This is to certify that the

thesis entitled

DISCRIMINATION OF FOREST TYPES THROUGH DENSITOMETRIC
ANALYSIS OF MULTILEVEL IMAGERY IN
PARANA STATE, BRAZIL

presented by

Flavio Felipe Kirchner

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Forestry

Carl W. Ramm

Major professor

Date April 25, 1980

DISCRIMINATION OF FOREST TYPES THROUGH DENSITOMETRIC ANALYSIS OF MULTILEVEL IMAGERY IN PARANA STATE, BRAZIL

By

Flavio Felipe Kirchner

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Forestry

1980

ABSTRACT

DISCRIMINATION OF FOREST TYPES THROUGH DENSITOMETRIC ANALYSIS OF MULTILEVEL IMAGERY IN PARANA STATE, BRAZIL

By

Flavio Felipe Kirchner

This study investigated the technological potential for discriminating four forest types using spectral densities measured on large-scale panchromatic and satellite imagery in Parana State, Brazil. Several levels of imagery were used, aerial photos in 1:40,000 scale, SKYLAB imagery, LANDSAT imagery, and ratio forms of LANDSAT channels. Discriminant analyses of densitometric measurements, supported by other multivariate techniques, were used to distinguish between the forest types at all imagery levels. Tests of the modeling process indicated that, from the fifteen imagery scales tested, the color composite gave the best results. Four variables were tested: (1) the arithmetic mean of grey density variation; (2) the mean quadratic deviation of grey density variations from the arithmetic mean of grey density; (3) the moment coefficient of skewness of grey density variations from the arithmetic mean of grey density; and (4) the moment coefficient of kurtosis of grey density variations from the arithmetic mean of grey density. Of the four, only the arithmetic mean and the mean quadratic deviation were used in the model. The two other variables, when used, accounted only for misclassification between the four forest types. Discriminant analyses procedures indicated that 100 percent correct classification for the four forest types was achieved using two discriminant functions.

This work is dedicated to my wife,
parents, grandfather, and in memory
of my grandmother.

ACKNOWLEDGMENTS

Through this opportunity, I would like to express my appreciation and profound gratitude to my major professor and chairman of my dissertation committee, Dr. Carl W. Ramm, for his assistance and patient guidance in the preparation of this manuscript and study at Michigan State University.

I am also grateful to Dr. Victor J. Rudolph, Dr. Dieter H. Brunnschweiler, Dr. Robert I. Wittick, and Dr. Ronald L. Shelton for serving on my guidance committee and assistance throughout my doctoral program.

Appreciation is extended to Dr. Wayne L. Myers for his invaluable help and encouragement during my study at Michigan State University.

I am also indebted to Federal University of Parana for providing me this opportunity. I am also indebted to the Brazilian government for the financial support to me during my study at Michigan State University.

I am also grateful to all persons who guided, encouraged, and supported me during the period of study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

Problem Statement

Forest resource planners require several types of information to achieve optimal decisions. Such information is provided through resource surveys, which quantify the extent of the resource base, current condition, and changes over time in terms of growth and loss. Such data may be obtained through field surveys, from aerial photos, from aerial photos supported by field surveys, or from satellite-based sensors. Selection of a data-gathering system for resource inventories, which utilize aerial photography or satellite photography, depends on the importance of the information, quality of photography obtainable, available analytical technology and cost considerations. In an extensive area such as Parana State, Brazil, field surveys are too costly and time consuming. The alternative survey methods, utilizing either aerial photos, or LANDSAT multispectral scanner (MSS) imagery, or some combination of the two, are evaluated in this study.

To inventory forest canopies by their reflectance properties, either on aerial photographs or satellite imagery, requires knowledge of spectral and spatial (e.g., horizontal and vertical structure) characteristics and their respective phenologies, growth characteristics, dependence on site conditions, as well as an understanding of how these characteristics affect remote sensors signals. These various phenomena can be investigated through direct observation, experimentation and measurement, or mathematical modeling and simulation. Mathematical modeling seems to have the greatest potential. Models that describe forest canopy reflectance incorporate the interactions of various characteristics with incoming radiation; this enables the calculation of the resultant signal for a particular remote sensor. The study of canopy reflectance offers the potential for identifying and assessing different forest canopy situations.

Forest canopy reflectance may be examined by either digital or analog analysis. Digital analysis requires relatively expensive hardware and software as well as experienced technicians. Analog analysis, in comparison, requires a lower level of hardware and supporting skills. Software for digital analysis is based on tapes containing data from LANDSAT MSS; such tapes are expensive and require special computer programs in pattern recognition. Analog analysis employs aerial photographs, multispectral photography, space photography (SKYLAB), or multispectral scanner frames (LANDSAT), which are relatively inexpensive in comparison to the data tapes.

In Brazil, ready access to computers is difficult; the only institution which offers processing and analysis of remotely-sensed data is the National Institute of Space Research (INPE), located in São Paulo State. Therefore, for use in Brazil, the densitometer -- an instrument that measures density of aerial or multispectral scanner imagery -- used in analog analysis is the most attractive solution for analysis of remotely sensed imagery.

Objectives

The objective of this study is to develop a method to distinguish the four major forest types of Parana State, Brazil. The forest types are: (1) Pines (Pinus taeda and Pinus elliottii); (2) Parana Pine (Araucaria angustifolia (Bert.) O. Ktze.); (3) Hardwoods; and (4) Mixed (Conifers and Hardwoods). In order to achieve this objective, considering the nature of the study area, the following analytical procedures were carried out:

a. determine the potential of LANDSAT MSS imagery to differentiate forest types through densitometric and discriminant analysis of forest canopy reflectance;

b. determine the potential for identifying forest types through analysis of canopy reflectance from low-level photography (1:40,000 scale), SKYLAB (1:950,000 scale), and LANDSAT MSS (1:1,000,000 scale) imagery.

In developing this methodology, the main considerations are:

a. Spatial properties such as texture, size, and shape to distinguish between forest types;

b. the precise spectral interpretation of the density characteristics of grey or color separation; and,

c. the use of spectral signatures to differentiate major forest types.

Background on Parana State

The State of Parana is located in the south of Brazil. It's total land area of approximately 201,203 square kilometers may be subdivided in five natural geographic regions: coastal zone; sea-mountain range; first plateau; second or Ponta Grossa plateau; and third or Parana Trapp plateau.

In a study done by Maack, in 1968, the natural vegetation cover types of Parana consisted mainly of pluvial tropical sub-tropical forest and Araucaria angustifolia (Bert.) O. Ktze. forest, which corresponded to 83 percent of the state's total land area. The other 17 percent included fields, swamps, beaches, and bays. Table 1 shows these natural vegetation cover types with their areas; and their distributions are further delineated in Figure 1 and Table 2.

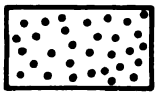
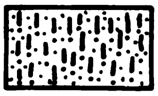
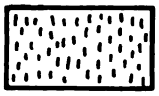
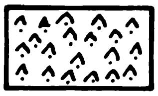
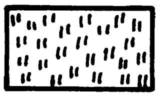
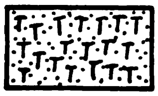
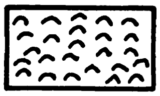
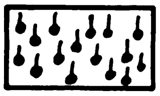
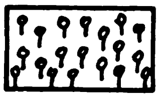
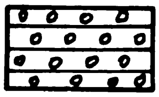
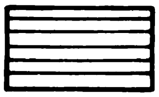
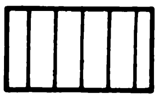
Since 1930, much of the primitive landscape of Parana has been changed beyond recognition. Forest lands have been affected the most, the Acungui Serie mountain region being one example. Its depletion, which began in the early 1930's and still continues, was caused by the exploitation of the forests for lumber, mainly of Araucaria angustifolia (Bert.) O. Ktze. In Parana's third plateau the depletion has been more intensive; it began in 1935 when coffee farmers moved west of the Tibagi river to clear new areas for cultivation. In 1963, Parana had approximately 6.1 million hectares of forest cover, comprising 30.27 percent of Parana's total land area. This resource was distributed irregularly across the state, with large concentrations in the west and

Table 1: Natural vegetation of Parana State[a] (Maack, 1968)

Cover	Area (%)
Pluvial tropical sub-tropical forest	46.74
Araucaria angustifolia (Bert.) O. Ktze. in plateaus and in the sub-tropical forest region	36.67
Fields	15.17
Swamp	0.88
Beach, islands vegetation	0.26
Bays with swamp strips areas	0.28

[a]Parana State contains 201,203 square kilometers.

Table 2: Phytogeographic map convention

	- Tropical pluvial forest less exuberant
	- Coastal tropical pluvial forest
	- Fields from the type of the savannas from the central region of Brazil
	- Secondary forest in the Araucaria zone
	- Fields with bushes
	- Araucaria angustifolia (Bert.) O. Ktze. forest region
	- Sub-tropical pluvial forest
	- Tropical pluvial forest from plateaus
	- Ilex paraguariensis region in the Araucaria region
	- Devastated secondary forest in the pluvial tropical region
	- Devastated zone in the pluvial sub-tropical forest
	- Devastated forest from the tropical pluvial forest less exuberant.

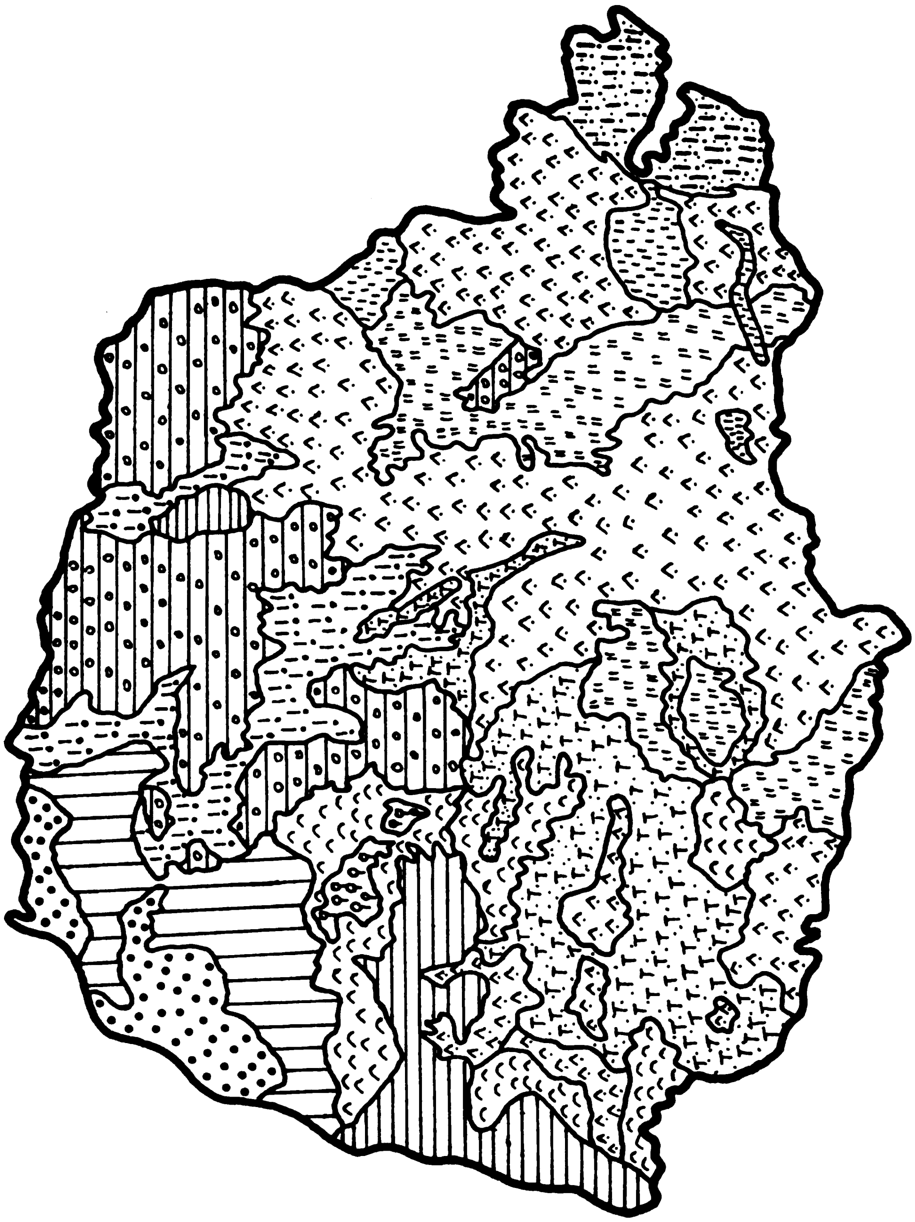


Figure 1: Phytogeographic map of Parana State (Maack, 196[illegible]).

southwest. In 1973, forest cover had decreased to 2.4 million hectares, or 60.93 percent of the 1963 forest cover was harvested in ten years. In the coastal zone this harvesting was mainly for charcoal production. These areas, including the coastal zone, have been replanted with Pinus elliottii, Pinus taeda, Araucaria angustifolia (Bert.) O. Ktze., and Eucalyptus spp (U.F.P., 1974).

However, the greatest impact in the forest sector has been caused by the Fiscal Incentive Law (1966); this allows income tax deduction of costs associated with forestation and reforestation. This law caused a large jump in reforested areas. In 1967, 5,954 hectares were planted; from 1968 to 1977, however, an average of 54,545 hectares/year were reforested. This increased reforestation has had a very significant positive impact in the technical, social, and economic aspect of life, as well as in the other sectors of the state economy.

Therefore, because of the various changes in the forest resource base, forest resource planners need a method to quantify changes in forest composition and forest cover over an extensive area. Remote sensing techniques, using either aerial photos or satellite imagery, can provide such information.

Definition of Terms

Since some readers may not be familiar with the terminology of remote sensing, several terms basic to discussion of the study will be defined. Densitometry or sensitometry involves the accurate measurement of the sensitivity of a photographic emulsion to light. However, since the effect of different exposures are apparent only after film development, sensitometry actually deals with both exposure and development and their relation to one another.

Density is a measure of the degree of blackening of an exposed film, plate, or paper after development, or of a direct image as in the case of printout material. Density measurements in color films can be divided into two general classes. One class, analytical density, measures the amount of dye in each of the layers of a color material. The use of analytical densities enables densitometric studies of cause-and-effect relationships in color image formation. The second class is integral densitometry, which is the measurement of the integrated effect of the combined color images. Integral density measurements of color films correspond to the same type of density measurements made on black and white films.

The photographic tone on black and white or on color photographs is the brightness of a lighted and developed area of a photographic layer. The size of the grey density value or of the color density is expressed by the percentage of light which is absorbed by the affected area (on prints or on films). Density (D) is generally defined as the logarithm to base 10 of opacity (0), or as the logarithm to base 10 of the inverse of transmittance in the film (T), or as the logarithm to base 10 of the inverse of reflectance capability on a paper print (R):

$$\begin{aligned} D &= \log_{10} 0 \\ &= \log_{10} {}^{1}/_{T} \\ &= \log_{10} {}^{1}/_{R} \end{aligned}$$

A grey density indicates a color density of 1, or that 10 percent of the incident or penetrating light is either reflected or transmitted. Therefore, density depends upon the total proportion of the illuminated area that is obscured by the silver grains. Resolution, in this sense, has been found to be a fairly good index of the detail-reproducing capabilities of photographic emulsions. This is because the image-forming characteristics are fairly similar, qualitatively, despite the quantitative differences. On the other hand, the image-forming properties of lenses vary from lens to lens because of differences in the way the distortions are balanced against one another. For a given lens, the distortions vary from one focal position to another and across the lens field for a given focal position. The result is that resolving power does not always indicate the relative sharpness with which the edges of large details are reproduced.

Photographic texture can be regarded as an extension of photographic tone. Texture involves both the mean photographic tone of an object or unit and the variation within this unit (the change of the grey density or color in the smallest area). In small scale imagery (e.g., LANDSAT MSS) where form and especially details of single objects can no longer be reproduced - the object being below the minimum resolution limits - texture has great significance in object identification.

Reflection refers to the reflected radiation by a surface, without change of frequency of radiation's monochromatic component. While reflectance is an expression of the ratio of reflected energy (in all directions) to incidence energy, brightness refers to the radiant flux per unit area of surface in a specified direction of observation.

