latitude and elevation. The annual precipitation varys from 1,500 to 3,000 millimeters (59 to 118 inches). Average mean annual temperatures are usually from 8 to 10 degrees centigrade. It is relatively dry during the summer. However, frequent fog and low clouds are helpful to reduce moisture stresses (Franklin and Dyrness, 1973).

The Coast range is a classic coniferous forests of the world. It is a

> "popularly known region of Douglas-fir (Pseudotsuga menziesii (Mirb.) Franco.) dominance, of western hemlock (Tsuga heterophylla (Raf.) Sarg.) climaxes, and coastal sitka spruce (Picea stitchenis (Bong.) Carr) 'rain forest'" (Franklin and Dyrness, 1973).

Some outstanding features of these forests are the nearly total dominance, the size, and the longevity of the dominant species. Other coniferous species consist of grand fir (Abies grandis (Dougl.) Lindl.) western red cedar (Thuja plicata Donn.), western larch (Larix occidentalis Nutt.), etc. The common deciduous hardwoods include Oregon white oak (Quercus garryana Dougl.), bigleaf maple (Acer macrophgllum Pursh), and red alder (Alnus rubra Bong.) (Franklin and Dyrness, 1973).

As mentioned above, Douglas-fir is a main species dominating more than half of the forested area. Alder is the most common hardwood species appearing in nearly pure stands in drainage bottoms, and from the bottom to top of ridges or mixed with the conifers. There are some old growth Douglas-fir and spruce stands still left within the study area, such as the old growth stands near Canal and Drift creeks located in the USGS TIDEWATER 7.5-minute topographic quadrangle. Most of the

forested areas are young growth, including mature Douglas-fir regenerated a century ago, small saw timber and pole size conifers, regeneration, mixed hardwood-coniferous stands, brush, and both new and old clearcuts.

PREVIOUS WORK

As part of another project, Richard McCreight[1] of the Department of Forest Management, Oregon State University, completed a preliminary vegetation mapping project using MSS digital data from a single date (May 22, 1983). He started out with several vegetation classes but condensed them into five management classes. These classes were based on some of the information needed by forest managers and planners. Little consideration was given to how well management classes might match the MSS spectral classes. In other words, the management classes were defined and then a best match was made with the spectral classes. The management classes were:

Harvest -- Mature conifer including small saw timber and larger. The age class may be 80 years and older;

Thinning -- Pole-size conifer sites including both thinned and not thinned;

Release -- Hardwoods, hardwoods mixed with few conifers, brush land, regeneration site mixed with brush and shrubs;

---

[1]Unpublished research by Richard McCreight of the Department of Forest Management, Feb. 19, 1985.

Planting -- Clearcuts, both burned or unburned, agriculture land, or grass land with more soil exposure; Other -- Water, urban, or build-up land.

McCreight's final results (unpublished) are shown in Table 1.

Table 1. Results of McCreight's (1985) Single Date MSS Classification.

Class	Acreage	percent(%)
Harvest	61,024	21.9
Release	100,883	36.2
Planting	36,269	13.0
Thinning	75,726	27.2
Other	4,935	1.7
TOTALS	278,837	100.0

METHODS

An objective of this study was to improve the mapping accuracy of McCreight's single date MSS classification using multispectral data and PCA. A logical approach might have been to compare the accuracy of the results of this study with McCreight's results. This was not feasible, however, because: (1) additional ground truth points were used in this project because it was decided that McCreight's project was inadequate in this respect and (2) different spectral class groupings were used in this project. To be consistent it was necessary to re-analyze the 1983 single data set to be compared with the results of PCA analysis using two dates. Thus, both the ground truth and spectral analysis were conducted by the same individual making the single date and multi-date analyses directly comparable.

Field Work and Ground Truth

Visiting the study area prior to classification usually helps the analyst understand the particular situation on the ground, such as geographic differences, species variety, seasonal changes, soil types, and moisture stress. All these factors help in the understanding of spectral reflectance patterns of vegetation and in the selection of training sites. The

knowledge gained from field work is useful for grouping spectral classes into the information classes.

However, due to cost and accessibility problems, National High Altitude Photography (NHAP) was used in this project as the primary source of ground truth with some field checking. Before the field trip, photo interpretation was done to locate several representative vegetation types found within the study area. During the field check, corrections were made to improve the quality of NHAP photo interpretation. Several typical sites in the Tidewater NW quadrangle were visited during August, 1986, which were mostly located along both sides of the Alsea River drainage. The sites visited provided representation examples of vegetation classes present.

The NHAP photography was flown in July, 1982. The scale was 1:58,000 (1 inch equals about 0.9 miles) and the photographs were color infrared positive transparencies (9in. X 9in. format). Each frame, acquired at 40,000 feet, covers nearly 68 square miles of terrain. The quality is much better than average. The vegetation classification system was based on a photo legend system developed by M. Hall of Environmental Remote Sensing Application Laboratory (ERSAL) for McCreight's study.

Landsat MSS Data

Two MSS scenes were used in the project and, due to the high cost of the new tapes, two existing tapes from the ERSAL Library were selected. One MSS scene was from the May 22, 1983, overpass and the other was from the July 29, 1979, overpass. In spite of four years difference it was assumed that the acreage of most natural forest vegetation changes were small and that man-induced changes could be accounted for. The primary man-induced changes include clearcuts, fire (both natural and man-caused), and thinning.

When comparing MSS data from scenes acquired four years apart and from different seasons, there are two major influences on the spectral response of vegetation. The dominant influences are the seasonal phenological differences resulting in different spectral characteristics of the May and July overpasses. Secondary influences are changes caused by man over the four year period. With the exception of clearcuts and other major changes, the spectral changes associated with vegetation within a four year time frame are minimal when compared with the May-to-July seasonal changes. In May, new leaves have broken bud, but are succulent and immature. By the end of July all new

leaves were fully developed and in a mature stage.

Registration

The 1983 data had been geometrically corrected as part of McCreight's study, and only the 1979 data were geometrically corrected in this study. The 1:24,000 scale USGS orthophotographic quadrangle maps were used as reference maps. For each map five ground control points were selected with 4 points near the 4 corners and one near the center. Thus, 40 control points with both Universal Transverse Mercator (UTM) and MSS image coordinates were recorded for regression. The linear regression equations used are:

$$Y_1 = a_0 + a_1X_1 + a_2X_2$$
$$Y_2 = b_0 + b_1X_1 + b_2X_2$$

where, Y_1 is column number of MSS data,
Y_2 is scan line number of MSS data,
X_1 is coordinate of UTM Easting,
X_2 is coordinate of UTM Northing.

The computed coefficients are:

$a_0 = 1377.57$ $\quad$ $b_0 = 10936.0$

$a_1 = 1.03797$ $\quad$ $b_1 = -0.302309$

$a_2 = -0.243179$ $a_2 = -1.30338$

Forty residuals were calculated for each regression equation. Because all residuals were smaller than the pixel dimensions, no control points were deleted. Nearest-neighbor resampling was used to generate a 1:24,000 scale line-printer map (Murray and Alexander, 1983). A common line-printer character size is 8 lines per inch and 10 characters per inch so that at 1:24,000 scale, one line printer character represents a 250 (high) by 200 (wide) ft^2 area which is about 1.15 acres.

Classification

A separate spectral classification was completed for each of two MSS scenes. An unsupervised minimum distance classifier was used to generate spectral classes. This required two steps. The first step was cluster building or prototype class generation (Murray, 1981). Decisions by the analyst were required to generate prototype classes, e.g., when should a new spectral class be formed or when should two spectral classes be merged? These decisions depended in part on previous experience, the particular research interest, minimum number of pixels in each class, and maximum number of classes to be separated. The final classification took several iterations of step 1 and

step 2. The next step was the assignment of pixels to one of the classes using the minimum distance classifier. A similarity measure (Murray, 1981) was used as follows:

$$D_{ik} = \sum_{j=4}^{7} ((X_{jk} - \mu_{ji}) / \sigma_{ji})^2$$

where D_{ik} is the similarity measure,

μ_{ji} is the mean value,

σ_{ji} is the standard deviation,

k is the index of the k^{th} pixel from a set of n (unknown) pixels; k = 1,2,3,...,n,

j is presently limited to values 4,5,6,7; i.e., the four Landsat bands,

i is the i^{th} class from a set of m "prototype" or spectral classes; i = 1,2,3,...,m.

Note that D_{ik} is calculated for all m classes for each pixel X_k. Then the pixel X_k is assigned to the class i for which D_{ik} is the smallest. If we have n pixels and m spectral classes there is a total of n X m D_{ik} values that must be calculated in each iteration. The means and standard deviations were calculated for each class based on the pixels assigned to the class by the minimum distance classifier (Murray, 1981). There were 68 spectral classes (62 symbols) in 1979 data and

49 spectral classes in 1983 data. Summaries of these classes statistics are in APPENDIX A and B.

Once the spectral classes were obtained, they were grouped into information (or management) classes (Murray, 1981). Because the spectral classes were generated mathematically by the similarity measure, usually there were spectral classes which were not in a one-to-one correspondence with the real information classes, or one spectral class contained two or more information classes. Where possible, these classes were identified. It was necessary to know the spectral characteristics of the study area well enough to correctly group a large number of spectral classes into a smaller number of information classes.

In this study, sites with known information classes were randomly selected and matched to the spectral class maps. All pixels within the matching ground sites were counted and used to categorize the spectral classes. The spectral classes with the most frequent occurrences best represent the corresponding information class. However, several problems occurred. Sometimes not all of the spectral classes appeared in the selected ground truth site. Another problem was that sometimes there was no strong association between a spectral class and one information class; instead, the spectral class was

associated with several information classes. This required an additional backwards-matching method in the grouping procedure. Large, spatially homogeneous blocks of spectral symbols were located and matched with an interpreted or known information class. Thus, it was possible to find stronger (more probable) associations, or that a spectral class represented two or more different information classes.

Graphic plots were helpful when grouping spectral classes. Figure 4 on next page is a 4-band spectral reflectance value plot of typical vegetation classes derived from the July 1979 data. The advantage of this kind of plotting is that changes in the spectral response values by class are shown over the four bands. The plots are similar to spectral response patterns or "spectral signature" plots. Thus, spectral class means were plotted and compared with known response patterns. However, when there were many spectral classes and especially when the range of the values was close, the lines representing the spectral classes became mixed and difficult to recognize. When spectral response curves for several classes were similar in shape and magnitude, they were grouped into a single information class.

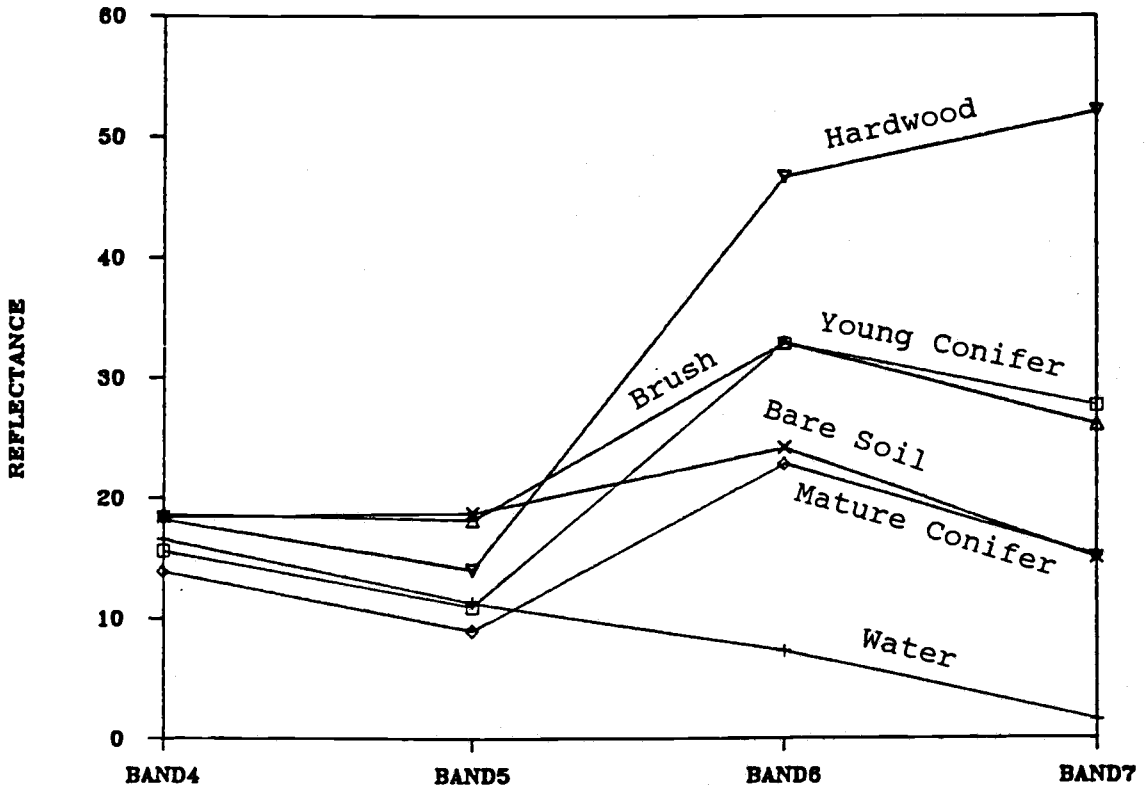

Figure 4. Spectral Signiture Plot of Selected Vegetation Classes from the July 1979 Data (+ Water, ◇ Mature Conifer, □ Young Conifer △ Brush, X Bare Soil, ▽ Hardwood).

A two-dimensional plot of the spectral classes produces a good pictorial representation of class proximity in spectral space. Usually a band 5 versus band 7 plot is used for Landsat MSS data (Murray, 1981). The shape of spectral classes indicates two-dimensional standard deviation limits. Figures 5 and 6 are plots of

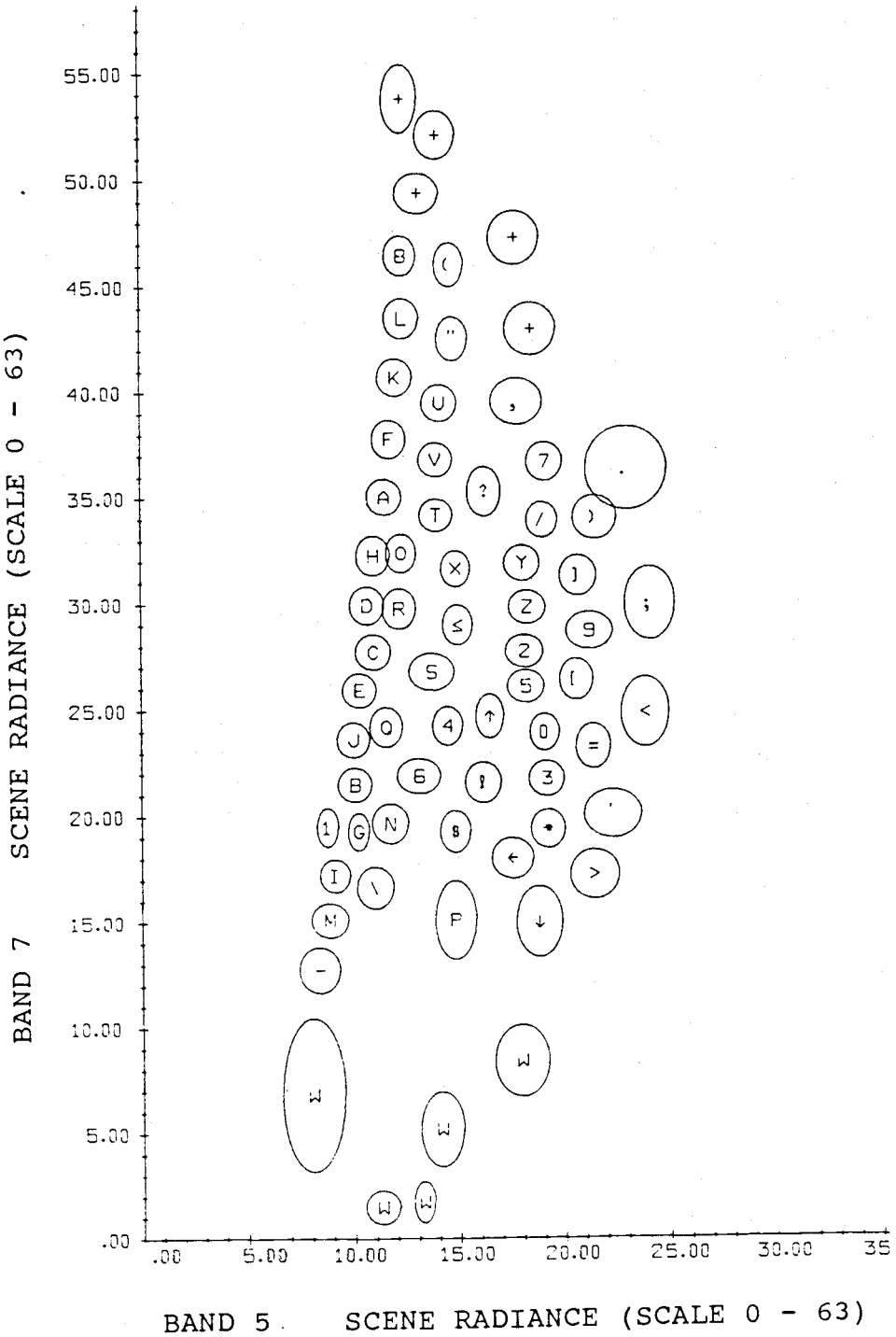

Figure 5. Band 5 versus Band 7 for 1979 Spectral Mean Classes and Standard Deviation Limits (See APPENDIX A for Symbol Assignments).

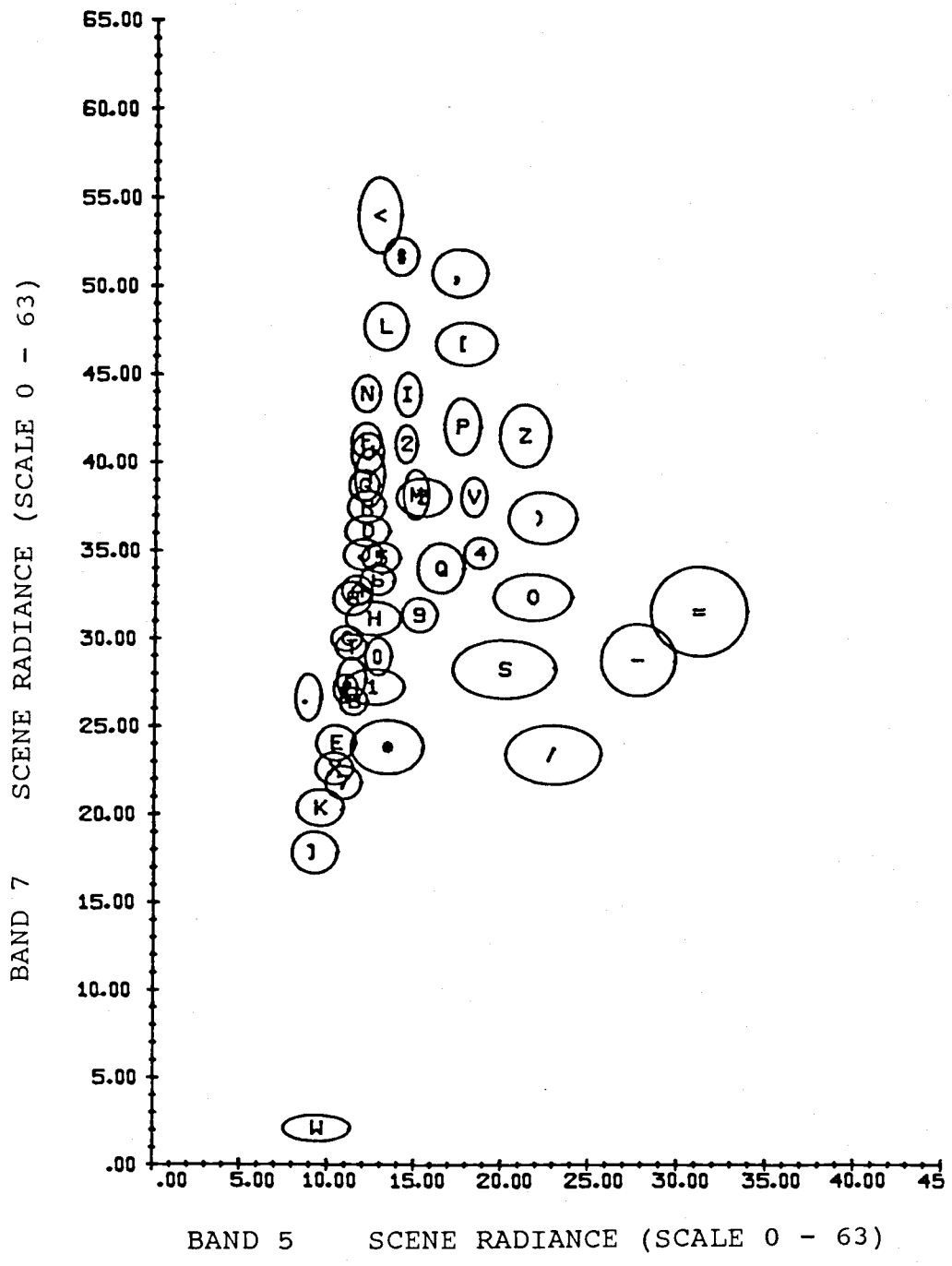

Figure 6. Band 5 versus Band 7 for 1983 Spectral Mean Classes and Standard Deviation Limits (See APPENDIX B for Symbol Assignments).

the spectral classes from the 1979 and 1983 data. From an understanding of the spectral features and comparisons with spectral class responses from known ground sites, we can predict what these spectral classes represent. For example, in Figure 5, the symbols M, I, G, B probably represent mature conifers because the values of both bands 5 and 7 are lower than most other classes; also, the symbols +, 8, L, K mostly represent hardwoods and brush because their values of band 7 are high and the values of band 5 are relatively low (cf. Figure 4).