CHAPTER II

LITERATURE REVIEW

Discriminant analysis in taxonomic problems began with Fisher (1936), who stated that: "When two or more populations have been measured in several characteristics, $x_1, \ldots, x_s$, special interest attaches to certain linear functions of the measurements by which the populations are best discriminated." Sneath and Sokal (1973) pointed out that the purpose of a discriminant function is to minimize the probability of wrong assignment of unknown individuals (misclassification) where there are clusters that group closely together between which identification must be as certain as possible.

A major problem is that the usual methods of discriminant analysis assume that the dispersion matrices of all the taxa are homogeneous (that is, the clusters all have much the same size, shape, and orientation in phenetic space) and that the clusters or groups follow multivariate normal distributions. If the group dispersion matrices are equal, then linear classification rules should be used. Otherwise, quadratic rules should be employed (Sneath and Sokal, 1973).

The application of discriminant analysis in studies of density, reflectance and or transmittance patterns on multilevel imagery of forest types requires some precautions when interpreting the imagery. Sayn-Wittgenstein (1978) emphasized that the recognition of species on aerial photographs depends on the scale of the photos.
Tree form characteristics such as crown shape and branching habit,

which are used for identification on large scale photographs, become progressively indistinct as the scale is decreased. Eventually, these features become so indistinct that they are replaced as key characteristics by photographic tone, texture and shadow pattern. These characteristics depend upon too many variables to make specific rules of interpretation possible. Such variables include sun altitude, length of exposure, method of printing and developing, atmospheric haze, and of camera and lens characteristics. The sun altitude involves the relationship between reflectance (in LANDSAT MSS) and difference between slope azimuth and sun azimuth (Sadowsky and Malila, 1977). This relationship is very important because as reflectance and azimuth decrease, so does percentage of reflectance. This can result in a diminishing capability in separating conifers from hardwoods. Species identification on small-scale aerial photographs is more of an art than a science if it is pursued by conventional interpretation methods (Sayn-Wittgenstein, 1978). Its success, to a high degree, depends upon the interpreter's skill and knowledge of the area. Interpretation remains a subjective procedure, even when using advanced methods (e.g., density measurements by means of densitometric analysis, computer-assisted pattern recognition) for the classification of digital data.

With the increase of data to be analyzed, the present conventional interpretation methods for analyzing and integrating the information from aerial photos were surpassed. Ray and Risher (1960) found that quantitative measurement of photographic tone, either in terms of optical density of film materials or in terms of light reflectance from paper prints, may also be useful in geologic research. Frequency of tone changes, as well as the magnitude of tone measurements, are useful in describing and comparing different terrain features. It is important to note how tone and texture were quantified in terms of density units.

Rosenfeld (1962) worked with oscilloscope photographs of video traces, using only black and white conventional aerial photography at a scale of 1:5,500. The specific terrain types investigated included hydrographic features, cultivated and uncultivated land, and various types of urban areas. Rosenfeld (1962a) discussed an approach to the automatic identification of the images of targets on aerial photographs. This approach involved the extraction from the photographic image of two basic types of information, one relating to the presence of figures having given shapes and sizes, and the other to the "textural" nature of the image. Goldstein and Rosenfeld (1964), in order to obtain rapid, simple measurements of the "degree" to which an image contains elements of a given shape, employed a simple two-channel noncoherent[1] optical correlator. They used this optical correlator for automating terrain type discriminations, to convert the distribution of elements of the given shape in the original terrain image into a correlogram of brightness distribution. Doverspike _et al_ (1965) used blue, green and red color components at a scale of

---

[1]noncoherent: having waves not in phase and of various wavelengths.

1:1,188 to identify even-aged conifer stands, land use classes, and other features through micrometric measurements.

Rib and Miles (1969) studied the identification of terrain features based on the interpretation of form, tone and texture. Measurements of tone and texture were performed with a densitometer adapted for continuous scanning. The factors investigated included film types and filters, seasonal effects, aperture size, and scale. Measurements were also performed on multichannel imagery (ultraviolet through far infrared), and spectral response signatures were developed for various target materials. A technique was also developed for the preparation of isochromal maps (maps showing uniform color zones) from densitometric scans of color photography. The technique developed for isochromal mapping offers a method for automatically mapping various tonal patterns present on color photography. This has immediate application for the identification of those terrain features which are directly related to color tonal patterns. Efforts to develop diagnostic patterns for various terrain features from measurements on a single film type were not successful. Analysis of the spectral response curves developed from multichannel imagery indicated that this approach offered greatest potential for delineating terrain features.

Von Steen _et al_ (1969) studied the relationship of film optical density to crop yield indicators. Cotton, grain sorghum, carrots, cabbage, and onions were included in the study. Film densities of the sample plots were obtained by scanning the transparencies with an isodensitracer. Densities for each of the three layers of film were measured by using appropriate filters in the isodensitracer. Average density readings were then related to plant counts and measurements obtained from ground sample plots. Statistically significant relationships were found between the preharvested yield indicators and film densities of aerial infrared film. Jackson et al (1971) used microdensitometer traces from 35 mm false-color transparencies to compare potato fields infected with various levels of late blight. A significant linear relationship was developed between microdensitometer readings and field-disease ratings.

Driscoll et al (1974) used color infrared photos to identify plant communities and their components with a microdensitometer modified to generate scan lines. They pointed out that image density differences in color infrared aerial photos can be used to discriminate individual shrub and tree species of a pinyion pine-juniper plant community. In addition, image density was successfully used to identify six general plant communities. However, different sites and cultural treatments within native grasslands and ponderosa pine forests could not be easily discriminated, even through visual differences were apparent in the photos.

Akca (1971) conducted a study to identify land use classes and forest types by means of microdensitometer and discriminant analysis. The generated densograms[2] were then described by the arithmetic mean of grey density variation, mean quadratic deviation of grey density

---

[2]densograms: optical density readings drawn on millimeter-paper.

variations from the arithmetic mean of grey density (amplitude), and mean spatial frequency of grey density variation per mm. The task was to combine these three parameters into some measurement such that the most advantageous separation of the individual groups was achieved. This was accomplished through discriminant analysis.

Akça (1971) also described various studies (Steiner and Haefner, 1965; Maurer, 1965; Steiner, Maurer and Kilchenmann, 1966; and Baumberger, 1969) done at the Geographical Institute of the University of Zurich. These studies utilized black and white and color aerial photos to identify agricultural crops through measurements of color components and general density. These measurements were analyzed by discriminant analysis to achieve differentiation between various agricultural crops.

Jordan et al (1978) demonstrated that manual densitometry was a valuable tool for cover classification in areas of surface-mining. Manual spot densitometers were used to obtain cover signatures for 118 strata from multi-temporal 1:24,000 scale color infrared and multi-spectral aerial photographs. Linear discriminant analysis and multi-seasonal imagery led to a reasonably accurate classification system.

Lillesand et al (1979) investigated stress levels in urban trees on the basis of spectral densities measured on large-scale color and color infrared photography. Factor analysis was used to develop quantitative "stress indices" based on ground data. Multivariate regression analysis was then used to develop a statistical model for predicting stress indices on the basis of image density measurements. Tests of the model indicated that the photo-based predictions were as reliable as ground estimators, particularly under drought conditions. However, the timing of aerial photography with respect to rainfall weighted heavily on the success of quantifying tree stress from density measurements.

The use of remote sensing techniques has long been accepted in forestry. At an early stage of development, aerial photography was recognized as a valuable tool in land resource management, and advantage has been taken of improvements in techniques. The launching of LANDSAT - 1 in June 1972 and SKYLAB missions between May 1973 and February 1974 introduced a new technology. The assessment of its potential value to forestry became an urgent need.

High-resolution sensors such as those on SKYLAB may play an important part in forest and rangeland surveys. Although conventional aerial photographs have been utilized in resource surveys for several decades, recent developments have moved toward more sophisticated photographic and non-photographic remote sensors and computer-assisted data analysis. The new technology is valuable for several reasons: (1) costs of acquiring resource data are rapidly increasing; (2) more resource data is required at shorter intervals to measure rapid changes in land use that affect the environment; (3) urban and recreational uses of land are encroaching upon available resources and alternate sources must be planned and provided for; and (4) there is a continuing need for up-to-date resource information for day-to-day land management decisions and program planning.

SKYLAB photography showed the possibilities of very small-scale imagery. The U.S. Forest Service has carried out a series of experiments (Francis, 1976) using microdensitometer evaluation of SKYLAB photographic products for classifying plant communities. A two-tailed t-test was used to determine presence of significant differences between optical density sample means for both region[3] and series[4] level classifications. Visual interpretation of SKYLAB and support aircraft photographic products could be used successfully for classification and aerial mapping of native plant communities up to the region level. Series level classifications were dependent on date, scale, and film type.

In terms of LANDSAT scanners, data for forestry must be in a useable form. The satellite's multispectral scanners (MSS) transmit photometric data received from the earth's surface; the data must be reconstructed for visual interpretation or densitometric measurements. The raw data is then converted into photographic form as black-and-white or color transparencies. Also, MSS produce magnetic tapes which are capable of producing, through computed-assisted analysis techniques, graphic representations of the earth as viewed by the satellite. Both the transparencies and tapes have low detail resolution, compared with conventional aerial photography.

---

[3]region: subdivisions of most general class of vegetation, associated regionally and therefore determined by sub-climates within continental climates: Montane Grassland, Temperate Mesophytic Coniferous Forest, Alpine Grassland, etc.

[4]series: a group of vegetation systems within the region category, with a common dominant climax species: Ponderosa Pine Forest, Fescue Grassland, Herbaceous Meadow, etc.

Driscoll and Francis (1975) used LANDSAT-1 imagery to classify plant communities in region and series level in central Colorado. A scanning microdensitometer was used to evaluate the June and August LANDSAT-1 color composite for classifying the plant communities at a scale of 1:1,000,000. Standard t-tests for unpaired plots with unequal sample sizes indicated highly significant differences between all vegetation combinations at two dates (July and August). Thus, it appears possible to automatically scan a LANDSAT color composite and relate density levels to regional level vegetation classes.

The LANDSAT study concluded that, with improved spectral and spatial resolution, satellite imagery could provide the first-level information required in extensive forest inventory sampling strategies. SKYLAB data provided an opportunity to investigate and substantiate the conclusion, using machine-assisted classification procedures in addition to those dependent on human skill and judgement.

CHAPTER III

RESEARCH METHODS

Theoretical Model of Individual Decision-Making

The approach used to discriminate the four major forest types of Parana State is based on texture type separation of density measurements. However, the subdivision of a photograph into forest types is not always easy for a human interpreter. In some cases the boundaries between forest types will be sharply defined, while in other cases the transitions between adjoining forest types may be quite gradual and difficult to delineate. Fortunately, mathematical models for the detection of subdivision points can be formulated using any of the textural variables defined in the previous chapter. The model used in this study utilized the following variables: (1) the arithmetic mean of grey density variation; (2) the mean quadratic deviation of grey density variations from the arithmetic mean of grey density; (3) the moment coefficient of skewness of grey density variations from the arithmetic mean of grey density; and (4) the moment coefficient of kurtosis of grey density variations from the arithmetic mean of grey density.

This model is used to discriminate between forest types by principal components analysis, multivariate analysis of variance, discriminant analysis, and cluster analysis of canopy reflectance characteristics.

Research Hypothesis

The heart of remote sensing discrimination between objects is the wavelength selectivity of the interaction of electromagnetic energy with these objects. The amount of energy in different wavelengths that is returned to the sensor from a given object defines a theoretically unique spectral signature, analogous to human fingerprints. Figure 2 shows typical spectral signatures for hardwoods and conifers based upon reflectance in visible and near visible wavelengths. It can be seen that the spectral signatures overlap through all visible wavelengths, and particularly through the green region (0.5 to 0.6 um). Therefore, both types of trees can display identical reflectances, i.e., they have essentially the same color. In the infrared region, however, reflective interactions are much greater for hardwoods than for conifers. This property makes infrared photography extremely useful for discriminating between these tree types. On infrared photos, hardwoods are easily detected as being relatively brighter than conifers. This model used spectral reflectance (density values) for forest types between 0.4 micrometers (um) and 1.1 micrometers.

One weakness of this canopy reflectance model is its assumption that canopy components are distributed within each canopy layer in a simple random fashion. The model assumes that leaves are distributed so as to be located equally in any position within a canopy layer in

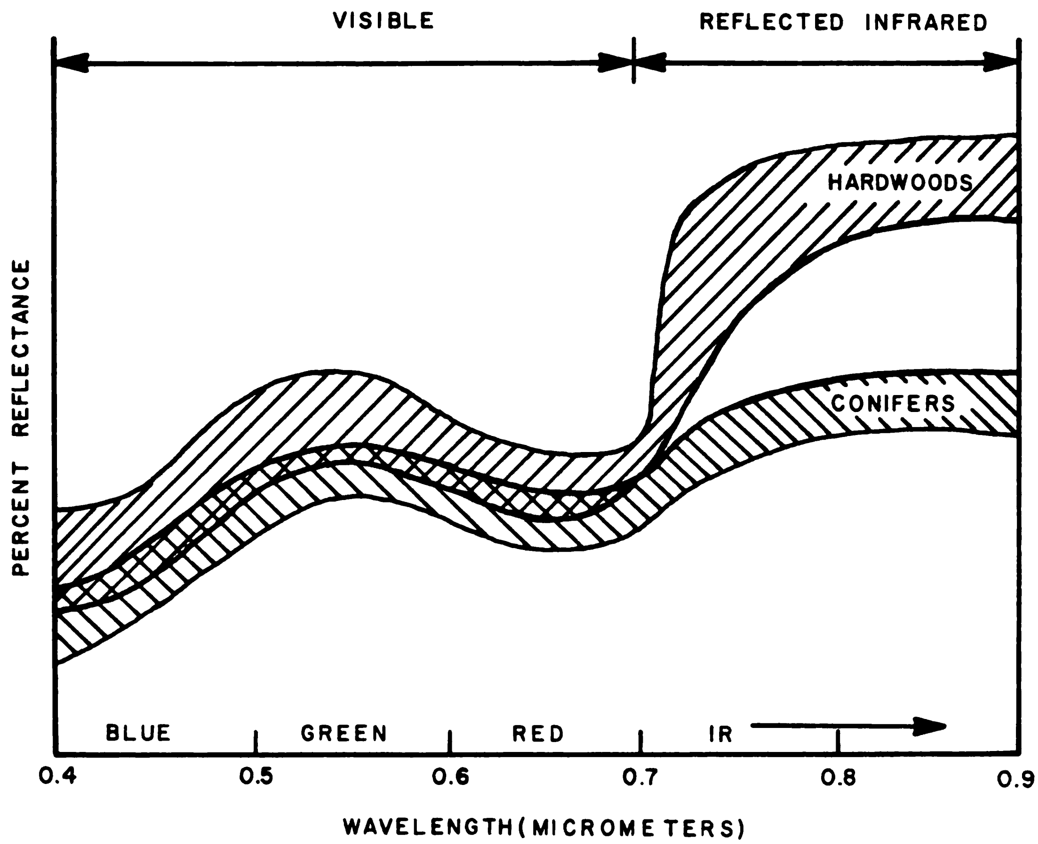


Figure 2: Spectral signature for hardwoods and conifers (Lillesand, 1976).

accordance with their concentration in that layer. This random distribution allows for the occurrence of occasional clumps of leaves with a predictable frequency, a situation perhaps best represented by grass canopies that contain more or less uniformly distributed components. However, it may not predict a situation that exhibits a distribution more clumpy than would be expected by the equally likely placement of a simple random distribution. A forest canopy represents such a situation, in that leaves and branches are associated with tree crowns as a clump with larger voids between crowns. In this way, the influences of clumped component distributions on canopy reflectance were believed to be properly represented by the radiometric effects that affect the sensor, and not by changes in percent cover (Sadowsky and Malila, 1977).

The study's hypothesis is that the spectral and spatial variations of the forest canopies signatures (reflectance) will cluster separately for each forest type in each imagery, and that the optical densities of forest canopies signatures are sufficiently discrete to allow discrimination between them.

Experimental Procedure

The study area for this research corresponds to a LANDSAT frame (Figure 3) with coordinate points: #1 S[5]23°55'26" - W[6]048°45'10"; #2 S23°39'27" - W050°29'29"; #3 S25°10'07" - W050°57'28"; #4 S25°26'22" - W049°11'55". It was chosen for the following reasons: (1) restricted

[5]S: corresponds to latitude south of the Equator.