Principal components analysis of spectral class mean reflectance values was also used in the grouping procedure. The first two components out of four possible components for a single scene of MSS data usually represent more than 90 percent of the variation, and plotting the first two principal components may provide additional information about spectral-to-information class assignments. Figures 7 and 8 are PCA plots for 1979 and 1983 data respectively. In comparison with band 5-7 plots, PCA plots usually provide more information about spectral class features (Murray, 1983, 1986). Comparing PCA plots with band 5-7 plots (Figure 5 versus Figure 7, Figure 6 versus Figure 8), it was found that the general trends are similar but the PCA plot shows a

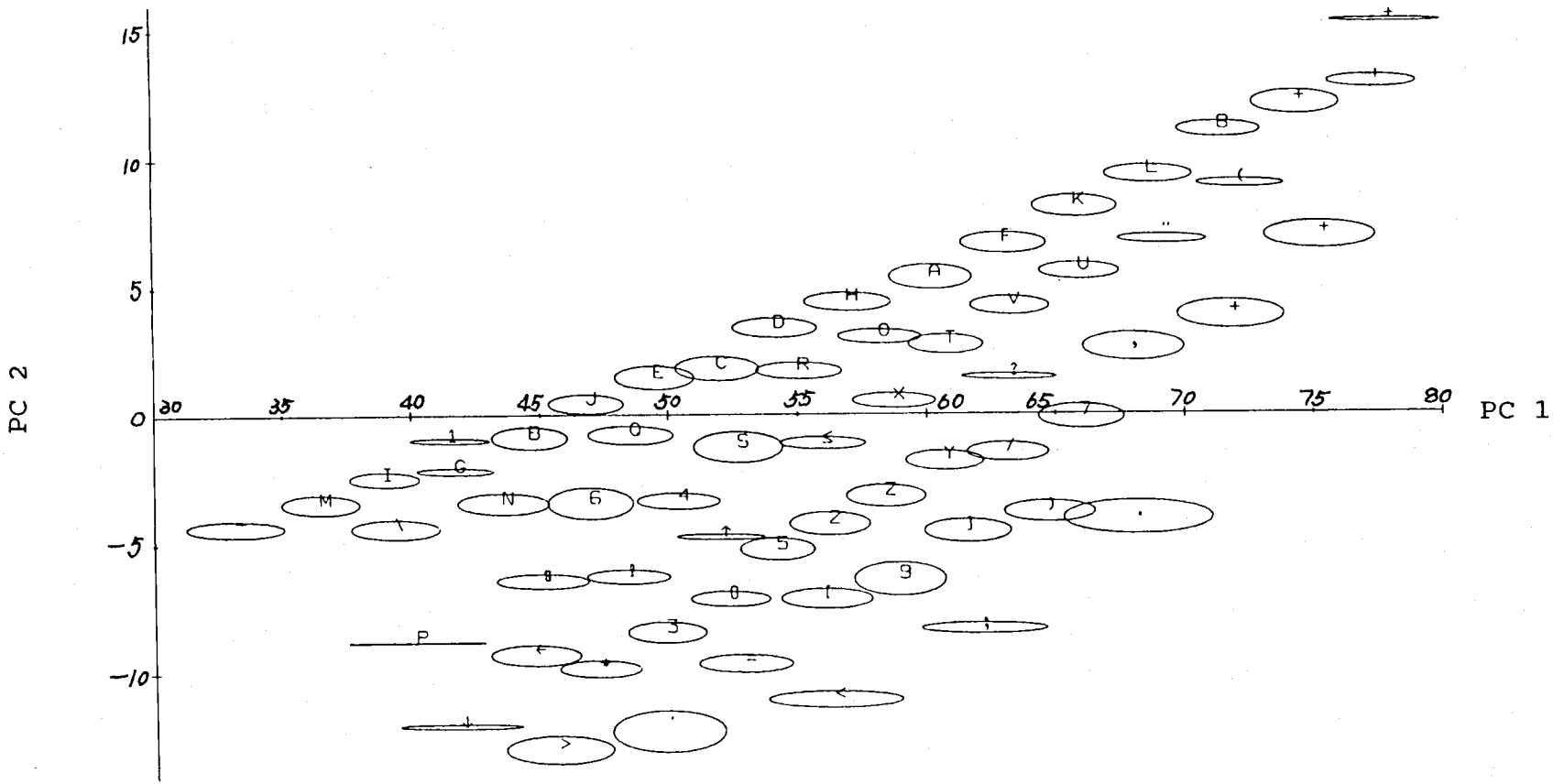

Figure 7. PCA Plot of 1979 Spectral Mean Classes and Standard Deviation Limits (See APPENDIX A for Symbol Assignments).

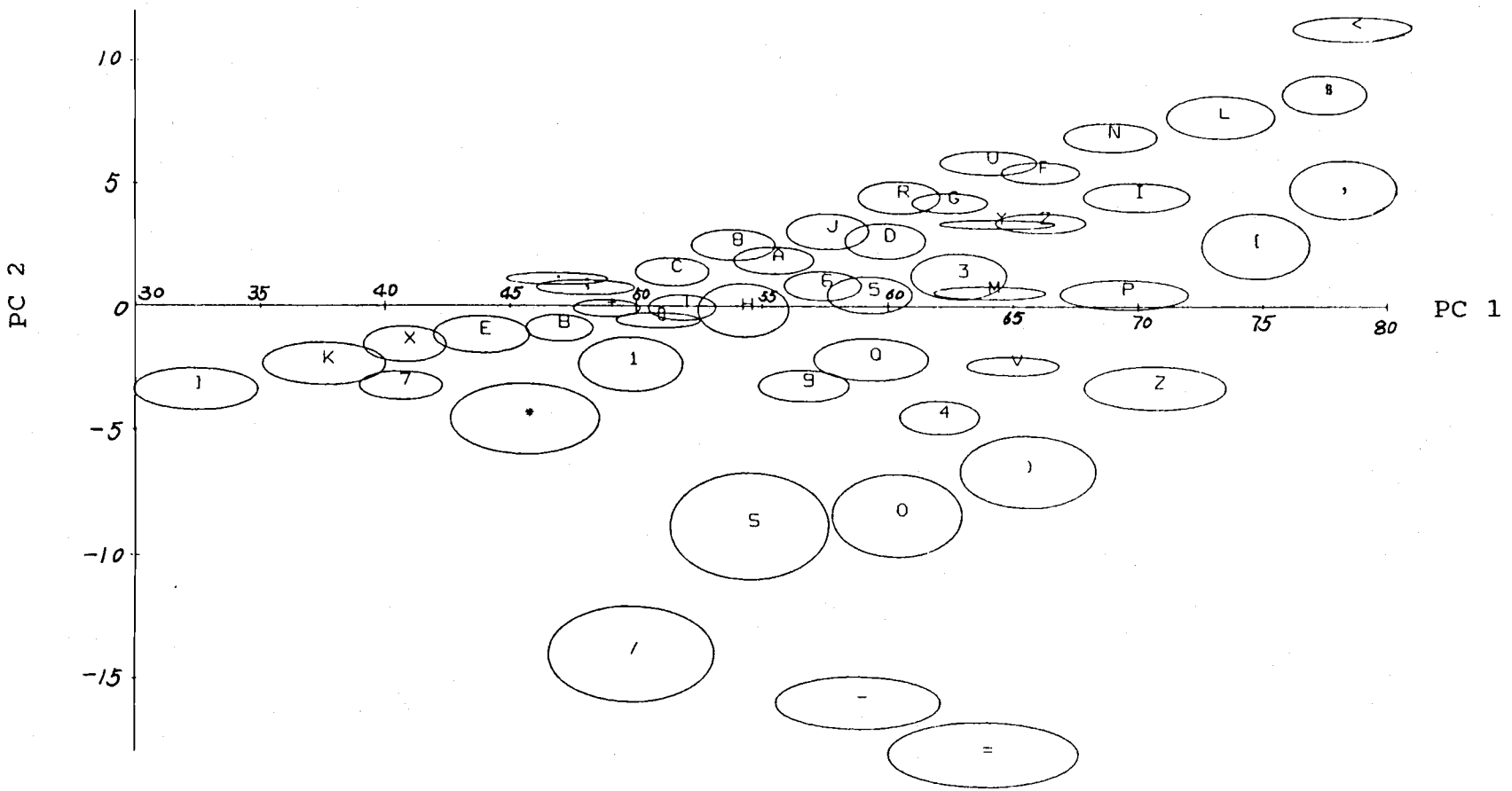

Figure 8. PCA Plot of 1983 Spectral Mean Classes and Standard Deviation Limits (See APPENDIX B for Symbol Assignments).

better separation for some classes. With the help of PCA plots and band 5-7 plots plus the two-way matching (from ground truth to spectral classes and from uniform spectral class blocks to ground truth), a grouping was done. The final grouping decisions were made after three iterations of grouping on the 1979 data and one additional iteration on the previously grouped 1983 data (McCreight's third grouping). The 68 spectral classes were grouped into 14 information classes for 1979 data and 49 spectral classes into 16 information classes for 1983 data. A final grouping into five management opportunity classes was completed for both dates. By using non-centered covariance matrices the principal components are on approximately the same scale for both years and have the same reference points. The error ellipses in Figure 7 and 8 are a mapping of the original one-standard deviation ellipses (Figures 5 and 6) into principal component space.

Verification of Single Date Classification

Verification of classification accuracy was done on both single date classifications. In each of the eight 7.5 minute quadrangles, four one-km by one-km verification sites were randomly selected by UTM grid coordinates. Then the locations of the 32 verification sites were transferred to the NHAP photos. After photo

interpretation was completed for those blocks, the boundaries of each information class interpreted were transformed back to the orthophotographic quadrangles using a stereo zoom transfer scope. By transferring boundaries from the photo interpretation to orthrophotos, the NHAP coordinates were mapped into UTM coordinates at 1:24,000 scale. The overlays with the ground truth matching the selected sites were then put on the top of computer-generated line printer maps at the same scale and the occurrences of spectral class symbols tabulated. Tables 2 and 3 (page 68 and 69) show the classification results for both dates.

Temporal Merging

The two geometrically registered scenes of MSS data were then merged and each 250 by 200 ft ground pixel (scaled at 1:24,000 and referenced to UTM grid coordinates) now had a two-symbol representation based on the classification results from two different dates. For example, a symbol pair "$=" corresponds to a pixel labelled as mature conifer in the 1979 classification and a clearcut in the 1983 classification. If both of classifications are correct and, if the geometric registration is accurate, this pixel would represent a clearcut made after July 29, 1979, and before May 22, 1983.

In the two-symbol mapping, there would be 288 possible combinations of information classes because there were 16 classes in 1979 data and 18 in 1983 data (16 X 18 = 288). However, because some combinations did not actually exist, only 255 combinations were found. These 255 combinations were sorted by symbol and by occurrence (see APPENDIX C for a sample page of the sorted results). Then 61 out of 255 combinations were selected for principal components analysis. The selection criteria were:

a) a pair must have more than 500 occurrences;

b) no unreasonable changes resulting from spatial mis-registration (e.g., no changes from mature conifer to thinning, clearcut (planting) to mature conifer);

c) combinations with no changes (e.g., thinning-thinning, release-release).

Thus, the 61 symbol pairs represent the most plausible and most frequent change combinations resulting from seasonal and year-to-year influences.

There were eight spectral class means for each combination. Because these information classes were grouped from spectral classes, these eight means were recalculated, e.g., the "$" symbol of 1983 came from a

combination of I, K, X, E, etc.; the means were recalculated as follows:

$$MG_k = \sum_{i=1}^{n} F_{ki}M_{ki} \; / \; \sum_{i=1} F_{ki} \qquad k = 4, 5, 6, 7$$

where MG_k = the grouped new mean of band k,

F_i = the frequency of i^{th} spectral class of band k,

M_{ki} = the mean of i^{th} spectral class of band k.

Principal components analysis was then conducted on the eight composite means for the 61 combinations (see APPENDIX D for the completed list of the means, 1979-1983 symbol pairs, and single symbols). The non-centered covariance matrix was generated, and the eigenvalues and eigenvectors were then calculated. The first five principal components for 61 combinations were printed. Finally, the first two components for 61 combinations were plotted. Figure 9 shows the first two components with 61 symbols. APPENDIX E contains the computer output for the PCA computation results. Using ground truth, the sorted 255 combinations, and the PCA plot, the 255 combinations were grouped into 20 symbols in a first grouping and 11 symbols in a second grouping. These 20 and 11 class groups were further compressed into four information classes (see APPENDIX D for the list of these symbol assignments).

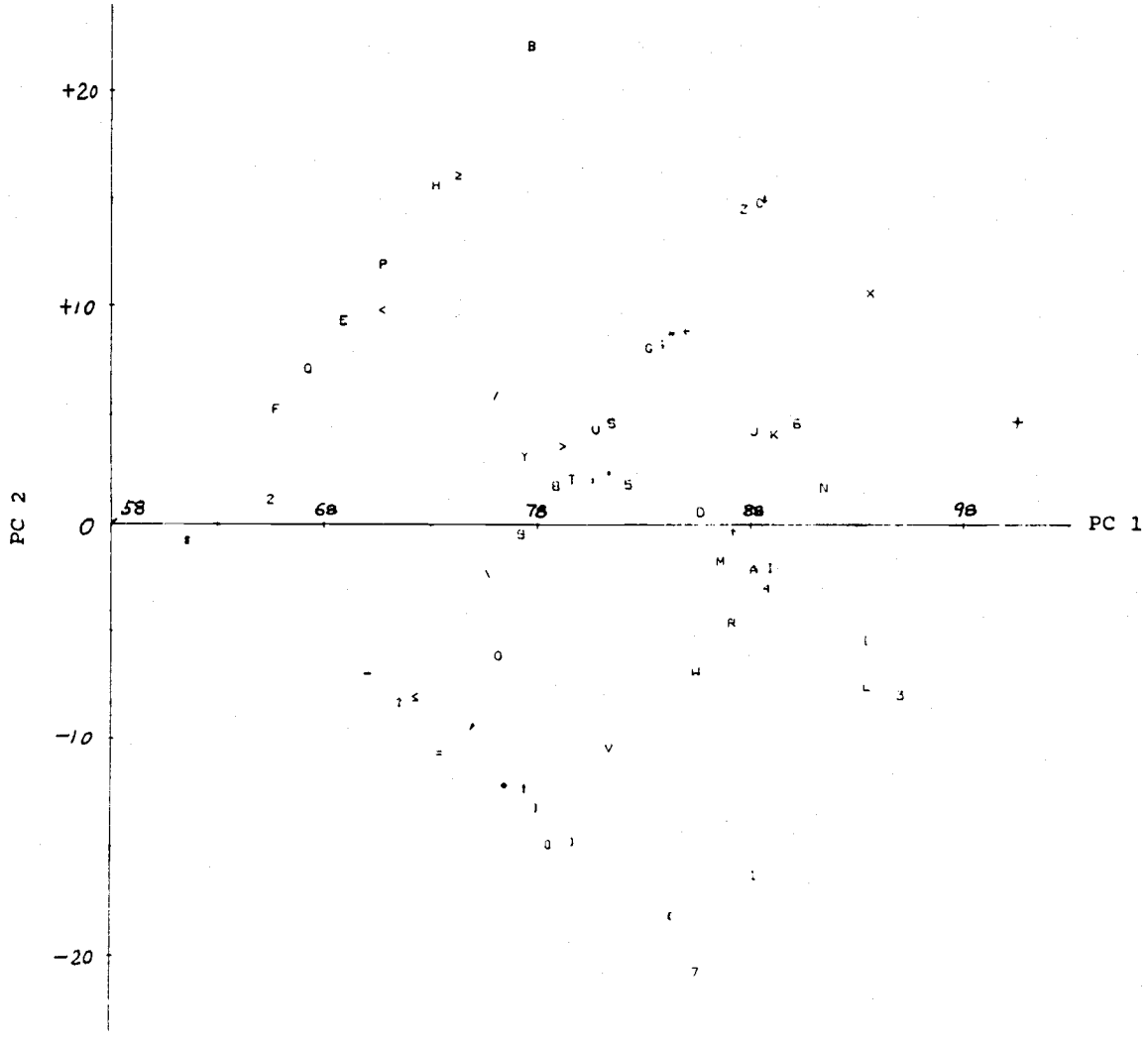

Figure 9. PCA plot of 61 Multitemporal Spectral Mean Classes and Standard Deviation Limits (See APPENDIX D for Symbol Assignments).

Programing for the Project

Several different computer programs were required for this project. The main program, "PIXSYS", was developed by the Milne Computer Center for ERSAL. In addition several new programs were developed specifically for this project by RJay Murray of the Milne Computer Center, assisted by the author. All new programs were coded in FORTRAN and run on the Milne Computer Center CYBER (Control Data Corporation) 170/720 computer.

Program TEMPMERG was developed to merge classification output from several MSS scenes. Up to five MSS scenes can be merged using this program, and a new header containing information from the merged scenes was created. The file <MERGED> is the primary output of this program. Using <MERGED> as input, program SYMATCH sorts and counts the occurrences of all class symbol combinations by either first symbol or by decreasing occurrences. Double hashing (Knuth, 1973) was used to find distinct combinations of class symbols, and matching collisions were used to count the occurrences for each combination. Program GRPMEAN was written for calculating the new means of an information class formed by grouping several spectral classes. An output file titled <NEWMEAN> contains the new mean values. Program

PCASAMP was used to generate a file of multitemporal means for PCA. Mean spectral responses, the <NEWMEAN> files for each date, the file <SORTED> from SYMATCH, and a file containing the temporal symbol string and replacement symbol were merged in PCASAMP. The output file <CLMEAN> was generated by PCASAMP. The last program developed is called TEMPREP which assigns a single symbol to the same pixel to produce the line printer map in merged scenes after final grouping.

RESULTS

Acreage Classification

Tables 2 and 3 show the acreage results for each single date classification (In McCreight's study [see Table 1, page 46], the corner of one quadrangle was clipped off, thus resulting a small difference in the reported acreage totals found in this study).

Table 2. July, 1979, Data Classification Results.

Class	Acreage	Percent(%)
Harvest	94,456	33.3
Release	108,431	38.2
Planting	22,083	7.8
Thinning	49,723	17.5
Other	8,847	3.2
TOTALS	283,540	100.0

Table 3. May, 1983, Data Classification Results.

Class	Acreage	Percent(%)
Harvest	74,681	26.3
Release	83,226	29.4
Planting	38,845	13.7
Thinning	77,269	27.3
Other	9,519	3.7
TOTALS	283,540	100.0

From these two tables a general trend is evident. Harvest and release acreages decreased about 15 percent and planting and thinning increased about 15 percent between the 1979 and 1983 analyses. However, changes of this magnitude are not expected over only a four year period. It is very likely that the major reason for this change is misclassification with either or both of the single date classification results. One could hypothesize that the July analysis is more accurate because the phenological condition of the vegetation is more stable at this time of year. In May there is considerable phenological difference between high and

low elevation vegetation. As will be shown later, the July single date analysis proved to be more accurate than the May single date analysis.

Acreage statistics of each management class are shown in Table 4 for the multitemporal classification using PCA. The percentages of harvest and release were close to the 1979 single date classification, however, planting and thinning changed substantially.

Table 4. Multitemporal Classification (July, 1979 and May, 1983) Results.

Class	Acreage	Percent(%)
Harvest	93,922	33.1
Release	102,654	36.2
Planting	48,319	17.1
Thinning	28,625	10.1
Other	10,020	3.5
TOTALS	283,540	100.0

Accuracy Comparison

Comparison of single date classifications

Tables 5 and 6 are the error matrices of the July, 1979, and May, 1983, single date classification results. The overall accuracy of the July, 1979, classification is 7.7 percent higher than for the May, 1983, classification. As stated earlier, this may due to a more stable phenological condition of the vegetation in late July than in May. Although the producer's accuracy (omission error) for thinning is better in 1983 date classification, the user's accuracy for thinning of both 1979 and 1983 classification is very low (less than 30 percent). In order to quantitatively compare the result, a contingency table analysis was performed. Congalton's (1987) program was used to calculate the KHAT and Z statistics. The KHAT of 1979 data is 0.65624 and 0.56353 for the 1983 data. The test statistic ($Z = 8.6982$) shows that the difference between two error matrices is significant (< 0.01). See completed print out in APPENDIX F for this test. The 1979 classification shows a significantly higher accuracy than 1983 classification.

Table 5. Error Matrix of the July, 1979, Data Classification Results.

Reference Data	Classified Data H	T	R	P	O	Total	Producer's Accuracies
H	2,211	387	103	12	4	2,717	81.4 %
T	94	309	103	1	0	507	60.9 %
R	165	394	2,030	46	2	2,637	77.0 %
P	19	20	156	458	0	653	70.1 %
O	46	1	25	74	7	153	4.5 %
Total	2,535	1,111	2,417	591	13	6,667	
User's Accuracies	87.2%	27.8%	84.0%	77.5%	53.8%		

Sum of the major diagonal = 5,015

Overall Accuracy = 5,015/6,667 = 75.2%

H - Harvest; T - Thinning; R - Release;

P - Planting; O - Other.

Table 6. Error Matrix of for the May, 1983, Data Classification Results.