[6]W: corresponds to longitude west of Greenwich.

availability of aerial photos regarding the date, scale, and coverage in Parana State; (2) restricted quality and coverage of the SKYLAB imagery; and (3) the area contains a good cross-section of the forest type-terrain combinations (Figure C5).

Four different levels of imagery were used in the study (Table 3); first, panchromatic aerial photos; second, color SKYLAB S-190 B imagery; third, subdivided in four sub-levels[7], LANDSAT-2 MSS imagery; and fourth, subdivided in nine sub-levels, were the ratio forms. Ratio form is an important procedure not only for normalization[8] but also for its enhancement of features (forest types, variations in rock chemistry in desert areas, etc.). Note that the ratio forms are only possible for LANDSAT MSS imagery, since the density values measured, for instance, in channel 4 were simultaneously measured in channels 5 and 7. These channels produce identical scales and formats in different spectral regions. The aerial photos were taken in 1976, SKYLAB imagery in August 1973, and LANDSAT MSS in September 1977. The imagery for disparate dates were used as they were the only ones available.

The forest types analyzed were: (1) Pines (*Pinus taeda* and *Pinus elliottii*); (2) Parana pine (*Araucaria angustifolia* (Bert.)

---

[7] The fourth sub-level, color composite, is a composition of channels 4, 5, and 7.

[8] normalization: eliminates or reduces the influence of image distortions involving scene radiation variations such as falloff in light intensity away from the center of a photographic lens, cross-field variations in lighting. Also,systematic variations in scene illumination and developing and printing procedures, scan lines variations arising from differences in viewing angle, and faulty equipment biases.

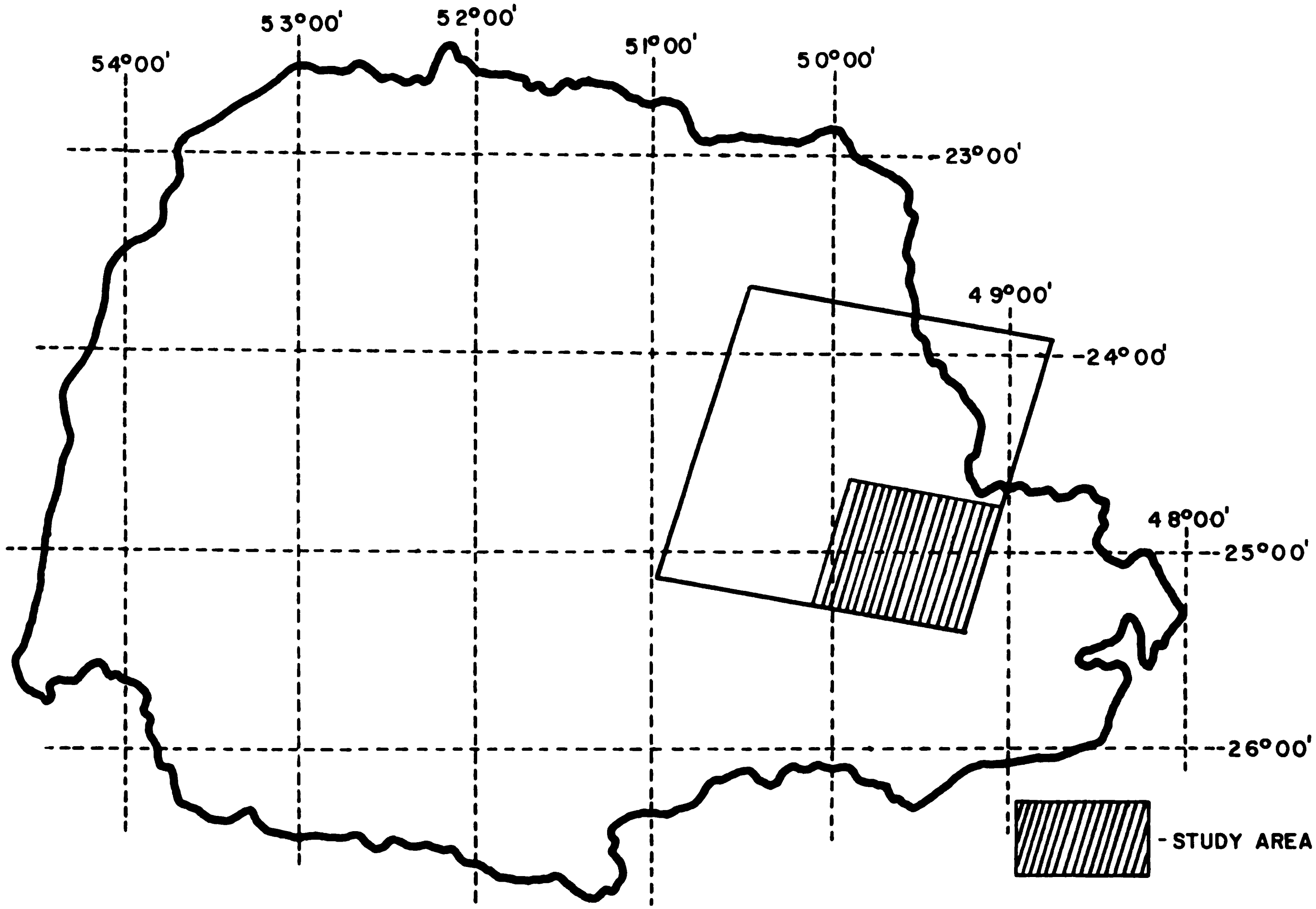


Figure 3: Localization of study area in LANDSAT frame in Parana State, Brazil.

Table 3: Levels of imagery used in the study.

Level[a]	Type of Imagery	Spectral Coverage (um)
1. Aerial Photo (1:40,000)	Panchromatic	0.4 - 0.7
2. SKYLAB (orbital) (1:950,000)	Color	0.4 - 0.7
3. LANDSAT (orbital) (1:1,000,000)	MS Scan[b] - Channel 4	0.5 - 0.6
	MS Scan - Channel 5	0.6 - 0.7
	MS Scan - Channel 7	0.8 - 1.1
	Color Composite	0.5 - 1.1
4. Ratio	Ratio 1: 4/5[c]	-
	Ratio 2: 4/7	-
	Ratio 3: 5/4	-
	Ratio 4: 5/7	-
	Ratio 5: 7/4	-
	Ratio 6: 7/5	-
	Ratio 7: 4/ 4+5+7	-
	Ratio 8: 5/ 4+5+7	-
	Ratio 9: 7/ 4+5+7	-

[a] First and second levels are paper print products and third level are film positive products.

[b] MS Scan - Multispectral scanner imagery.

[c] 4/5 - represents Channel 4 divided by Channel 5, and so on.

O. Ktze.; (3) Hardwoods; and (4) Mixed (conifers and hardwoods). These forest types were chosen for several reasons. Pines were chosen because they are an exotic species introduced in Brazil and now contributing significantly to forest coverage. Parana pine was chosen because it is the only native commercial coniferous species. Hardwoods and Mixed were chosen because they are both natural forest types (Appendix C).

Forest types were delineated on the imagery in the following manner. First, the forest types were identified and delineated through photointerpretation of the aerial photography (1:40,000 imagery) with a mirror stereoscrope. These delineations were used as ground-truth based on former photointerpretation and author's background on the area. These forest types were then transferred to the SKYLAB imagery and to the individual frames in the LANDSAT MSS imagery, using a magnifying comparator. Since each level of imagery is independent, the sample area is proportional to the size of the forest types delineated in the 1:40,000 scale after photointerpretation.

Random sampling procedures were then used to select observational units. Five areas for each forest type were delineated on the aerial photos, SKYLAB imagery, and on the individual LANDSAT channels, and 25 density measurements were taken in each area. This procedure provided 125 density measurements for each forest type in each scale of imagery. These measurements were taken at random with the densitometer.[9] Past studies (Doverspike, 1965) suggested that a minimum of 100 replicates are needed for each forest type in each scale of imagery. For the ratio forms, caution is necessary when collecting

---

[9] See Appendix A for description of the densitometer.

such density measurements, because each density value taken in channel 4 should also be taken in channels 5 and 7. Therefore, each channel was enlarged in the densitometer until each individual line and pixel[10] could be seen. After the density values were taken for each individual channel, the ratio forms were obtained by dividing each individual density value from one channel by its respective density value in the other channels.

Statistical Parameters Used in Discrimination.

The mathematical model for the detection of subdivision forest types uses statistical moments as texture variables. Specifically, these moments are: the arithmetic mean of grey density variations, the mean quadratic deviation of grey density variations from the arithmetic mean of grey density, the moment coefficient of skewness of grey density variations from the arithmetic mean of grey density, and the moment coefficient of kurtosis of grey density variations from the arithmetic mean of grey density.

For a given set of data $\{x_1, x_2, \ldots, x_n\}$, the first moment, the mean, is defined as:

$$\bar{x} = \frac{1}{n} \Sigma x_i , \qquad (1)$$

which identifies the degree of opacity of the photographic texture. The second moment, variance, is defined as the square of the standard deviation:

$$m_2 = \frac{1}{n} \Sigma (x_i^2 - \bar{x}^2) \qquad (2)$$

[10]pixel - picture element — in a LANDSAT MSS imagery, each pixel has 79 x 79 meters ground resolution cell and a nominal ground resolution of 56 x 79 meters because of cell spacing.

This study used the standard deviation (or mean quadratic deviation), which distinguishes the contrast differences in the texture or the smoothness of the texture:

$$s = m_2^{\frac{1}{2}} \qquad (2a)$$

The third moment, skewness, is defined as the degree of asymmetry or departure from symmetry of a distribution (Figure 4), and, in our case, of differences in texture between samples. Skewness is defined as:

$$m_3 = \frac{1}{n} \Sigma x_i^3 - \frac{3}{n} \bar{x} \Sigma x_i^2 + 2\bar{x}^3 \qquad (3)$$

In this study, the moment coefficient of sknewness was used, because it is easier to calculate and is scaleless:

$$\gamma_1 = \frac{m_3^{\;3/2}}{m_2} \qquad (3a)$$

The fourth moment, kurtosis, is defined as the degree of "peakedness" of a distribution (Figure 4). In this study, kurtosis defined the differences in texture between samples. Kurtosis is defined as:

$$m_4 = \frac{1}{n} \Sigma x_i^4 - \frac{4}{n} \bar{x} \Sigma x_i^3 + \frac{6}{n} \bar{x}^2 \Sigma x_i^2 - 3\bar{x}^4 \qquad (4)$$

In this study, the moment coefficient of kurtosis was used, again because it is easier to calculate and is in dimensionless form to avoid particular units:

$$\gamma_2 = \frac{m_4}{m_2^2} \qquad (4a)$$

For perfectly symmetrical curves, such as the normal curve, skewness is zero and kurtosis has a value of three.

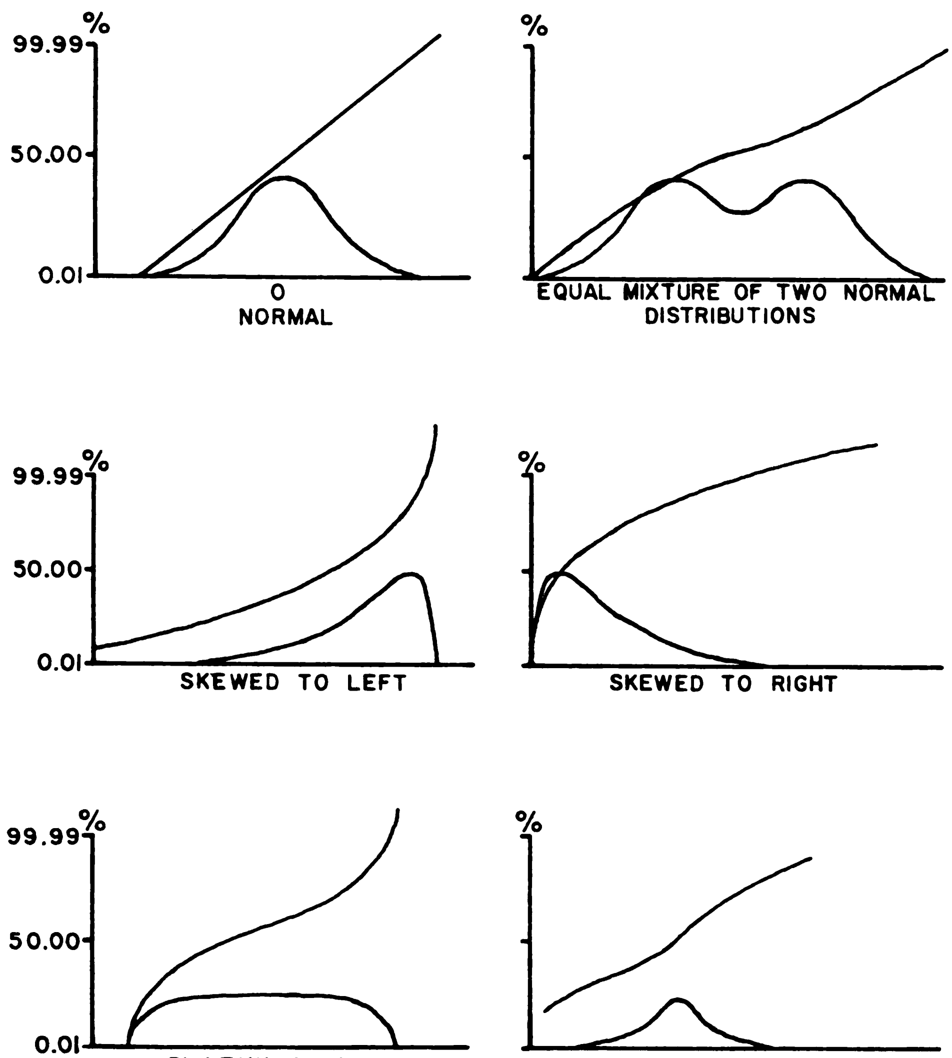


Figure 4: Examples of basic frequency distributions with their cumulative distributions plotted with the ordinate in normal probability scale (Sokal and Rohlf, 1960).

The four parameters defined above were used in the discrimination model. To illustrate its use, assume there are three distributions, D1, D2, and D3, as in Figure 5. D1 has a high peak while D3 is almost normal. Therefore, discrimination between D1 and D3 could be achieved through the use of differences between means, or by kurtosis. Comparing D2 and D3, D3 has larger skewness and different standard deviation than D2. Therefore, discrimination could be achieved through the use of differences in skewness and standard deviation.

Techniques for Data Analysis

The statistical analyses that were used to discriminate between forest types were principal component analysis, multivariate analysis of variance, discriminant analysis and cluster analysis.

Principal components[11] are linear combinations of random or statistical variables which have special properties in terms of variances. For example, the first principal component is the normalized (that is, the sum of the squares of the coefficients being one) linear combination with maximum variance.

The principal components are the eigenvectors of the covariance matrix. Thus the study of principal components can be considered as putting into mathematical terms the developments of eigenvalues and eigenvectors (for positive semi-definite matrices). In many exploratory studies the number of variables under consideration is too large to handle. Since it is the deviations in these studies which are of

[11]See Morrison 1976, Anderberg 1973, and Anderson 1958 for the theory behind principal components.