Reference Data	Classified Data H	T	R	P	O	Total	Producer's Accuracies
H	2,013	590	106	33	12	2,754	73.1 %
T	66	413	75	9	0	563	73.4 %
R	169	644	1,550	149	4	2,516	61.6 %
P	8	13	146	512	1	680	75.3 %
O	40	1	27	63	16	147	10.9 %
Total	2,296	1,661	1,904	766	33	6,660	
User's Accuracies	87.2%	24.9%	81.4%	66.8%	48.5%		

Sum of the major diagonal = 4,504

Overall Accuracy = 4,504/6,660 = 67.6%

H - Harvest; T - Thinning; R - Release;

P - Planting; O - Other.

Multitemporal versus single date

The error matrix for the multitemporal classification is shown by Table 7. An overall accuracy of 82.3 percent was achieved which is much better than for either single date classification. Both producer's and user's accuracies for harvest, release and planting show improvement. Producer's thinning accuracy declined slightly while user's accuracy increased. This is considered to be a very satisfactory result when MSS data is used for vegetation mapping. The KHAT value is the highest for the multitemporal classification. The Z-test statistics were more significant (19.1428 and 10.1411) as compared to the 1983 and 1979 data sets for single date classification.

Table 7. Error Matrix of Multitemporal (1979 and 1983) with PCA Classification.

Reference Data	Classified Data H	T	R	P	O	Total	Producer's Accuracies
H	2,340	197	118	34	0	2,689	87.0 %
T	81	252	104	6	0	443	56.9 %
R	133	178	1,990	137	8	2,446	81.4 %
P	4	12	51	815	12	894	91.2 %
O	39	3	24	14	8	88	9.1 %
Total	2,597	642	2,287	1,006	28	6,560	
User's Accuracies	90.1%	39.3%	87.0%	81.0%	28.6%		

Sum of the major diagonal = 5,405

Overall Accuracy = 5,405/6,560 = 82.4%

H - Harvest; T - Thinning; R - Release;

P - Planting; O - Other.

CONCLUSIONS

It is clear from the classification results that the multitemporal classification method developed in this study shows significant improvement over single date classification. All but the thinning class indicated accuracies exceeding 80 percent. Most management classes obtained from MSS data have 80 to 90 percent accuracies. From either single date or multitemporal classification the accuracy of thinning class is low. This class apparently is not associated with a distinct spectral response resulting from MSS data. Multitemporal MSS scenes provide more information over single date, especially when seasonal changes are involved.

Another advantage of multitemporal classification is that it provides information about change over time (usually defined as change detection). However, the burden of data handling, processing, and analysis increases dramatically when using more than two scenes. There is a tradeoff between using multitemporal scenes and using higher resolution imagery depending on the cost, accuracy, and management requirement. The PCA-based classification of multitemporal spectral mean classes derived from single date classification is a new

way to increase efficiency and accuracy for vegetation mapping using remotely-sensed multitemporal data.

A combination of two Landsat overpasses is the simplest combination of multitemporal data; yet, this combination is an excellent example for investigating the more complicated problem of information extraction from multilayered Geographic Information System (GIS) data. Furthermore, the logic and decisions needed to resolve the assignment of two attributes from the same object, often with the attributes in conflict or very uncertain, could be the basis for a computer-based expert system that would provide satisfactory assignments with minimal human interpretation.

SUMMARY

Advantages of satellite imaging are synoptic coverage and repeated (multitemporal) coverage of the same geographic area over time. These attributes enable us to obtain information for resource management in rapid and economic manner. This project was undertaken to investigate these two questions: (1) how can classification accuracy be improved using the relatively coarse resolution data provided by the Landsat Multispectral Scanner (MSS) systems aboard the earth resource satellites, and (2) how can we better utilize multitemporal information that is available due to repeated coverage over time. Two MSS image segments from different seasons over the study area were selected and classified. Then multitemporal analysis was conducted using Principal Components Analysis (PCA) as an aid in producing a new and more accurate vegetation map. Results from single date classification and multitemporal classification were compared. The increased information resulting from phenological change in forest sites and associated vegetation canopies decreased the uncertainty inherent in a single date classification. The use of PCA digital data from several overpasses at different dates (seasons) could be mode-

lled and an optimal combination of temporal-spectral variables could be selected.

BIBLIOGRAPHY

American Society of Photogrammetry and Remote Sensing, 1983. Manual of Remote Sensing, Second Edition. Falls Church, Va.

Bishop, Y., S. Fienberg, and P. Holland, 1975. Discrete Multivariate Analysis - Theory and Practice. MIT Press, Cambridge, Mass. 575p.

Cohen, J., 1960. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement, Vol.20, NO.1, pp.37-46.

Congalton, Russell and Richard G. Oderwald, 1983. Assessing Landsat Classification Accuracy Using Discrete Multivariate Analysis Statistical Techniques. Photogrammetric Engineering and Remote Sensing, Vol.49, No.12, December 1983, pp. 1671-1678.

Congalton, Russell, 1987. Comments on the Remote Sensing Brief entitled "Correct Formulation of the Kappa Coefficient of Agreement". Photogrammetric Engineering and Remote Sensing, Vol. 53, No.4, April 1987, pp.422.

Everett, J., and D. S. Simonett, 1976. "Principles, Concepts, and Philosophical Problems in Remote Sensing," Chapter 3 in Remote Sensing of the Environment, J. Lintz and D. S. Simonett, Eds. Reading, Mass.: Addison-Wesley. pp. 85-127.

Franklin, J. F. and G. T. Dyrness, 1973. Natural Vegetation of Oregon and Washington. USDA Forest Service Gen. Tech. Report PNW-8.

Ford, Gary E. and Claudio I. Zanelli, 1985. Analysis and Quantification of Errors in the Geometric Correction of Satellite Images. Photogrammetric Engineering and Remote sensing. Vol.51, No.11, November 1985, pp.1725-1734.

Gates, D. M., H. J. Keengan, J. C. Schlelter, and V. R. Weidner, 1965. Spectral Properties of Plants. Applied Optics, 4(1), pp.11-20.

Green, W. B., 1983. Digital Image Processing. Van Nostrand Reinhold, New York.

Heaslip, G. G., 1975. Environmental Data Handling. John Wiley & Sons, New York.

Hixson, M., D. Scholz, N. Fuhs, and T. Akiyama, 1980. Evaluation of Several Schemes for Classification of Remotely Sensed Data. Photogrammetric Engineering and Remote Sensing, Vol.46, pp.1547-1553.

Holkenbrink, P. E., 1978. Manual on Characteristics of Landsat Computer-compatible Tapes Produced by the EROS Data Center Digital Image Processing System. U.S. Department of Interior, Geologic Survey, Washingtion, D.C.

Hudson, William D. and Carl W. Ramm, 1987. Correct Formulation of the Kappa Coefficient of Agreement. Photogrammetric Engineering and Remote Sensing, Vol.53, No.4, April 1987, pp.421-422.

Jensen, John R., 1986. Introductory Digital Image Processing, A Remote Sensing Perspective. Prentice-Hall, a Division of Simon & Schuster, Englewood Cliffs, New Jersey.

Johnson, Richard A. and Dean W. Wichern, 1982. Applied Multivariate Statistical Analysis. Prentice-Hall, Englewood Cliffs, New Jersey.

Knipling, E. B., 1970. Physical and Physiological Basis for the Reflection of Visible and Near Infrared Radiation from Vegetation. Remote Sensing of Environment, 1(3). pp.155-159.

Knuth, D. E., 1973. The Art of Computer Programming. Chapter 6: Searching, Vol.3: Searching and Sorting. Addison-Welsey, Reading, Massachusetts, 722 pp.

Lillesand, T. M. and R. W. Kiefer, 1979. Remote Sensing and Image Interpretation. John Wiley and Sons, New York. 612 p.

Murray, RJay, 1981. Landsat-Based Vegetation and Landuse Inventory for Five Columbia Basin Counties in Oregon, Final Report, Environmental Remote Sensing Applications Laboratory, Oregon State University, March 1, 1981.

Murray, R. and C. J. Alexander, 1983. Practical Aspects of Producing Thematic Line Printer Maps from Landsat Data. Technical Papers, Fall Convention, American Society of Photogrammetry American Congress on Mapping and Surveying, September 19-23, 1983, Salt Lake City, pp.358-366.

Murray, RJay, 1983. Abstract Factor Analysis of Multispectral Data. Proceedings of International Conference on Renewable Resource Inventories for Monitoring Changes and Trends, School of Forestry, Oregon State University, Corvallis, Oregon. pp.648-651.

Murray, RJay, 1986. Spectral Characteristics of Oregon Estuaries. In: Fourteenth Year Projects and Activities: A Progress Report to the National Aeronautics and Space Administration (NASA Grant NGL 38-002-053), Environmental Remote Sensing Applications Laboratory, Oregon State University, Corvallis, OR. pp.215-236.

Noy-Meir, I., 1973. Data Transformation in Ecological Ordinations. I. Some Advantages of Non-centering. J. Ecol., Vol.61, pp.329-341.

Paine, David P., 1981. Aerial Photography and Image Interpretation for Resource Management. John Wiley and Sons, New York. 571 p.

Strahler, A. H., 1980. The Use of Prior Probabilities in maximum likelihood classification of remotely sensed data. Remote Sensing Environment, Vol.10, pp.135-163.

Story, Michael and Russell G. Congalton, 1986. Accuracy Assessment: A User's Perspective. Photogrammetric Engineering and Remote Sensing, Vol.52, No.3, March 1986, pp.379-399.

Swain, P. H. and S. M. Davis, 1978. Remote Sensing: The Quantitative Approach. McGraw-Hill Book Company, New York.

Townshend, J. R. G., 1981. Terrain Analysis and Remote Sensing, J. R. G. Townshend, Ed. Geoge Allen & Unwin, London. pp.59-108.

Waring, R. H., J. D. Aber, J. M. Melillo, and B. Moore III, 1986. Precursors of Change in Terrestrial Ecosystems. Bioscience, Vol.36, No.7, July/August 1986, pp.433-438.

Wolf, Paul R., 1974. Elements of Photogrammetry. McGraw-Hill Book Company, New York.

APPENDICES

APPENDIX A

SPECTRAL MEAN CLASSES AND FINAL SYMBOL ASSIGNMENTS FOR 1979 DATA SET

S ----- Symbol on the plot,
#PIX ----- Number of pixels,
M4-M7 ----- Means of band 5 to band 7,
ST4-ST7----- Standard deviation of band 5 to band 7.