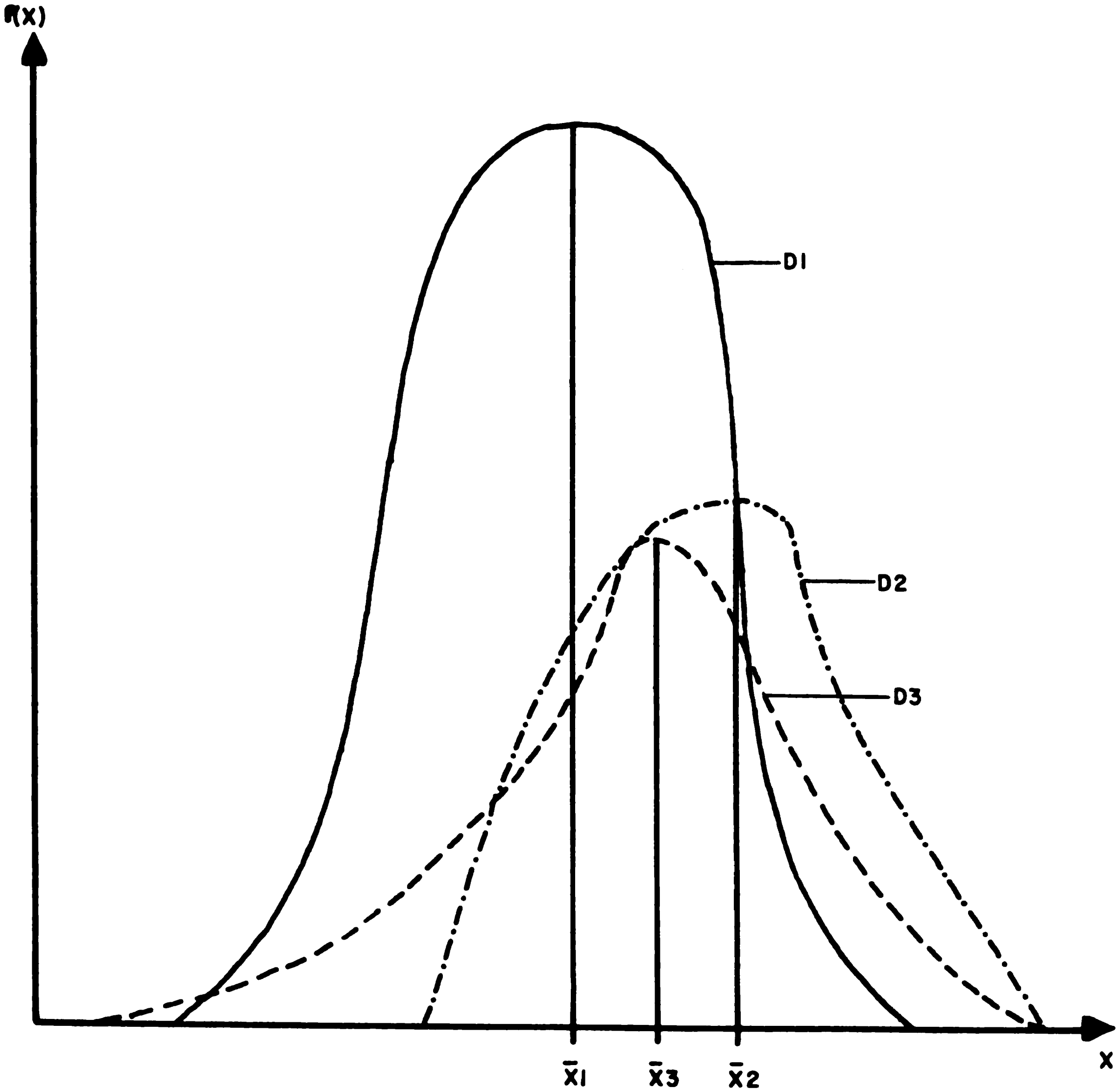


Figure 5: Illustration showing three hypothetical distributions.

interest, a way of reducing the number of variables is to discard those linear combinations which have small variation and study only those with large variation. In this study, we were interested in describing and analyzing how density measurements of forest types differ in the statistical moments as texture variables. Thus, we wanted to know which statistical moments or combinations of moments showed considerable variation.

In multivariate analysis of variance[12] (MANOVA), we are concerned with the study of group differences (e.g., forest types) in location in a multi-dimensional space. The distinctive multivariate nature of MANOVA is that the dependent variable is a vector variable. This dependent vector variable is assumed to be multivariate normal in distribution with the same dispersion, or variance-covariance matrix, for each forest type. Equality of dispersion matrices is the MANOVA extension of the assumption of homogeneity of variances in ANOVA designs. The main concern of MANOVA is equality among the population centroids, or mean vectors. That is, concern whether some or all of the populations are centered at different locations in the measurement space spanned by the dependent vector variable.

In discriminant analysis[13] the objective is to examine how far it is possible to distinguish between members of various groups (forest types) on the basis of observations made upon them. Discriminant

---

[12] See Morrison 1976 for the theory behind MANOVA.

[13] See Morrison 1976, Green and Carroll 1976, Rao 1952, and Anderson 1958 for the theory behind discriminant analysis.

analysis is thus an extension to multivariate observations of the ordinary analysis of variance within and between groups. A collection of discriminant variables are selected by forming one or more linear combinations in the form of:

$$D_i = d_{i1}Z_1 + d_{i2}Z_2 + \dots + d_{ip}Z_p,$$

where D_i is the score on discriminant function i, the d's are weighting coefficients, and the Z's are the standardized values of the p discriminating variables used in the analysis. The possible rank n of the discriminating function (subspace) depends on the relative sizes of g, the number of groups, and p, the number of discriminating variables. Therefore, the maximum number of functions possible is either one less than the number of groups or equal to the number of discriminating variables, if there are more groups than variables.

Discriminant analysis sometimes fails because the resulting discriminant function separates the groups very poorly due to two factors: either the groups overlap in the chosen measurement space, or the groups cannot be separated by a function of the form adapted for the analysis.

If discriminant analysis gives disappointing results because the data set contains fewer identifiable groups than prior knowledge would suggest, then cluster analysis should be in the first rank of diagnostic tools.

The purpose of cluster analysis[14] is to divide the data into groups in the hope of detecting some sort of natural grouping. Basic

---

[14] See Anderberg 1973, Sneath and Sokal 1973, Hartigan 1975, and Dubes and Jain 1979 for the theory behind cluster analysis.

to cluster analysis is the assumption that is reasonable to seek clusters in the data characterized by possession of the properties of compactness[15] and isolation[16]. The objective is to have clusters as much compacted and isolated as possible.

The choice of a precise definition of "cluster" is designed to incorporate the idea that a grouping together of objects in terms of measure of dissimilarity[17] always reflects some aspect of their structure, but that there is a possibility of objects becoming separated accidentally.

Cluster analysis consists of two methods; hierarchic cluster methods, and non-hierarchic cluster methods. In hierarchic cluster methods, the end point of the process is a dendrogram, or tree diagrm. The kind of algorithm employed does not use the internal structure of the groups formed at each stage, but simply treats the set of groups as a new set of objects. Because of this, it is necessary to find a different kind of approach to make possible the description of non-hierarchic (overlapping) cluster method.

The non-hierarchic cluster methods allow clusters at each level to overlap. For a general system of subsets of a set we no longer

---

[15]compactness: number of internal edges.

[16]isolation: number of linking edges.

[17]dissimilarity: If X is a class of individuals known to be one of A,B, then its identification as A or as B gives information about the character states which describes it, further to the information given by knowing simply that X is one of A,B, but not which one. The amount of information given in this way is not well defined since it may differ for the two outcomes of the identification. Dissimilarity is a suitable typical value of the expected information gain for each identification.

have a convenient descriptive tool such as the equivalence relations, if we were to ask that non-hierarchic cluster methods should be completely general.

Computational Algorithms[18]

For discriminant analysis the following algorithms, available in the Statistical Package for the Social Sciences (SPSS) Subprogram DISCRIMINANT, were used:

(1) DIRECT - all independent variables are entered into the analysis;

(2) WILKS - the criterion is the overall multivariate F-ratio for the test of differences among group centroids;

(3) MAHAL - seeks to maximize the Mahalanobis distance between the two closest groups;

(4) MAXMINF - maximizes the smallest F-ratio between pair of groups;

(5) MINRESID - criterion which tends to separate groups that are close together, which minimizes R, the residual variation;

(6) RAO - is Rao's V, a generalized distance measure.

For hierarchical cluster, the algorithms used were:

(1) Single-Linkage - joins together the two closest objects to form a cluster;

---

[18] See Appendix B for description of algorithms and computer programs used in this study.

(2) Complete-Linkage - characterized by the longest link needed to connect every member of a cluster to every other member;

(3) Average-Linkage within the new group - characterize a cluster by the average of all links within it;

(4) Average-Linkage between merged groups - consists of evaluating the potential merger of clusters i and j in terms of the average similarity for links between two clusters;

(5) Centroid - merges at each state those two clusters with the most similar mean vectors or centroids;

(6) Ward - maximize an objective function based on the error sum of squares.

For non-hierarchical clusters, the algorithm available was based on a squared criterion. It searches for clustering with the smallest squared-error.

The objective of testing all these algorithms was to find the specific algorithm best suited to our unknown data structure.

CHAPTER IV

ANALYSIS AND DISCUSSION

Characteristics of the Sampled Population

This study develops a methodology for the discrimination of four forest types based on the reflectance levels (signatures) from their canopies through densitometric analysis. However, it is important to understand how the sampled population behaves before going further into the analysis. From the fifteen scales showed in Table 3, only the five which gave the best discrimination were analyzed in the study. The image types analyzed were: (1) Panchromatic, (2) LANDSAT: Channel 7, (3) LANDSAT: Color composite, (4) LANDSAT: 7/5, and (5) LANDSAT: 7/ 4+5+7 .

When analyzing density on a transparency or in a paper print, the process normally involves placing the film in a beam of light that passes or reflects through it. The darker the images, the less light is allowed to pass or reflect, the lower the transmittance or reflectance, the higher the opacity, and the higher the density. Some sample values of transmittance, opacity, and density are given in Table 4.

Table 4. Sample transmittance, opacity, and density values (Lillesand and Kiefer, 1979).[19]

% Transmittance	Transmittance	Opacity	Density
100	1.00	1	0.00
50	0.50	2	0.30
25	0.25	4	0.60
10	0.10	10	1.00
1	0.01	100	2.00
0.1	0.001	1000	3.00

[19]See Chapter I for definition of terms.

A primary advantage of measuring image densities is the ability to quantify the radiometric dimension[20] of photography to remove subjectivity from the interpretation process. Therefore, a graphical analysis of the sampled population (raw data) was done for each of the five scales. These analyses were based on a cumulative frequency distribution in which the shape of an observed distribution was examined for departures from normality. Also, such analyses permit estimation of the mean, standard deviation, skewness and kurtosis, as well as the range of densities for each forest type.

In the image type, (1:40,000 Figure 6) notice that the hardwoods separate very well from the other three forest types. The others, especially Araucaria and Mixed, have some overlapping of densities. This is due to the fact that the Araucaria included in the Mixed forced the densities to higher values. Pinus, being coniferous, has higher density values (see Figure 2). Relating the frequency distributions of Figure 4 to Figure 6, notice that Hardwoods, Araucaria and Mixed have almost the same skewness configuration. Pinus, however, has a skewness to the right. For kurtosis, again Pinus is the only one which differs appreciably from the other forest types. In channel 7 (Figure 7) the shape of the distributions change. First, the range of densities are lower because the reflectivity of the scene in channel 7 in the range of 0.8 um to 1.1 um (reflected infrared) is generally more uniform. Second, there is a better differentiation of the forest types, even though the distributions have quite similar skewness and kurtosis.

---

[20] Radiometric resolution is the smallest difference in exposure that can be detected in a given film analysis. It is not a characteristic of a film per se but is set by the ability of a given densitometer to discriminate between density levels.

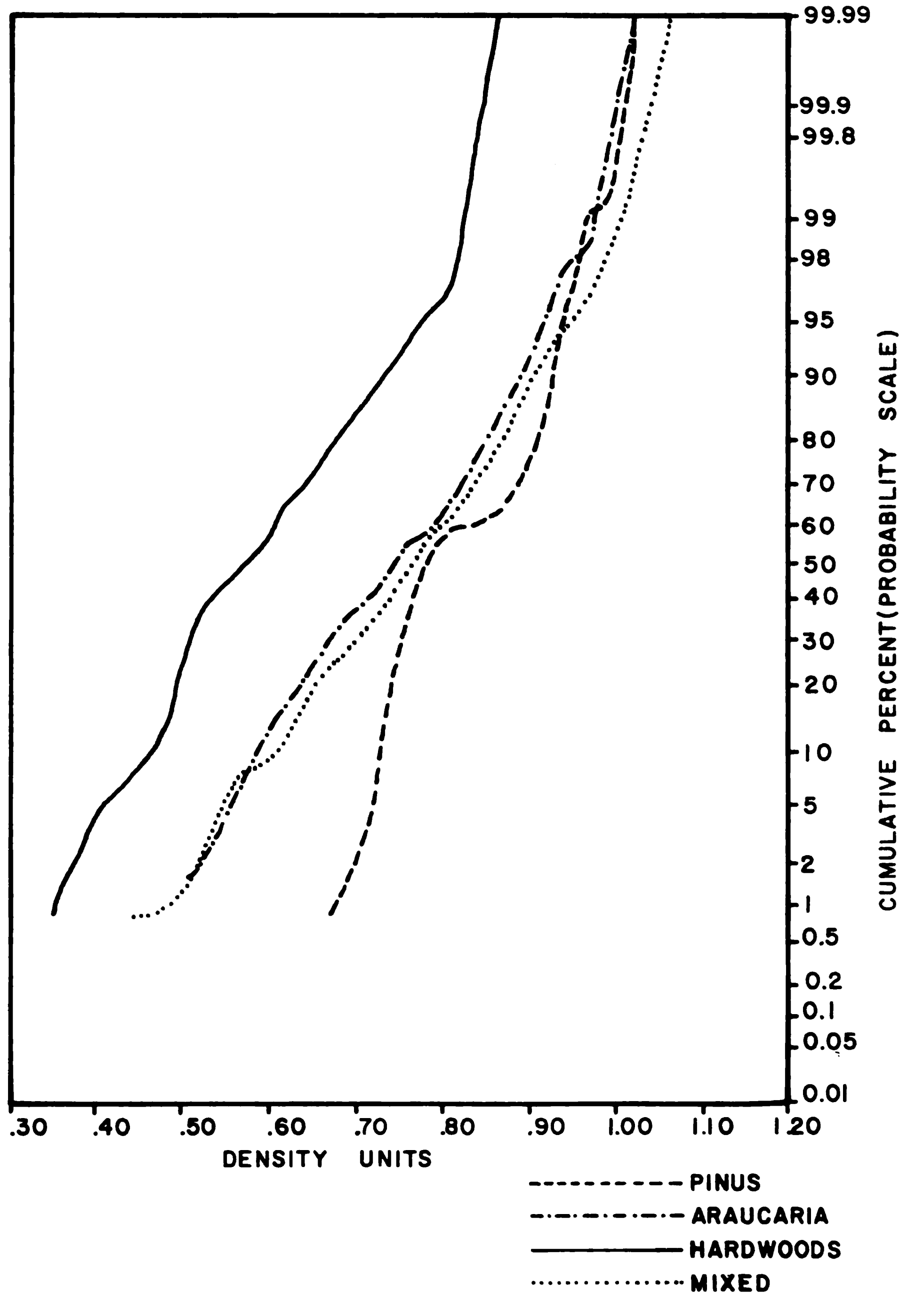


Figure 6: Frequency distributions with their cumulative distributions for 1:40,000 scale.

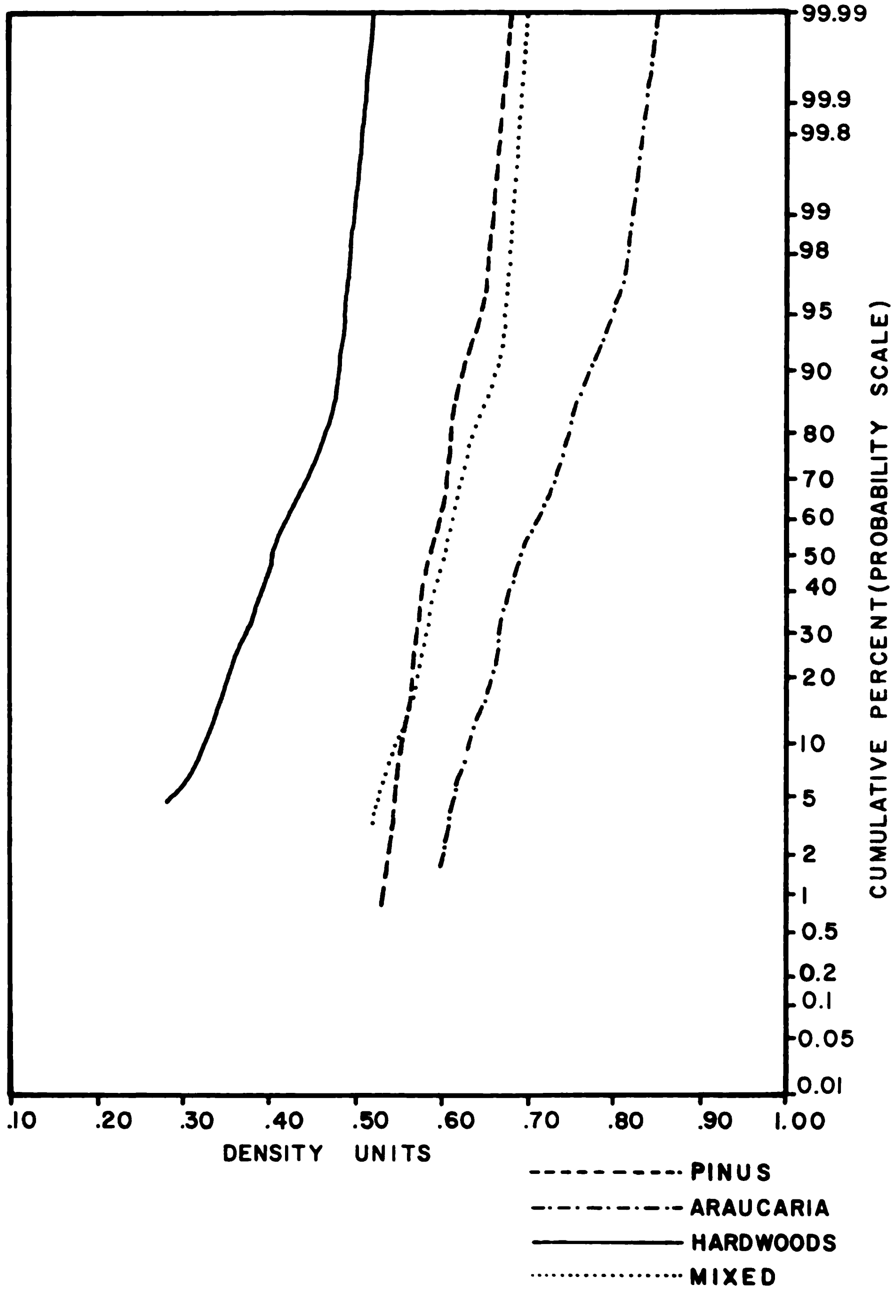


Figure 7: Frequency distributions with their cumulative distributions for channel 7.

The best separation of the forest types is in the color composite (Figure 8). The range of the densities is much larger (0.5 µm to 1.1 µm) due to the fact that channels 4 and 5 are in the chlorophyll absorption region and channel 7 is in the reflective infrared region (see Figure 2 and Table 3). Also significant is that Pinus and Araucaria have very similar high density values (lower transmittance, Table 4), while Hardwoods and Mixed group together with lower density values (higher transmittance). This grouping is due to the reflectance properties of coniferous needles and deciduous leaves. Conifer needles, because of their internal structure, absorb more light while deciduous leaves (Mixed and Hardwoods) reflect more light - especially infrared light. Otherwise, the four forest types have similar distributions, with Mixed having only a slightly skewness and kurtosis.

In ratios 6 and 9 the distributions are quite similar (Figures 9 and 10) with the forest types in the same order, starting with Hardwoods in the lower density values, followed by Pinus, Mixed and Araucaria. The only significant point is the difference in density ranges for ratio 6 to ratio 9. The higher range of ratio 6 is because it is a ratio of channel 7 by channel 5, which caused this broad range of density values (channel 5 is in the chlorophill absorption band).

Planned Comparisons

In considering the different analyses used (principal components analysis, multivariate analysis of variance, discriminant analysis, and cluster analysis) for building a meaningful canopy reflectance model,

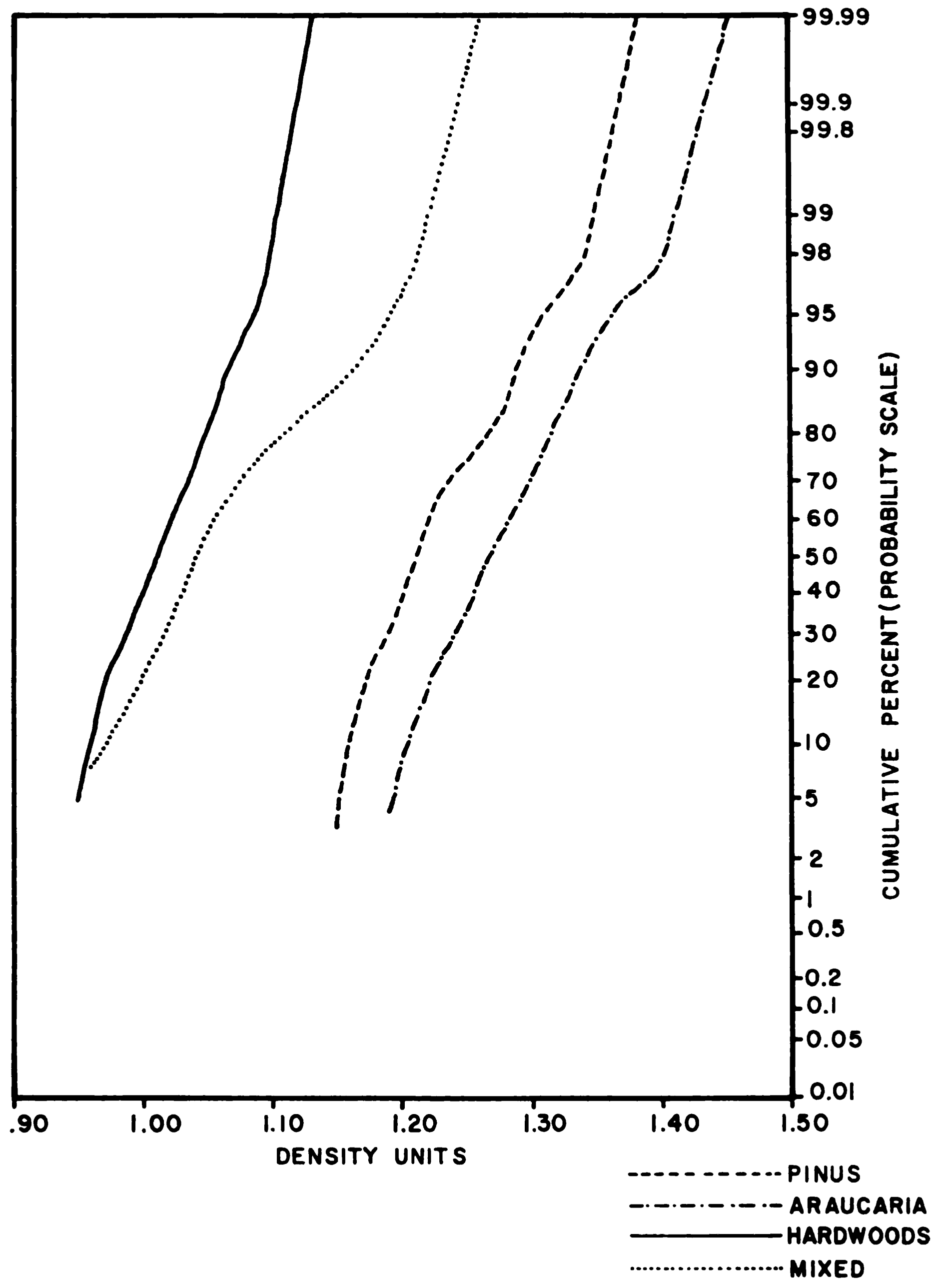


Figure 8: Frequency distributions with their cumulative distributions for color composite.

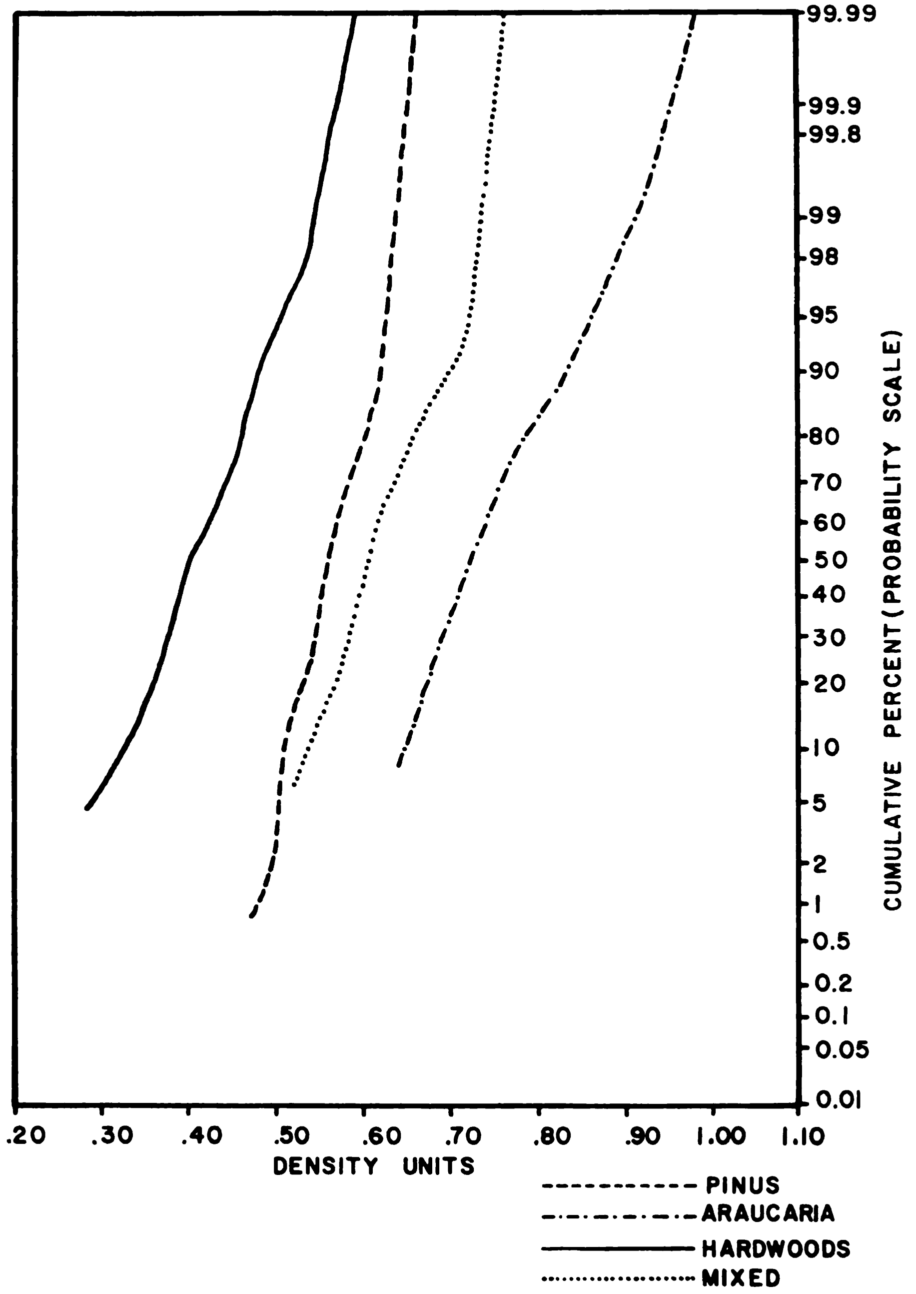


Figure 9: Frequency distributions with their cumulative distributions for ratio 6.

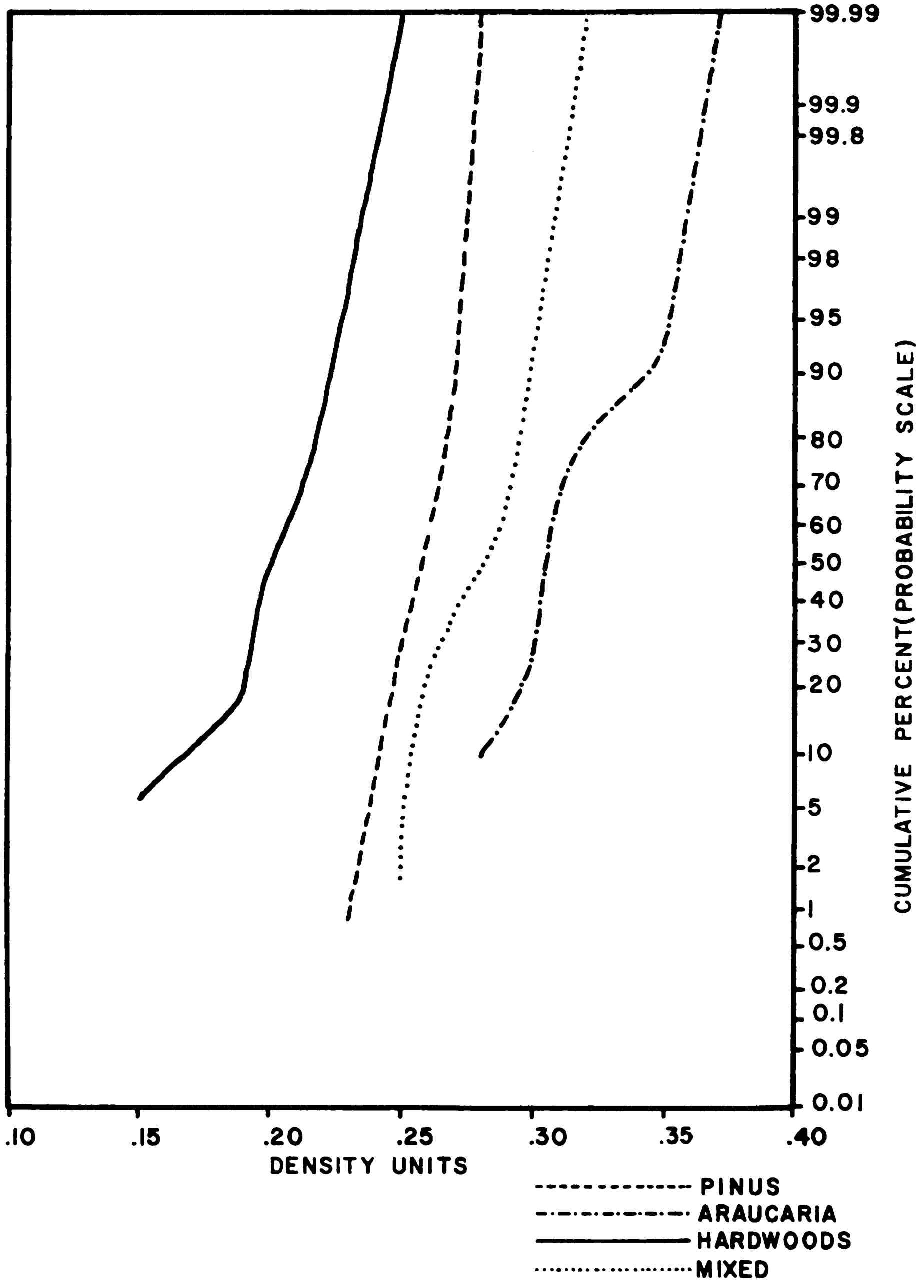


Figure 10: Frequency distributions with their cumulative distributions for ratio 9.

the first step was to test for differences between sample means within the individual forest types for all levels and sublevels of imagery (see Table 3). The ratio estimators were not included beacuse they are only a manipulation of the data, not actually measured density values. Table 5 shows the one-way analysis of variance for each scale of imagery for each forest type. Almost all were significant,indicating that the samples for each forest type are different. This is probably due to the following properties: (1) Tone variations from sample to sample causing differences in density readings; (2) pattern variations, especially in 1:40,000 scale and SKYLAB; (3) sun elevation, the main cause for tone variations; (4) terrain slope; (5) texture variations, especially in LANDSAT; and (6) crown (shape) differences between forest types. The error mean square, a measure of the accuracy of the sampling plan, gives the average dispersion of the 5 (n=25) samples in each group around the group centroid. In Table 5 we see that for all scales for all species the error mean square is very small, indicating that the sample means group very closely around the group centroid.

The next step was to study the internal structure of the variables used in the canopy reflectance model. This was accomplished by principal component analysis. The purpose here was to remove intercorrelations among the elements of a vector variable; if an uncorrelated vector variable is desirable the data must be transformed. As the dispersion of a standardized variable is a correlation matrix (R), the sphericity test (Bartlett's test criterion, Bartlett, 1950) was done to test the null hypothesis that R was an identity matrix. That is, the null

Table 5: One-way analysis of variance for individual species between samples (to test differences in samples)

Scale	Species	Error Mean Square	F-Value[a]
1:40,000	Pinus	0.001055	164.95**
	Araucaria	0.006844	31.40**
	Hardwoods	0.010202	6.79**
	Mixed	0.013640	2.93*
SKYLAB	Pinus	0.000581	253.93**
	Araucaria	0.000364	63.92**
	Hardwoods	0.000851	6.45**
	Mixed	0.000276	8.89**
Channel 4	Pinus	0.001000	10.58**
	Araucaria	0.001496	39.95**
	Hardwoods	0.000980	21.76**
	Mixed	0.001141	10.42**
Channel 5	Pinus	0.002664	1.58
	Araucaria	0.001674	9.97**
	Hardwoods	0.001429	3.81**
	Mixed	0.003330	4.14**
Channel 7	Pinus	0.000944	0.80
	Araucaria	0.002592	6.15**
	Hardwoods	0.002836	9.22**
	Mixed	0.001683	1.74
Color Composite	Pinus	0.002074	7.99**
	Araucaria	0.002473	6.36**
	Hardwoods	0.001369	10.47**
	Mixed	0.004746	1.73

* - significant at 0.05 level (95%).
** - significant at 0.01 level (99%).