S	#PIX	M4	M5	M6	M7	ST4	ST5	ST6	ST7
C	20997	15.61	10.95	32.82	27.68	.92	.86	.82	.86
A	20883	16.39	11.50	38.10	35.03	.97	.83	.80	.86
B	19368	14.70	10.07	28.25	21.46	.84	.81	.72	.82
E	17067	15.05	10.30	31.36	25.90	.88	.81	.80	.81
F	17067	16.67	11.74	39.68	37.77	.91	.81	.86	.92
H	16661	15.71	10.96	36.30	32.26	.86	.80	.90	.95
J	17026	14.71	10.00	29.70	23.58	.78	.78	.74	.84
D	15469	15.35	10.66	34.48	29.90	.84	.82	.86	.91
I	14971	14.03	9.12	24.39	17.17	.61	.72	.71	.79
R	12831	16.73	12.18	34.62	29.75	.81	.80	.89	.97
O	12965	16.96	12.28	36.83	32.37	.80	.72	.90	.93
K	11636	16.91	12.03	40.99	40.69	.90	.83	.85	.90
G	11016	14.45	10.23	25.94	19.25	.75	.50	.84	.90
Q	11617	15.87	11.55	30.48	24.18	.88	.78	.89	.95
V	9059	17.67	13.94	39.41	36.80	.67	.83	.87	.83
M	8835	13.83	8.86	22.76	15.07	.65	.88	.83	.82
1	8548	14.19	8.78	26.24	19.46	.69	.50	.93	.93
T	8478	17.59	13.96	38.03	34.17	.67	.82	.79	.78
L	7748	17.26	12.36	42.51	43.46	.83	.82	.92	.96
U	7558	17.87	14.14	40.89	39.48	.63	.86	.85	.89
S	7494	17.15	13.71	32.80	26.78	.87	1.10	.88	.89
X	6776	17.85	14.88	36.69	31.64	.73	.70	.98	.87
≤	6168	17.85	14.97	34.83	28.98	.72	.72	.96	.96
"	4516	18.16	14.77	42.60	42.52	.66	.75	.98	1.05
-	3995	13.55	8.37	20.26	12.71	.73	.96	1.11	1.10
8	3720	17.33	12.33	43.94	46.42	.82	.76	.83	.96
?	3600	18.61	16.24	39.17	35.30	.66	.80	1.02	1.20
4	3418	17.39	14.47	31.12	24.23	.83	.73	.87	.92
Z	3033	18.80	18.25	35.81	29.82	.72	.89	.86	.80
N	3531	15.57	11.72	27.17	19.63	.91	.86	1.00	.96
\	3371	14.88	11.00	24.28	16.60	.86	.87	.88	1.00
6	2955	16.51	13.09	29.17	21.88	.93	1.05	.82	.84
Y	2878	18.80	18.01	37.50	31.91	.71	.86	.77	.86
(	2600	18.12	14.65	44.13	46.02	.68	.70	.96	1.04

S	#PIX	M4	M5	M6	M7	ST4	ST5	ST6	ST7
,	2381	19.30	17.81	41.57	39.56	.90	1.24	.96	1.11
2	2718	18.68	18.12	34.50	27.74	.72	.91	.88	.79
9	2347	20.30	21.22	35.51	28.70	.98	1.12	.89	.89
0	2215	18.87	19.08	31.60	23.93	.76	.71	.84	.88
[	2214	19.88	20.60	33.79	26.40	.95	.81	.96	.98
]	2166	20.28	20.69	37.40	31.31	.97	.89	.80	.97
=	1968	20.14	21.40	31.36	23.26	.93	.83	.95	1.07
5	2123	18.66	18.18	32.92	26.08	.69	.89	.76	.78
/	2177	19.34	18.98	38.65	33.95	.81	.76	.88	.85
+	2123	17.70	13.14	45.27	49.35	.73	1.06	.88	.96
	2202	18.19	16.47	31.88	24.68	.63	.67	1.00	1.05
<	2141	21.77	23.85	33.21	24.85	1.19	1.14	1.34	1.66
;	2075	22.14	24.08	37.05	29.99	.99	1.20	1.10	1.74
'	1980	20.44	22.30	29.19	20.06	1.28	1.36	1.01	1.15
7	1995	19.75	19.11	40.24	36.71	.98	.87	.80	.93
3	2120	18.75	19.17	29.89	21.74	.73	.86	.79	.85
!	1844	18.06	16.15	29.51	21.56	.62	.86	.88	.97
)	1382	21.03	21.49	39.01	34.05	.75	1.05	.88	1.04
$	1509	17.32	14.83	27.54	19.25	.86	.73	1.05	1.00
*	1508	18.60	19.23	28.17	19.37	.70	.83	.84	.92
P	1495	17.16	14.82	24.03	15.07	.98	.98	1.41	1.87
	1240	18.44	18.79	24.20	14.99	.79	1.11	1.20	1.69
	1058	18.14	17.52	26.66	17.98	.67	1.00	1.00	.97
>	975	19.52	21.44	26.24	17.21	1.00	1.17	1.09	1.17
+	917	19.96	18.48	43.60	42.97	1.11	1.22	.91	1.25
.	938	22.41	22.99	40.72	36.36	1.26	1.93	1.08	1.98
+	520	18.19	14.03	46.66	52.08	.64	.97	.82	1.15
+	316	19.75	17.73	45.45	47.25	1.05	1.21	1.04	1.27
+	236	17.61	12.35	46.96	53.81	.73	.84	1.01	1.62
W	556	16.58	11.26	7.31	1.47	1.00	.81	2.09	.81
W	508	17.20	14.12	13.99	5.16	1.01	1.04	1.91	1.78
W	499	18.54	17.93	18.61	8.39	.95	1.30	1.52	1.71
W	109	17.21	13.23	8.31	1.72	.78	.50	1.03	.99
W	488	13.68	8.07	13.68	6.80	1.29	1.47	4.70	3.64

MANAGEMENT CLASSES	SYMBOLS
HARVEST	- M I \ 1 G N B 6 J Q E
THINNING	C D R H O
RELEASE	4 S 5 2 ≤ Z X Y / T ? 7 A V F U
	, K " + L 8
PLANTING	P > ' * $! 3 = 0 < [9] ;) .
OTHER	W

APPENDIX B

SPECTRAL MEAN CLASSES AND FINAL SYMBOL ASSIGNMENTS FOR 1983 DATA SET

S ----- Symbol in the plot,
#PIX ----- Number of pixels,
M4-M7 ----- Means of band 5 to band 7,
ST4-ST7----- Standard deviation of band 5 to band 7.

S	#PIX	M4	M5	M6	M7	ST4	ST5	ST6	ST7
A	11810	18.66	11.42	45.91	32.75	.84	.87	.91	.85
J	14305	17.93	11.77	47.51	34.78	.84	1.12	.87	.88
E	15274	16.65	10.32	37.86	24.06	1.03	1.11	1.07	1.01
H	15031	19.59	12.40	44.77	31.15	.89	1.60	.86	.96
C	15492	17.33	10.86	43.25	30.01	.71	.89	.89	.71
L	13531	20.98	13.02	57.50	47.72	1.37	1.29	.95	1.36
G	13706	19.89	11.79	50.34	38.69	.52	.87	.92	.88
B	11461	17.27	11.29	39.74	26.43	.79	.79	.64	.79
F	10993	20.19	11.90	53.21	41.24	.70	.87	.83	1.00
D	19222	19.93	12.02	48.68	36.12	.58	1.29	.88	.89
K	11422	15.93	9.46	32.63	20.38	1.11	1.31	1.69	1.05
N	11520	20.14	11.91	55.05	43.90	1.20	.82	1.00	1.05
I	11278	21.45	14.34	55.20	43.84	1.47	.78	1.10	1.25
S	11099	22.87	20.03	43.46	28.26	1.80	2.98	1.35	1.67
R	9393	17.78	11.95	49.36	37.49	.74	1.09	.90	.91
O	8666	24.27	21.62	47.25	32.38	1.54	2.26	1.18	1.33
P	8091	22.89	17.51	54.27	42.05	1.49	1.04	1.33	1.61
8	10498	17.08	11.16	44.86	32.28	.77	1.08	.88	.96
2	9260	20.51	14.26	52.53	41.04	1.08	.66	.99	1.08
M	7143	22.43	14.85	50.87	38.18	.99	.77	1.33	1.41
6	10274	19.84	12.60	47.24	33.33	.65	1.03	.87	.86
X	6575	15.72	10.23	35.23	22.61	.93	1.06	.85	.91
Q	8055	21.05	16.32	47.68	34.01	1.32	1.36	1.07	1.42
0	7185	17.48	12.68	42.64	28.98	.69	.77	.94	1.08
T	6574	19.48	11.12	43.14	29.53	.54	.90	.79	.72
1	7432	19.68	12.45	41.56	27.25	.77	1.78	1.18	1.02
)	6741	24.88	22.14	50.58	36.89	1.66	1.97	1.29	1.47
+	8206	17.39	10.81	41.85	27.15	.59	.69	.73	.82
U	7051	18.25	11.98	51.75	40.52	.88	.95	1.14	1.14
Z	5348	24.73	21.12	54.40	41.55	1.59	1.48	1.41	1.77
Y	5743	21.90	12.10	51.63	39.14	.92	.86	1.27	1.64
.	5729	17.37	8.74	40.19	26.65	.91	.73	1.12	1.34
3	7640	20.06	15.28	50.20	38.02	.72	1.61	1.01	1.08
V	4025	23.28	18.20	50.79	38.06	.94	.75	1.06	1.11

S	#PIX	M4	M5	M6	M7	ST4	ST5	ST6	ST7
;	5562	15.54	11.14	41.00	27.70	.80	.82	1.15	1.23
[	4218	23.72	17.70	57.52	46.73	1.45	1.77	.88	1.22
]	3695	15.31	9.16	27.95	17.84	1.24	1.29	1.53	1.18
7	3683	17.60	10.74	34.62	21.82	.72	1.03	.94	.94
W	16081	18.09	9.39	4.83	2.11	1.75	1.91	1.68	.74
5	3337	21.71	12.80	47.97	34.60	.85	1.13	.94	.87
4	2632	23.57	18.58	48.94	34.88	.99	.94	.81	.88
$	2957	22.09	13.88	59.47	51.68	1.28	1.03	.65	1.08
*	6411	19.06	13.27	38.42	23.83	1.47	2.14	1.60	1.53
-	3148	26.88	27.63	44.91	28.79	1.37	2.10	1.71	2.06
/	2958	23.82	22.85	39.47	23.40	1.80	2.75	1.62	1.68
9	2249	22.28	15.08	45.87	31.35	.88	1.05	1.04	.98
,	1537	23.90	17.30	59.33	50.74	1.49	1.61	.77	1.36
<	1914	20.95	12.60	60.14	54.03	1.35	1.27	.68	2.14
=	1648	29.24	31.02	47.56	31.57	1.45	2.70	1.82	2.56

MANAGEMENT CLASSES	SYMBOLS
HARVEST	] K 7 * X E B 1 + ; . 0
THINNING	T C 8 A J D R G
RELEASE	H 6 5 3 M P Y 2 [I F U N L , $ <
PLANTING	/ - = S O) 9 4 Q V Z
OTHER	W