[a]degrees of freedom for test: numerator - 4
denominator - 120

hypothesis states that the elements of the standardized vector variable Z are already uncorrelated, and that the observed correlations in R differ from zero only by chance. If so, the population shape in the space of the standardized variables is spherical, thus the name of the test. For any equi-density surface in a mapping of points in a space of uncorrelated variables of unit variance, the shape is a sphere (Anderson, 1958). If we let P stand for the population correlation matrix, then for standardized variables in matrix Z, $|P|$ is the generalized variance and $|R|$ is its estimator. What Bartlett has devised is an approximate chi-square test of Ho: $|P| = 1$. Of course, $|P| = 1$ only if $P = I$. Results of the test are shown in Table 6. Without significant test results, interest in accounting for the relation among these four variables (the four moments) would normally cease, since we cannot reject the null hypothesis that $P = I$. In this case, we continued the analysis.

Table 6: Sphericity test (Bartlett's test criterion)

Scale	Chi-square
1:40,000	13.60*
Channel 7	10.53*
Color Composite	31.48*
Ratio 6	42.84*
Ratio 9	25.43*

* - significance at 0.10 level (90%) based on 6 degrees of freedom.

In the next step (Table 7) the percent of variation accounted by each principal component was examined, to determine the dimensions of the data space. Most of the variation between densities resides in the first three linear combinations or principal components, indicating a three-dimensional space. That is, variations in density values can be explained by means, standard deviations, and skewness coefficients. The other linear combination, kurtosis, is relatively constant from one density to the next; hence further study would explain little of density variation between the forest types. Table 8 gives the correlation of the variables with the two first principal components (loadings). In the 1:40,000 scale, the mean and standard deviation have the higher correlations. This indicates that the raw data values do not require any manipulation or transformation. The same thing happened in channel 7; however the skewness has a smaller correlation than the mean. But, when the data is transformed, as in the color composite and ratios, the higher moments have larger correlations; the lower moments (mean and standard deviation) were therefore used in the second principal component.

Until now, only the internal structure of the variables were described or considered. The main objective was to study group differences (forest types) in a multidimensional measurement space. That is, to study the effectiveness of density measurements in the separation of the forest types. The distinctive multivariate nature of MANOVA is that the dependent variable is in vector form. This dependent vector (variable) is assumed to be multivariate normal in distribution with the same dispersion, or variance-covariance matrix, for each population. The purpose of the MANOVA is to test for differences among population

Table 7: Percent of variation accounted by each principal component

Scale	Principal Component	Eigenvalue	Percent of trace (%)	Cumulative Percent (%)
1:40,000	1	1.693714	42.34	42.34
	2	1.209908	30.25	72.59
	3	0.749285	18.43	91.32
	4	0.347093	8.68	100.00
Channel 7	1	1.614126	40.35	40.35
	2	1.200265	30.01	70.36
	3	0.786391	19.66	90.02
	4	0.399218	9.98	100.00
Color Composite	1	1.863771	46.59	46.59
	2	1.010261	25.26	71.85
	3	1.006755	25.17	97.02
	4	0.119213	2.98	100.00
Ratio 6	1	2.061248	51.53	51.53
	2	1.054201	26.35	77.88
	3	0.809411	20.23	98.11
	4	0.075140	1.89	100.00
Ratio 9	1	1.813308	45.33	45.33
	2	1.302945	32.57	77.90
	3	0.702530	17.45	95.46
	4	0.181217	4.54	100.00

Table 8. Correlation of variables with the two first principal components (loadings).

Scale	Variables	Principal Component 1	Principal Component 2
1:40,000	Mean	0.7609	-0.1413
	Standard Deviation	-0.7413	-0.5498
	Skewness	-0.2088	0.9307
	Kurtosis	0.7222	-0.1463
Channel 7	Mean	0.8909	-0.0101
	Standard Deviation	-0.3425	0.7549
	Skewness	0.8074	0.0885
	Kurtosis	-0.2258	0.7889
Color Composite	Mean	0.0756	0.6189
	Standard Deviation	-0.0452	0.7916
	Skewness	0.9586	0.0079
	Kurtosis	0.9680	-0.0192
Ratio 6	Mean	0.5797	0.4954
	Standard Deviation	0.3064	0.7908
	Skewness	0.9485	-0.2366
	Kurtosis	0.8553	-0.3566
Ratio 9	Mean	0.7288	-0.5237
	Standard Deviation	-0.1710	0.7934
	Skewness	0.9364	0.1510
	Kurtosis	0.6130	0.6134

centroids or mean vectors. Significant results would indicate that the populations are centered at different locations in the measurement space spanned by the dependent vector variable.

Let H_2: $\mu_k = \mu$ be the null hypothesis that forest types have a common mean vector. This test may be viewed as a test of the effects of the treatments represented by the groups. In applying the test, significant results were found in all fifteen scales. This led to the conclusion that there was a difference between the group centroids, but it was not known which groups were different. This problem was examined through discriminant analysis.

The Lambda test of the H_2 hypothesis assumes that the four dispersion matrices are based on samples of four multivariate normal population with the same dispersion matrices. A test criterion for the null hypothesis (H_1) of the equality of the four group dispersion matrices, extended from a development of Bartlett's (1937), was presented by Box (1949). When this test was applied, significant results were found for all 15 scales. This indicated that the four dispersion matrices were different, that they did not have the same shape and form in multidimensional space. In this case, a quadratic discriminant function should be used, but because of the lack of quadratic discriminant algorithms and because the MANOVA H_2 test is fairly robust to unequal dispersion matrices, quadratic forms were not used. We proceeded with the analysis using linear discriminant functions.

The first procedure in discriminant analysis is to apply the overall test for discrimination of the four forest types. The test, Wilk's Lambda criterion, was presented by Wilks (1932) and is based on the discriminatory power between the group centroids. When the test was applied for all fifteen scales, only ratio 1 and ratio 3 did not produce significant results, indicating that the ratios of the spectral regions from channels 4 and 5 were not appropriate for densitometric analysis of forest types.

Before any discriminant function was derived, the test for the discriminatory power that exists in the variables being used was applied. This procedure tests for the ability of the variables to individually distinguish between group means, and was accomplished by a one-way analysis of variance (Table 9). The test is based on F-ratio with its associated lambda. The larger the lambda is, the less discriminatory power is present. For the five scales in Table 9, only the mean and standard deviation were each able to individually distinguish between the group means. In addition, for ratio 9, skewness was able to distinguish between group means. But, the above test does not tell which group or groups are different.

Associated with this test, that is, based on the two variables selected (mean and standard deviation) by the stepwise procedure used, is the test for significance of the Mahalanobis distance between groups (forest types). Thus, a matrix of F-ratios is generated under the assumptions of random sampling and multivariate normal distributions. Within each scale or imagery level, the 4 forest types were compared to determine significant differences between their mean vectors. These comparisons, composed of six total pairs, are shown in Table 10.

Table 9: One-way analysis of variance for equality of group centroids on a single discriminating variable.

Scale	Wilks Lambda(A) / F-Ratio (B)[a]	Mean	Standard Deviation	Skewness	Kurtosis
1:40,000	A	.35	.15	.85	.77
	B	9.75**	29.49**	.94	1.56
Channel 7	A	.03	.23	.70	.82
	B	164.87**	17.67**	2.24	1.14
Color Composite	A	.03	.36	.75	.79
	B	142.06**	9.16**	1.76	1.36
Ratio 6	A	.03	.41	.76	.90
	B	146.71**	7.43**	1.68	.55
Ratio 9	A	.03	.24	.56	.91
	B	128.67**	16.25**	4.15**	.48

[a]F-table value for the F-ratio with 3 and 16 degrees of freedom.

** - significant at .01 level (99%).

Table 10: Mahalanobis distance between forest types based on the discriminant power of mean and standard deviation variables (F-ratio[a] matrix for each pair of forest types to test equality of mean vectors).

Scale		Type 1[b]	Type 2	Type 3
1:40,000	Type 2	13.42**		
	Type 3	35.32**	7.53**	
	Type 4	37.49**	6.64**	8.78**
Channel 7	Type 2	67.49**		
	Type 3	89.54**	241.12**	
	Type 4	6.44**	36.34**	101.91**
Color Composite	Type 2	7.42**		
	Type 3	147.11**	219.99**	
	Type 4	167.07**	239.02**	12.10**
Ratio 6	Type 2	52.80**		
	Type 3	68.92**	216.21**	
	Type 4	9.36**	28.11**	88.43**
Ratio 9	Type 2	54.91**		
	Type 3	52.82**	173.24**	
	Type 4	18.91**	14.14**	88.60**

[a]F-table value for the F-ratio with 2 and 15 degrees of freedom.

[b]Forest type 1 - Pinus
Forest type 2 - Araucaria
Forest type 3 - Hardwoods
Forest type 4 - Mixed
** - Significant at .01 level (99%).

There were significant differences of mean vectors for all groups for all five scales. This indicates that each group has a unique position in the multidimensional space. The importance of this statistic is shown where the discriminant functions were derived in the following pages. This table can be related to Figures 6, 7, 8, 9, and 10 in the sense of quantifying the distributions.

"To determine the number of discriminant functions, the maximum number is one less the number of groups or equal to the number of discriminant variables. The maximum number of dimensions needed to completely describe a set of points is one less than the number of points. In discriminant analysis, each group (as measured by its centroid) was treated as a point and each discriminant function was an orthogonal[21] dimension describing the location of that group relative to the others" (SPSS, page 442).

Table 11 reports the information necessary for selecting the number of discriminant functions to be derived. "The first criterion for selecting the number of discriminant functions was the relative percentage of the eigenvalue associated with the function. It is a measure of the relative importance of the function, as the sum of the eigenvalues is a measure of the total variance existing in the discriminant variables. When it is expressed as a percentage of the total sum of eigenvalues, the eigenvalue provides a measure of the relative importance of the associated function. Since discriminant functions were derived in the order of their importance, the process could be

---

[21]orthogonal: axes are right angle to each other.

Table 11: Determination of the number of discriminant functions by stepwise procedure with their associated statistics.

Scale	D.F.[a]	Eigen-value	Canonical Correlation	Percent of trace (%)	Wilks Lambda	Chi-square	Degrees of Freedom
1:40,000	1	6.5071	.93	83.6	.055	43.24**	12
	2	1.1861	.73	15.2	.420	13.01*	6
	3	.0891	.28	1.1	.918	1.27	2
Channel 7	1	34.6640	.98	89.0	.005	79.32**	12
	2	4.2301	.89	10.9	.180	25.71**	6
	3	.0612	.24	.2	.942	.89	2
Color Composite	1	56.2121	.99	96.1	.004	80.46**	12
	2	2.0353	.81	3.5	.267	19.76**	6
	3	.2301	.43	.4	.812	3.11	2
Ratio 6	1	32.1294	.98	92.4	.006	75.72**	12
	2	2.1441	.82	6.2	.212	23.22**	6
	3	.4958	.57	1.4	.668	6.04	2
Ratio 9	1	27.5882	.98	89.5	.007	72.91**	12
	2	3.1698	.87	10.3	.221	22.61**	6
	3	.0829	.27	.3	.923	1.19	2

[a] D.F. - discriminant function

** - significant at .01 level (99%)

* - significant at .05 level (95%)

stopped whenever the relative percentage of the new function was judged to be too small" (SPSS, page 442). Thus, in Table 11, the third eigenvalue is quite small, hence the third discriminant function was disregarded.

The second criterion was associated with the canonical correlations. "Canonical correlation is a measure of association between the single discriminant function and the set of (g-1) dummy variables which define the "g" group membership. It indicates how closely the function and the group variable are related, providing another measure of the function's ability to discriminate among groups. The squared canonical correlation can be interpreted as the proportion of variance in the discriminant function explained by the groups" (SPSS, page 442). The first two discriminant functions were highly correlated with the groups, but the third function had a very low correlation (Table 11). Thus, the third discriminant function was eliminated under this second criterion.

"The third criterion for eliminating discriminant functions was to test for the statistical significance of discriminating information not already accounted for by earlier functions. Wilks' Lambda (Wilks, 1932) was computed as each function was derived. Lambda is an inverse measure of the discriminating power in the original variables; the larger lambda is, the less information remaining in the data. Lambda values can be transformed into chi-square statistics for an easy test of statistical significance. From Table 11, it was clear that the third function had little discriminating power. It would not be useful to derive the third (and last) discriminant function, since it would not significantly add to our ability to discriminate between the forest types" (SPSS, page 442). Consequently, the two first functions were quite adequate for describing the four forest types.

Of course, the two first functions did not completely utilize all of the information in the discriminating variables. But, as the remaining information was not statistically significant, the third function was ignored. This is often the case, the last function may not mathematically disappear, due to sampling and measurement errors, even though it did not actually exist as a separate dimension in the population.

Table 12 gives the discriminant functions derived from Table 11, with their standardized coefficients. Notice that only the mean and standard deviation were selected as variables. They were used to compute the discriminant scores for the standardized variables: the sum of the product of each variable was multiplied by its corresponding coefficient. The discriminant scores were used to produce a graphical representation of the clustering produced by a discriminant function. There will be a separate score for each observation for each forest type on each function. The discriminant scores produced were in standard form - over all observations in the analysis, the score from one function will have a mean of zero and standard deviation of one. Thus, any single score represents the number of standard deviations that each observations is away from the mean for all observations on the given discriminant function. If there are several discriminant

Table 12: Discriminant functions and number of variables in each function as selected by stepwise procedure.

Scale	Standardized discriminating function coefficients (D_i)[a]	Percent of separation[b]
1:40,000	$D_1 = -.52507\ Z_1 + 2.18655\ Z_2$ $D_2 = 1.45164\ Z_1 + .84966\ Z_2$	85.00
Channel 7	$D_1 = 5.45278\ Z_1 + .52705\ Z_2$ $D_2 = .29471\ Z_1 + 1.93716\ Z_2$	100.00
Color Composite	$D_1 = 6.49041\ Z_1 - 1.43490\ Z_2$ $D_2 = .30574\ Z_1 + 1.44722\ Z_2$	100.00
Ratio 6	$D_1 = 5.15142\ Z_1 - .40466\ Z_2$ $D_2 = -.21819\ Z_1 + 1.43821\ Z_2$	100.00
Ratio 9	$D_1 = 4.55198\ Z_1 - .28615\ Z_2$ $D_2 = .66398\ Z_1 + 1.82388\ Z_2$	100.00

[a] Z_1 - standardized mean

Z_2 - standardized standard deviation

[b] percent of separation between the forest types.

functions, each observation will have a score for each function.

The group mean for a specific discriminant function was calculated by averaging the scores for all observations within a particular group (forest type). "For a single group, the means on all the functions are referred to as the group centroid, the most typical location of an observation from that group in the discriminant function space. A comparison of the group means on each function tells how the groups are spread along that dimension. But, more importantly, the functions were arranged in order of decreasing importance, so that a given difference between group means on the third function is not as meaningful as the same difference on the first function" (SPSS, page 443).

Figures 11, 12, 13, 14, and 15 show the scores of the groups plotted using the first and second discriminant functions. The boundary lines delineate the region of each group defined by the linear discriminant functions. Examination of the scores was particularly useful in studying the separation of the group centroids and their relative locations. In Figure 11, we see that Pinus and Mixed were very well defined, but Araucaria and Hardwoods had some misclassification. In Figures 12, 14, and 15 the groups fell into the same positions. Hence, since channel 7 alone gave the same results as ratio 6 and ratio 9 (respectively, Figures 14 and 15), we could discard these two ratios.

Figure 13, the color composite discrimination, showed the best classification. Pinus and Araucaria grouped closely, as did Hardwoods and Mixed. And, as important, the two "pairs" were well separated.

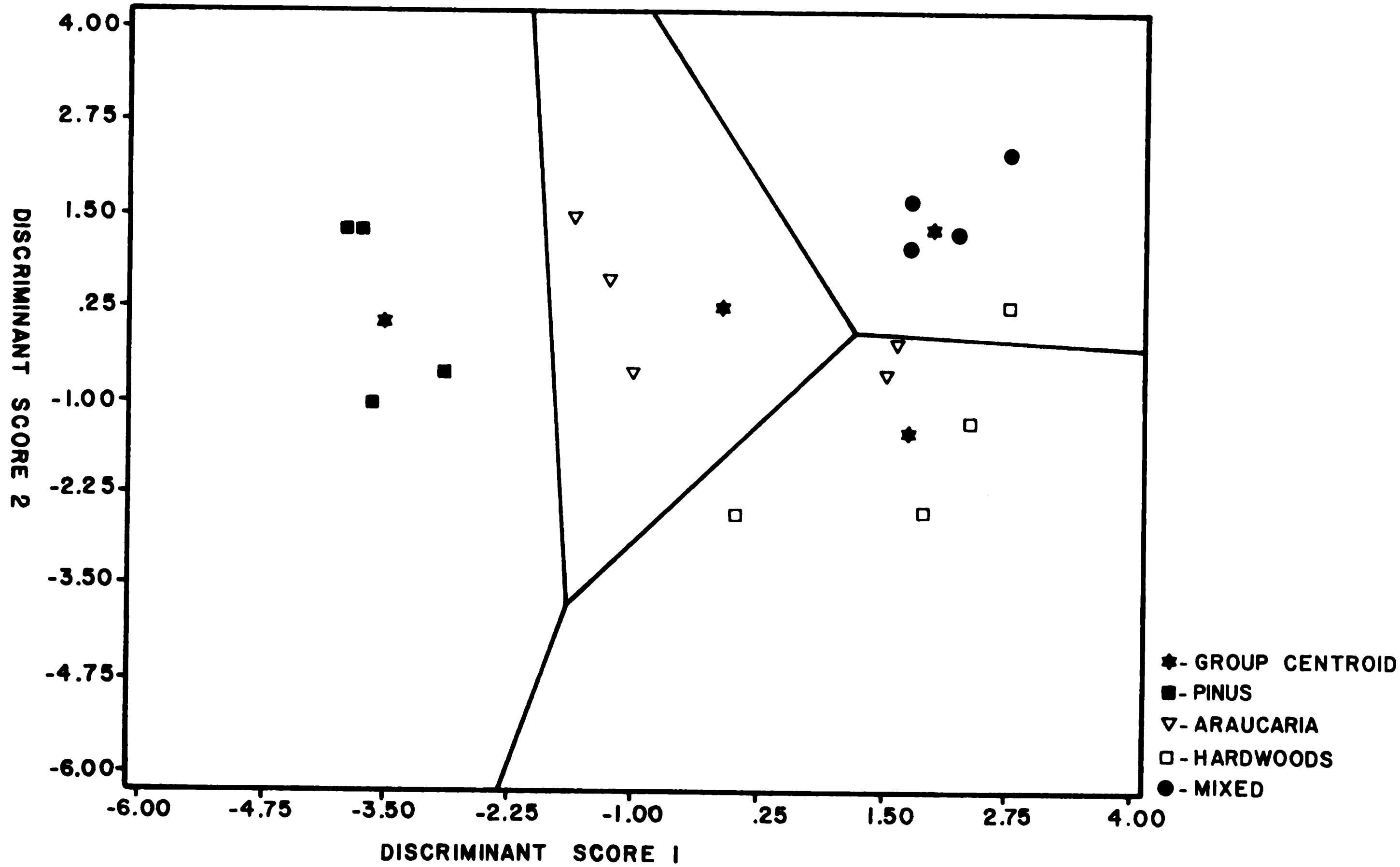


Figure 11: Plot of discriminant scores (group means) for 1:40,000 scale.

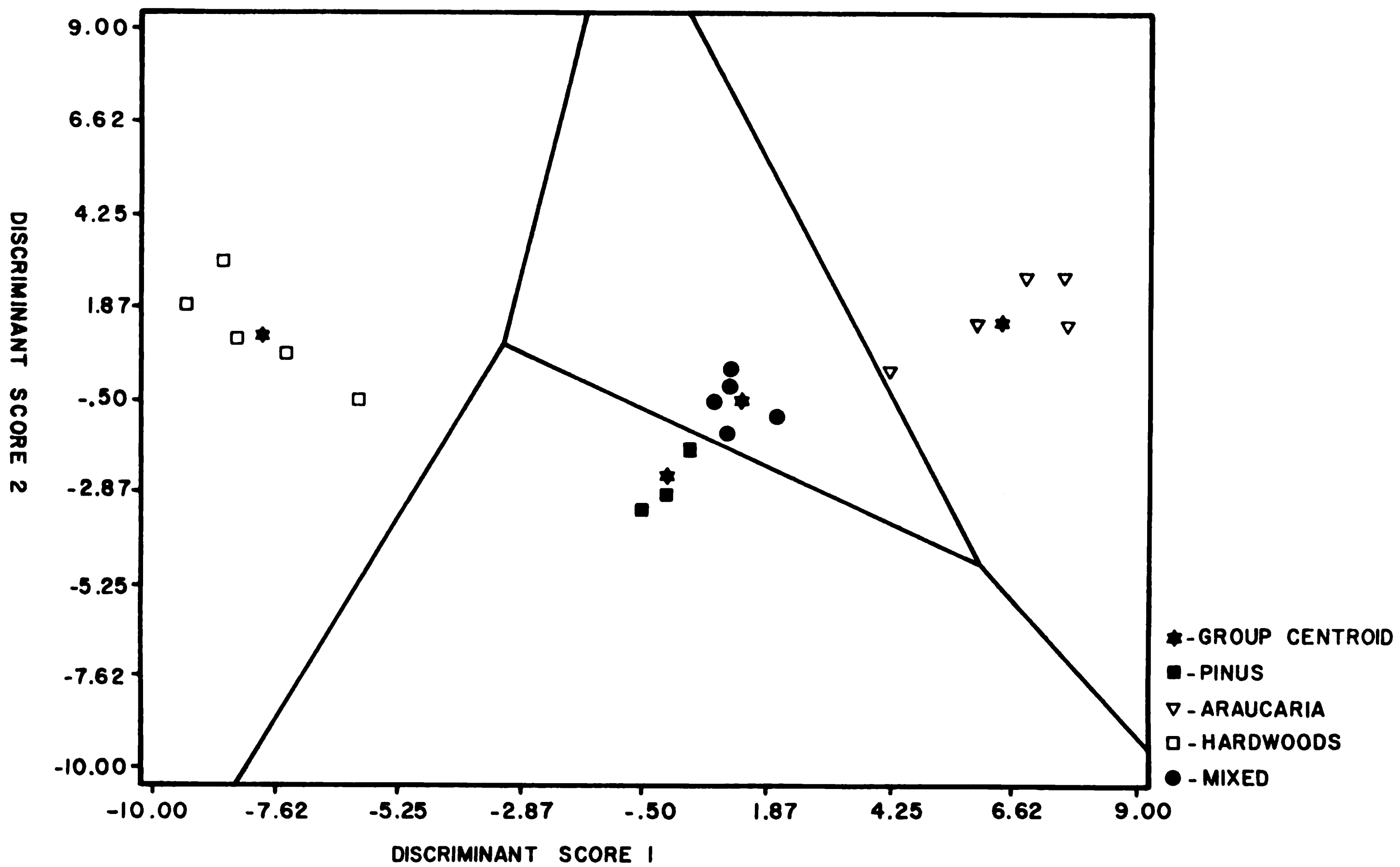


Figure 12: Plot of discriminant scores (group means) for channel 7.

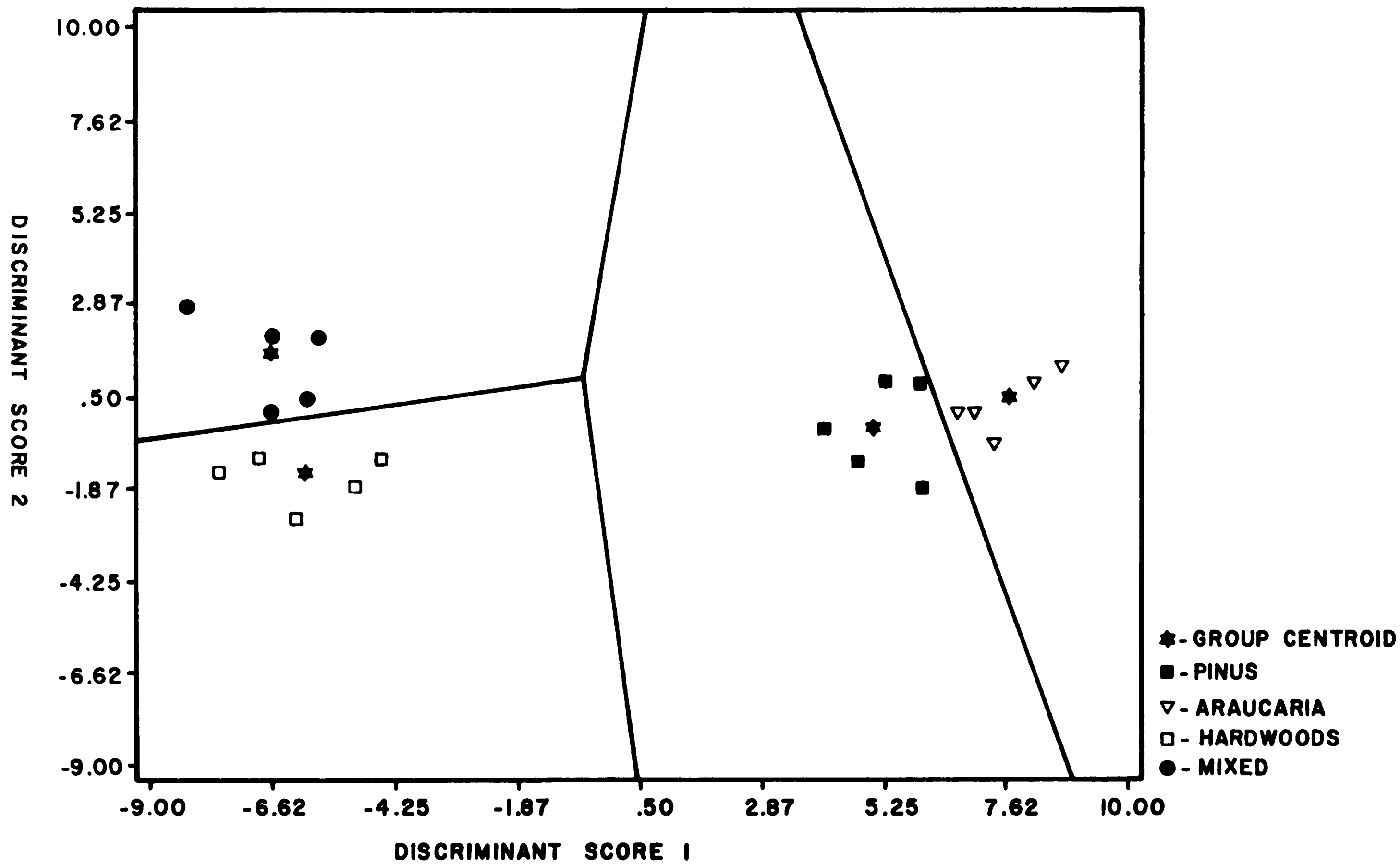


Figure 13: Plot of discriminant scores (group means) for color composite.

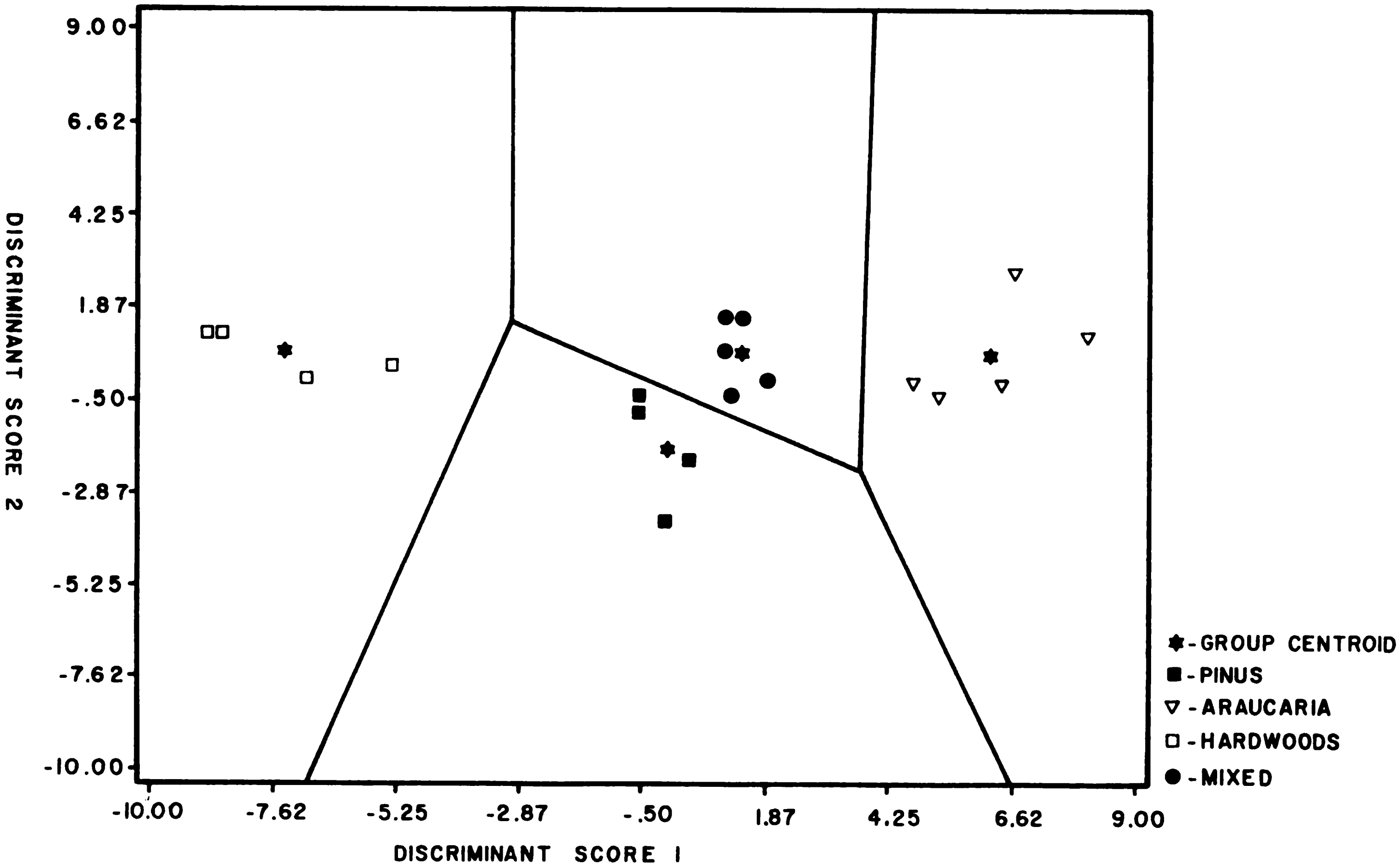


Figure 14: Plot of discriminant scores (group means) for ratio 6.