APPENDIX C

A SAMPLE PAGE OF THE SORTED RESULT ON 255 MULTITEMPORAL SYMBOLS

NUMBER OF CHARACTER COMBINATIONS = 255
TOTAL OCCURRENCES(=NUMBER OF PIXELS) 247021.
SYMBOLS AND OCCURRENCES FOR SCENE 1

	A	C	E	G	I	O	S	V	W	4	8	@
NUM(N)	8241	10681	7973	25179	14966	7465	3989	31929	1218	17091	15547	18400

	$	=	%
NUM(N)	74317	3533	5776	714

SYMBOLS AND OCCURRENCES FOR SCENE 2

	G	H	I	N	O	V	W	Y	6	8	*	*
NUM(N)	5754	5459	13165	17679	11586	18469	796	1688	4831	61563	29685	2090

	/	$	=		%	÷
NUM(N)	1212	57288	2575	5785	1712	5684

CHARACTER: * COMBINATIONS: 16 TOTAL OCCURRENCES: 18400
% TOTAL PIXELS: 7.449

	*	N	8	I	V	O	$	G	=	6	H	Y
NUM(N)	5668	3964	3254	1031	1008	847	809	737	338	171	154	149
N/OCC	.308	.215	.177	.056	.055	.046	.044	.040	.018	.009	.008	.008
ACC N/O	.308	.523	.700	.756	.811	.857	.901	.941	.960	.969	.977	.985
N/TOT	.023	.016	.013	.004	.004	.003	.003	.003	.001	.001	.001	.001
ACC N/T	.023	.039	.052	.056	.060	.064	.067	.070	.071	.072	.073	.073
N/SC1	.308	.000	.209	.069	.032	.113	.011	.029	.096	.000	.000	.000
N/SC2	.191	.224	.053	.079	.055	.073	.014	.126	.131	.035	.028	.068

	:	%	/	@
NUM(N)	96	74	60	40
N/OCC	.005	.004	.003	.002
ACC N/O	.991	.995	.998	1.000
N/TOT	.000	.000	.000	.000
ACC N/T	.074	.074	.074	.074
N/SC1	.000	.104	.000	.000
N/SC2	.017	.043	.050	.019

CHARACTER: $ COMBINATIONS: 18 TOTAL OCCURRENCES: 74317
% TOTAL PIXELS: 30.085

	$	8	:	+	V	O	N	H	I	*	6	G
NUM(N)	41790	16033	2603	2152	1731	1722	1691	1428	1226	1110	859	627
N/OCC	.562	.216	.035	.029	.023	.023	.023	.019	.016	.015	.012	.008
ACC N/O	.562	.778	.813	.842	.865	.889	.911	.930	.947	.962	.973	.982
N/TOT	.169	.065	.011	.009	.007	.007	.007	.006	.005	.004	.003	.003
ACC N/T	.169	.234	.245	.253	.260	.267	.274	.280	.285	.289	.293	.295
N/SC1	.562	1.031	.000	.117	.054	.231	.000	.000	.082	.000	.000	.025
N/SC2	.729	.260	.456	.072	.094	.149	.096	.262	.093	.531	.178	.109

	=	/	%	Y	W	
NUM(N)	547	345	268	160	19	6
N/OCC	.007	.005	.004	.002	.000	.000
ACC N/O	.989	.994	.998	1.000	1.000	1.000
N/TOT	.002	.001	.001	.001	.000	.000
ACC N/T	.298	.299	.300	.301	.301	.301
N/SC1	.155	.000	.375	.000	.016	.001
N/SC2	.212	.285	.157	.095	.024	.001

APPENDIX D

COMBINED SPECTRAL MEAN CLASSES AND FINAL SYMBOL ASSIGNMENTS FOR 61 MULTITEMPORAL COMBINATIONS (1979-1983)

SP ----- 1979-1983 symbol pairs,
SS ----- Single symbols on the PCA plot,
M1-M4 ----- Means of band 5 to band 7 for 1979 data,
M5-M8 ----- Means of band 5 to band 7 for 1983 data.

SP	SS	M1	M2	M3	M4	M5	M6	M7	M8
$$	$	14.64	10.00	26.87	20.03	16.58	10.28	37.29	24.13
$8	?	14.64	10.00	26.87	20.03	18.36	11.51	46.35	33.55
G8	G	16.52	11.61	38.81	36.26	18.36	11.51	46.35	33.55
48	U	16.26	11.54	36.53	32.31	18.36	11.51	46.35	33.55
V+	L	18.21	16.12	35.57	30.08	21.90	14.84	56.62	46.22
88	8	15.98	11.35	34.54	29.83	18.36	11.51	46.35	33.55
++	+	17.29	12.58	42.46	43.42	21.90	14.84	56.62	46.22
VV	M	18.21	16.12	35.57	30.08	23.25	18.92	50.00	36.50
V8	,	18.21	16.12	35.57	30.08	18.36	11.51	46.35	33.55
C8	9	15.61	10.95	32.82	27.68	18.36	11.51	46.35	33.55
GN	N	16.52	11.61	38.81	36.26	19.71	11.92	53.58	42.10
I8	"	17.70	14.01	39.39	36.72	18.36	11.51	46.35	33.55
+8	Z	17.29	12.58	42.46	43.42	18.36	11.51	46.35	33.55
E8	\	15.05	10.30	31.36	25.90	18.36	11.51	46.35	33.55
A+	3	20.98	22.06	36.13	29.42	21.90	14.84	56.62	46.22
E$	F	15.05	10.30	31.36	25.90	16.58	10.28	37.29	24.13
V$	<	18.21	16.12	35.57	30.08	16.58	10.28	37.29	24.13
AV	A	20.98	22.06	36.13	29.42	23.25	18.92	50.00	36.50
C$	Q	15.61	10.95	32.82	27.68	16.58	10.28	37.29	24.13
8$	E	15.98	11.35	34.54	29.83	16.58	10.28	37.29	24.13
VO	.	18.21	16.12	35.57	30.08	23.48	20.73	45.12	30.07
$+	(	14.64	10.00	26.87	20.03	21.90	14.84	56.62	46.22
OV	V	18.64	18.39	29.60	21.43	23.25	18.92	50.00	36.50
$V	!	14.64	10.00	26.87	20.03	23.25	18.92	50.00	36.50
$O	≤	14.64	10.00	26.87	20.03	23.48	20.73	45.12	30.07
$N	)	14.64	10.00	26.87	20.03	19.71	11.92	53.58	42.10
G$	H	16.52	11.61	38.81	36.26	16.58	10.28	37.29	24.13
4N	I	16.26	11.54	36.53	32.31	19.71	11.92	53.58	42.10
4$	P	16.26	11.54	36.53	32.31	16.58	10.28	37.29	24.13
$H	-	14.64	10.00	26.87	20.03	19.59	12.40	44.77	31.15
GG	J	16.52	11.61	38.81	36.26	19.89	11.79	50.34	38.69

(CONTINUATION OF APPENDIX D)

SP	SS	M1	M2	M3	M4	M5	M6	M7	M8
OO	O	18.64	18.39	29.60	21.43	23.48	20.73	45.12	30.07
IV	6	17.70	14.01	39.39	36.72	23.25	18.92	50.00	36.50
$I	]	14.64	10.00	26.87	20.03	20.94	14.76	51.30	39.23
O+	1	18.64	18.39	29.60	21.43	21.90	14.84	56.62	46.22
GV	K	16.52	11.61	38.81	36.26	23.25	18.92	50.00	36.50
4+	[	16.26	11.54	36.53	32.31	21.90	14.84	56.62	46.22
=V	0	18.98	19.42	26.24	17.16	23.25	18.92	50.00	36.50
GO	;	16.52	11.61	38.81	36.26	23.48	20.73	45.12	30.07
+V	X	17.29	12.58	42.46	43.42	23.25	18.92	50.00	36.50
8N	R	15.98	11.35	34.54	29.83	19.71	11.92	53.58	42.10
=O	=	18.98	19.42	26.24	17.16	23.48	20.73	45.12	30.07
I$	≥	17.70	14.01	39.39	36.72	16.58	10.28	37.29	24.13
+O	C	17.29	12.58	42.46	43.42	23.48	20.73	45.12	30.07
4O	S	16.26	11.54	36.53	32.31	23.48	20.73	45.12	30.07
VH	>	18.21	16.12	35.57	30.08	19.59	12.40	44.77	31.15
+$	B	17.29	12.58	42.46	43.42	16.58	10.28	37.29	24.13
4G	D	16.26	11.54	36.53	32.31	19.89	11.79	50.34	38.69
8O	T	15.98	11.35	34.54	29.83	23.48	20.73	45.12	30.07
AI	4	20.98	22.06	36.13	29.42	20.94	14.76	51.30	39.23
8H	Y	15.98	11.35	34.54	29.83	19.59	12.40	44.77	31.15
$G	*	14.64	10.00	26.87	20.03	19.89	11.79	50.34	38.69
A8	5	20.98	22.06	36.13	29.42	18.36	11.51	46.35	33.55
S6	↓	18.15	14.73	43.16	43.80	18.36	11.51	46.35	33.55
IO	←	17.70	14.01	39.39	36.72	23.48	20.73	45.12	30.07
V:	/	18.21	16.12	35.57	30.08	18.60	12.56	42.09	28.10
$=	'	14.64	10.00	26.87	20.03	27.69	28.79	45.82	29.75
CN	W	15.61	10.95	32.82	27.68	19.71	11.92	53.58	42.10
O$	2	18.64	18.39	29.60	21.43	16.58	10.28	37.29	24.13
=+	7	18.98	19.42	26.24	17.16	21.90	14.84	56.62	46.22
4Y	↑	16.26	11.54	36.53	32.31	21.90	12.10	51.63	39.14

MANAGEMENT CLASSES	SINGLE SYMBOLS
HARVEST	$? F < Q E ( ) H P -] ≥ B * /
THINNING	U 8 9 \ D 5
RELEASE	G L + , N " Z I J [R > Y ↓ W ↑
PLANTING	M 3 A . V ! ≤ O 6 1 K 0 ; X = C S T 4 ← ' 2 7

APPENDIX E

PCA OUTPUT OF 61 MULTITEMPORAL COMBINATIONS (1979-1983 SPECTRAL MEAN CLASSES)

```
 FXCORL   VERSION 3.0   2 MARCH 1982     87/05/28.   12.50.33.

 ENTER  # OF COLUMNS,  # OF ROWS
? 8 61
 ENTER INPUT FILE FORMAT
 FREE FORMAT = 0
 READ BINARY = 1
? 0
 INPUT DATA FROM INMAT(LUN 4) -- FREE FORM
 ENTER NUMBER PLACES TO SHIFT INPUT VALUES LEFT
<CR> DEFAULT = 0
?
 ENTER FILE TYPE
 RM = CORRELATION ABOUT MEAN
 RO = CORRELATION ABOUT ORIGIN
 CM = COVARIANCE ABOUT MEAN
 CO = COVARIANCE ABOUT ORIGIN
? CO
 ENTER ID RECORD--MAX 2 LINES OF 60 CHARS
 TERMINATE WITH EOF
? SEN WANG 61 CLASS SELECTION 1ST
? 61 SPECTRAL CLASS COMBINATIONS X 8 COLS

 CORRELATION REDUCTIONS VIA COOLEY-LOHNES  8 COLUMNS    61 ROWS
 FILE ID -
 [[[[ID]]COSEN WANG 61 CLASS SELECTION 1ST
 61 SPECTRAL CLASS COMBINATIONS X 8 COLS

 DO YOU WANT TO PRINT C OR R MATRIX? -- <CR> = NO
? YES

 COVARIANCE (OR CORRELATION) MATRIX

 SECTION  1

 ROW            1           2           3           4           5           6           7           8
     1     17721.988
     2     14387.773   12000.062
     3     35689.154   28447.541   73493.433
     4     30689.809   24275.015   64112.638   56552.121
     5     21158.263   16991.364   42655.433   36598.742   25850.348
     6     15086.658   12170.855   30269.372   25919.375   18724.662   13944.439
     7     48920.165   39241.897   98780.804   84814.648   59304.133   42195.754  138069.654
     8     35552.572   28559.669   71701.162   61523.270   43140.259   30541.970  101130.185   74502.42
```

```
/BEIGEN,VMATRX,OUTRAN
 EIGENVALUES AND VECTORS
 22 SEPTEMBER 1981 VERSION 1.0      87/05/28.   12.53.13.


 *** FILE ID ***
 [[[[ID]]]COSEN WANG 61 CLASS SELECTION 1ST
 61 SPECTRAL CLASS COMBINATIONS X 8 COLS


EIGENVALUES, SMALLEST TO LARGEST

        1          2          3          4          5          6          7          8
    1.4251     5.0492     7.8977     174.27     936.99     1668.2     5358.3     .40398E+06

 EIGENVECTORS AS COLUMNS CORRESPONDING TO EIGENVALUES ABOVE.

          1          2          3          4          5          6          7          8
  1    .82387    -.11288     .11071     .25120     .38214    -.20788   -.23856E-01-.20812
  2   -.36284     .52518E-01 .80015E-01-.38181     .75660    -.33293     .47941E-01-.16681
  3   -.17948     .32288E-01-.66731     .35333     .86805E-01-.21521E-01-.45775    -.42316
  4    .68390E-01-.11132E-01 .44246    -.38893    -.17420     .16609    -.67569    -.36562
  5   -.26698E-01 .80237     .20923     .18708    -.22841    -.38551     .11341    -.25112
  6    .70006E-01-.38050    -.21607    -.34218    -.43303    -.67425     .11150    -.17885
  7   -.29330    -.38174     .36776     .37543    -.58447E-01 .12889     .36353    -.58299
  8    .24709     .22142    -.33846    -.47115    -.23261E-01 .44503     .41665    -.42568

 ENTER NUMBER OF COLUMNS AND ROWS IN ORIGINAL DATA MATRIX
? 8 61


                SUM        RMS        REAL       IMBED                            PERCENT
   N   (C-N)   EV(2)       ERR.       ERR.       ERR.       MALINOWSKI    VAR.       RATIO
   -   ----   -----      ----       ----       ----       ----------  ---------   -------
   1     7    8152.1     4.0872     4.3694     1.5448     .89171E-01    98.02      75.39
   2     6    2793.9     2.3927     2.7629     1.3814     .76747E-01    99.32       3.21
   3     5    1125.6     1.5188     1.9211     1.1764     .76844E-01    99.73       1.78
   4     4    188.65     .62175     .87928     .62175     .54955E-01    99.95       5.38
   5     3    14.372     .17161     .28024     .22155     .31138E-01•   100.00     22.07
   6     2    6.4743     .11518     .23037     .19950     .57591E-01    100.00      1.56
   7     1    1.4251     .54039E-01 .15285     .14297     .15285        100.00      3.54

 AVERAGE EIGENVALUE =  51517.

 RMS ERR. ... EST RMS OF (RECONST. DATA - ORIGINAL DATA)
 RE (REAL ERR.) ... EST. OF (PURE DATA - ORIGINAL DATA)
 IE (IMBED. ERR.) ... EST. OF (PURE DATA - RECONST. DATA)

 ENTER NUMBER OF FACTORS
/ 5


      ABSTRACT COLUMN MATRIX [C] = TR([Q]) :
        -.208    -.167    -.423    -.366    -.251    -.179    -.583    -.426
        -.024     .048    -.458    -.676     .113     .111     .364     .417
        -.206    -.333    -.022     .160    -.386    -.674     .129     .445
         .382     .757     .087    -.174    -.228    -.433    -.058    -.023
         .251    -.382     .353    -.389     .187    -.342     .375    -.471
```

[[[[ID]]COSEN WANG 61 CLASS SELECTION 1ST
61 SPECTRAL CLASS COMBINATIONS X 8 COLS

1	-61.422	.93274	-1.4025	1.0245	3.7784
2	-71.381	8.4902	2.4419	-.66337	2.6536
3	-83.027	-7.9094	3.9537	-.51759	.41767
4	-80.552	-4.1940	3.4241	-.17974	1.1097
5	-93.365	7.7105	4.4964	1.1903	-2.2120
6	-78.713	-1.6098	3.1765	-.17123	1.3734
7	-100.38	-4.6049	7.9336	-3.5654	-3.8454
8	-86.437	1.8621	-3.9539	-.27183	-1.2613
9	-80.500	-2.0747	1.1442	4.3358	.37902
10	-77.056	.61995	3.0665	-.39004	1.6616
11	-91.294	-1.5200	7.8937	-1.6249	-.78407
12	-84.086	-8.3988	2.9733	1.7194	-.17622
13	-87.511	-14.390	4.5814	-.41985	-1.2543
14	-75.562	2.4732	3.1351	-.91249	1.9455
15	-94.928	8.1188	1.8213	6.9066	-3.3296
16	-65.603	-5.0842	-.70926	.77535	3.0703
17	-70.542	-9.6321	-2.7002	6.0236	1.5038
18	-88.000	2.2704	-6.6290	5.4445	-2.3788
19	-67.097	-6.9375	-.77785	1.2978	2.7863
20	-68.755	-9.1672	-.66786	1.5166	2.4981
21	-81.237	-2.3631	-8.7534	-.67336	-.64018
22	-84.246	18.275	5.7941	-3.8089	.62570E-01
23	-81.216	10.538	-6.1072	2.5986	-.76515
24	-77.317	12.427	-2.6562	-5.2710	1.0133
25	-72.117	8.2018	-7.4557	-5.6725	1.6344
26	-79.647	14.880	6.3818	-1.7707	1.4519
27	-73.068	-15.467	.10939	1.1703	1.5424
28	-88.819	2.1955	7.3641	-1.2871	-.91995E-01
29	-70.593	-11.751	-.42025	1.5081	2.2345
30	-69.906	7.1546	.95913E-01	-1.1815	3.1168
31	-87.975	-4.1126	5.9768	-1.3411	-.31567
32	-76.016	6.3129	-10.907	2.1970	-.14404
33	-90.023	-4.4621	-2.1247	-2.8882	-1.8165
34	-77.913	13.311	2.4217	-3.0814	1.2064
35	-88.145	16.386	2.3430	4.0607	-1.7159
36	-88.964	-3.9727	-1.1443	-5.1252	-1.2226
37	-93.417	5.5912	6.7763	-3.3253	-1.4813
38	-78.476	15.003	-7.1577	3.9599	-.59948
39	-83.763	-8.1979	-5.9439	-5.5267	-.60153
40	-93.448	-10.453	-.51667	-5.0275	-2.8946
41	-86.980	4.7797	7.1165	-1.2786	.17162
42	-73.275	10.777	-11.957	3.5584	.21624E-01
43	-74.128	-15.956	-.87103	3.4072	.94853
44	-88.248	-14.678	-5.3162	-5.4290	-2.2735
45	-81.288	-4.4824	-6.4735	-5.1889	.90548E-01
46	-79.026	-3.4103	-1.2018	3.8176	.84218
47	-77.553	-21.947	.73703	1.2680	-.12956
48	-85.500	-.39716	5.4472	-1.0032	.37640
49	-79.450	-1.8982	-6.7211	-5.1804	.35416
50	-88.596	3.1546	-1.5511	7.6340	-2.1857
51	-77.239	-2.9454	.83052	-.68939	1.8365
52	-76.329	12.287	4.4650	-1.4868	1.9203
53	-82.063	-1.6664	-1.5310	10.052	-.73853
54	-88.484	-14.885	3.7349	1.5300	-1.7596
55	-84.823	-8.6873	-6.9243	-3.2898	-1.1954
56	-75.945	-5.7498	-2.6308	4.2021	1.0331
57	-74.887	9.6990	-14.565	-10.158	.77632E-01
58	-85.323	7.0094	7.0065	-1.4974	.45986
59	-65.321	-.95607	-4.8535	8.8940	1.9999
60	-85.404	20.851	1.2925	5.4220	-1.5502
61	-87.004	.52179	4.8298	-1.6824	.91864

APPENDIX F

OUTPUT OF CONTINGENCY TABLE ANALYSIS FOR 1979, 1983 SINGLE AND MULTITEMPORAL CLASSIFICATION

PCA (79-83)
****

THE ORIGINAL ERROR MATRIX IS:

2340.	197.	118.	34.
81.	252.	104.	6.
133.	178.	1990.	137.
4.	12.	51.	815.

LOWER LIMIT	KHAT	UPPER LIMIT
.74448	.75754	.77061

TH1	TH2	TH3	TH4	VARIANCE
.836485	.325588	.562909	.466048	.00004444

THE Z STATISTIC IS: 113.63500

DATE79
****

THE ORIGINAL ERROR MATRIX IS:

2211.	387.	103.	12.
94.	309.	103.	1.
165.	394.	2030.	46.
19.	20.	156.	458.

LOWER LIMIT	KHAT	UPPER LIMIT
.64166	.65624	.67083

TH1	TH2	TH3	TH4	VARIANCE
.769514	.329507	.536948	.477038	.00005534

THE Z STATISTIC IS: 88.21726

THE ORIGINAL ERROR MATRIX IS:

2013.	590.	106.	33.
66.	413.	75.	9.
169.	644.	1550.	149.
8.	13.	146.	512.

LOWER LIMIT	KHAT	UPPER LIMIT
.54857	.56353	.57849

TH1	TH2	TH3	TH4	VARIANCE
.690887	.291788	.438163	.376538	.00005828

THE Z STATISTIC IS: 73.81763

SUMMARY TABLE AND COMPARISONS

MATRIX	LOWER LIMIT	KHAT	UPPER LIMIT
1	.74448	.75754	.77061
2	.64166	.65624	.67083
3	.54857	.56353	.57849

COMBINATION	TEST STATISTIC
1 2	10.1411
1 3	19.1428
2 3	8.6982