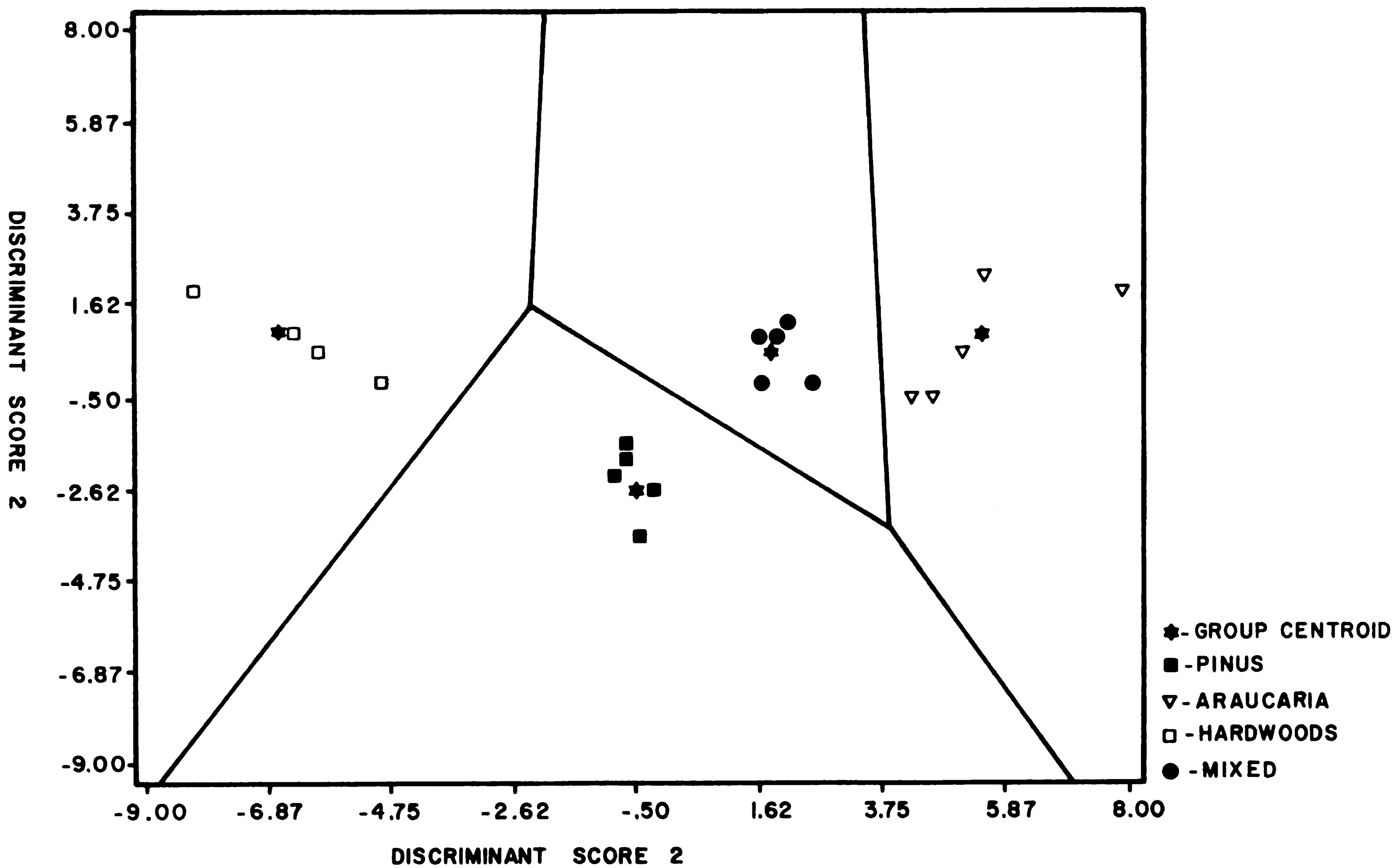


Figure 15: Plot of discriminant scores (group means) for ratio 9.

This is the most desired situation, because Pinus and Araucaria are conifers and the other two groups are deciduous. Even with Pinus and Araucaria at one extreme and Hardwoods and Mixed at the other the separation of the close groups was perfect. Therefore, the color composite is the best image for densitometric analysis of canopy reflectance for forest types.

Table 13 shows a classification of the information for each group for each scale. This complemented the information provided by Figures 11, 12, 13, 14, and 15. Also, Table 13 gives the percent of correct classification for each forest type. In Table 12, the percent of separation corresponds to the overall average of correct classification in each scale between the four forest types. At the 1:40,000 scale, some misclassification occurred: two groups of Araucaria were classified as Hardwoods and one group of Hardwoods was classified as Mixed. This probably was caused by tone variations from sample to sample, pattern variations, sun elevation, terrain slope, etc., as explained previously. The other four scales achieved 100 percent classification.

Normally, cluster analysis would be done prior to discriminant analysis. The purpose of applying cluster analysis after discriminant analysis was to show and describe how the technique should be carried out and its effectiveness. A second purpose was to discover which cluster algorithm best fitted densitometric analysis.

Table 13: Classification information and prediction results for each scale based on five cases.

Scale	Group	Predicted Pinus (%)	Predicted Araucaria (%)	Predicted Hardwoods (%)	Predicted Mixed (%)
1:40,000	Pinus	100.0	0	0	0
	Araucaria	0	60.0	40.0	0
	Hardwoods	0	0	80.0	20.0
	Mixed	0	0	0	100.0
Channel 7	Pinus	100.0	0	0	0
	Araucaria	0	100.0	0	0
	Hardwoods	0	0	100.0	0
	Mixed	0	0	0	100.0
Color Composite	Pinus	100.0	0	0	0
	Araucaria	0	100.0	0	0
	Hardwoods	0	0	100.0	0
	Mixed	0	0	0	100.0
Ratio 6	Pinus	100.0	0	0	0
	Araucaria	0	100.0	0	0
	Hardwoods	0	0	100.0	0
	Mixed	0	0	0	100.0
Ratio 9	Pinus	100.0	0	0	0
	Araucaria	0	100.0	0	0
	Hardwoods	0	0	100.0	0
	Mixed	0	0	0	100.0

The analysis started by standardizing the variables to a zero mean and unit variance. Euclidean distance was then used as a measure of dissimilarity. From the six algorithms used for hierarchical clustering, Ward's method (based on error sum of squares as an objective function) gave the best results. Figures 16, 17, 18, 19, and 20 show the five resulting dendrograms for the five scales. In Figure 16, three clusters were defined. In cluster 1, Hardwoods, one Araucaria group was included. This follows from the discrimination shown in Figure 11: in the Hardwood region one Araucaria group was closer to the group centroid. In the second cluster (Figure 16), one Araucaria and one Hardwood group were included in the Mixed group. Looking again at Figure 11, we see that in the Mixed group one Hardwood and one Araucaria group were located near the group centroid. In the third cluster (Figure 16), the conifers clustered together. Thus, one reason to run discriminant analysis after cluster analysis would be to discriminate Pinus from Araucaria, and to check the validity of the clusters.

In Figure 17, four clusters were defined. Cluster 4 was considered an outlier, that is, not belonging to any group. This is because Ward's method is based on error sum of squares and observations in cluster 4 had an error sum of squares too large in relation to each of their group centroids, which caused them to form one separate cluster. Again, referring to Figure 12, we see that the Mixed and Pinus groups clustered together but still were well discriminated. In Figure 17, cluster 1 was defined by the Pinus and Mixed groups, cluster 2 defined as Araucaria and cluster 3 as Hardwoods.

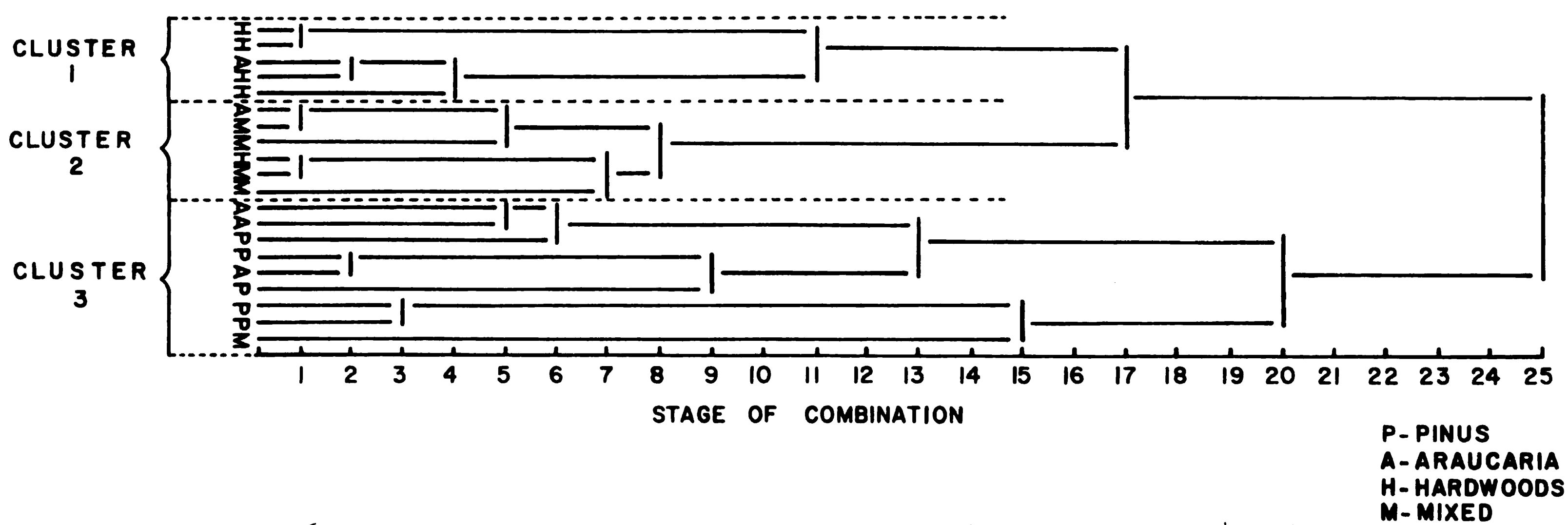


Figure 16: Hierarchical clustering based on Ward's method for 1:40,000 scale.

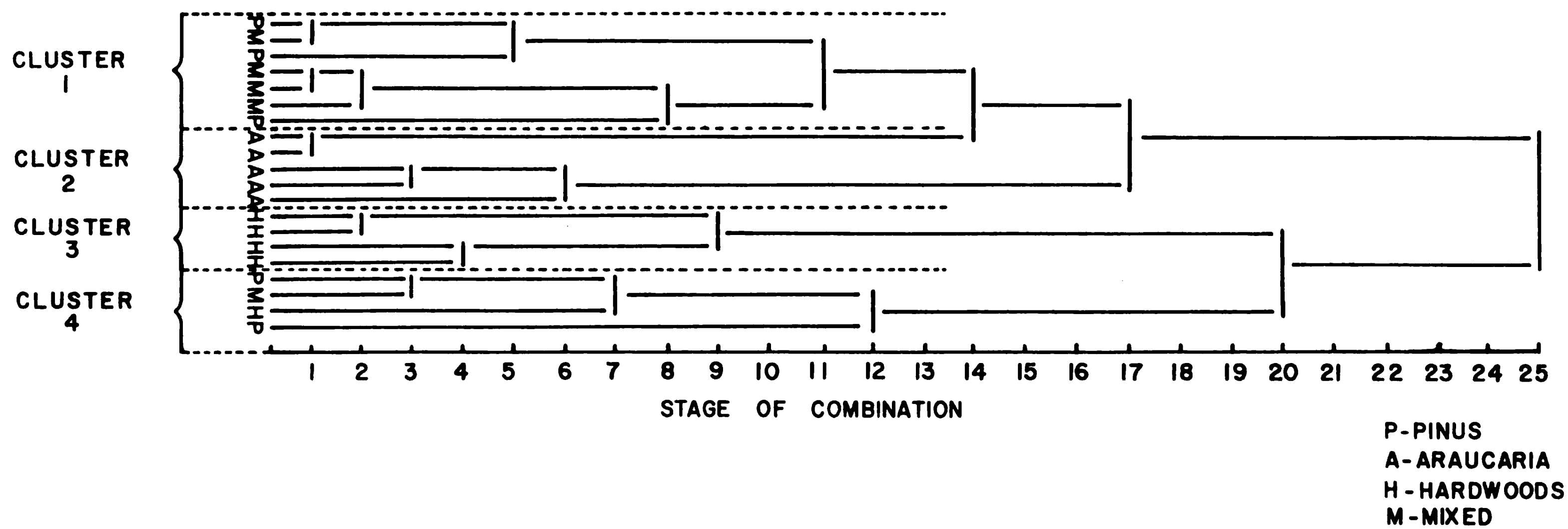


Figure 17: Hierarchical clustering based on Ward's method for channel 7.

Without doubt, the color composite again showed the best results (Figure 18). Cluster 1 consisted of the conifer groups, Pinus and Araucaria, cluster 2 contained Hardwoods, and cluster 3 included Mixed. These were same results as in Figure 13.

As shown in Figures 19 and 20, both ratios produced the same results. Cluster 1 consisted of Mixed and Araucaria groups, cluster 2 consisted of the Pinus group, and cluster 3 consisted of the Hardwoods group. Since in cluster 1 Mixed and Aracaria grouped together, the color composite (Figure 18), again proved to be the best to discriminate between the four forest types. The association between cluster analysis and discriminant analysis should now be clear. Discriminant analysis checks for the validity of the clusters formed by clusters analysis. Even though the two analyses have completely different approaches and purposes, they have the same objective in achieving a classification and discrimination of the densitometric measurements.

Since hierarchical clustering methods do not use the internal structure of the groups formed, a non-hierarchical method was used to study the internal structure of the groups by allowing the groups to overlap. The approach employed was based on a squared criterion and searched for clusters with the smallest squared-error. Several statistics were applied to an analysis of variance for each feature (dimension). One should note that the features used are the four statistical moments. Hence, one can ask whether the clusters were significantly different in a particular feature. Table 14 shows the analysis of variance for each feature in each of the five scales.

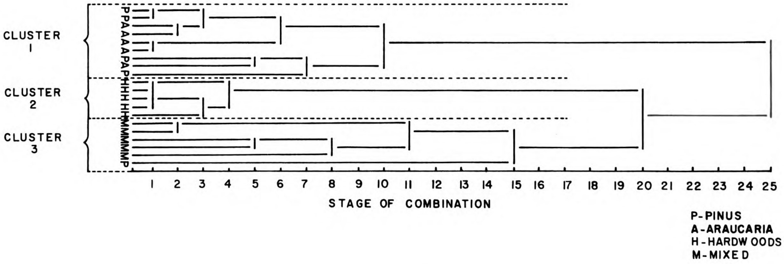


Figure 18: Hierarchical clustering based on Ward's method for color composite.

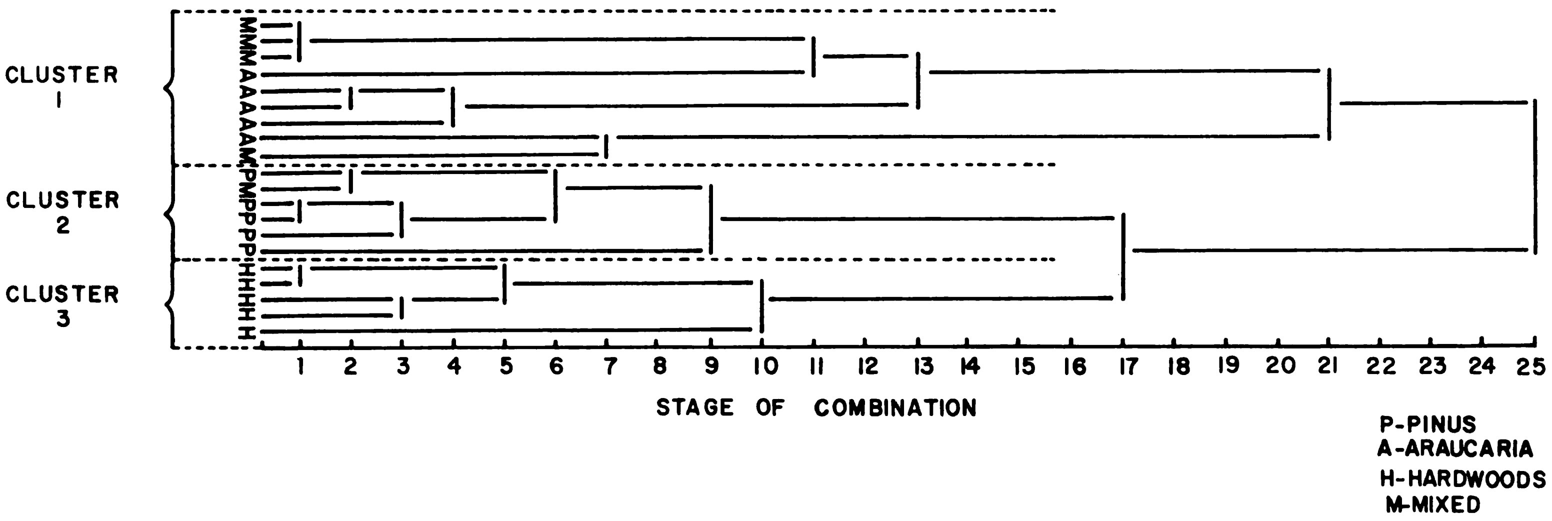


Figure 19: Hierarchical clustering based on Ward's method for ratio 6.

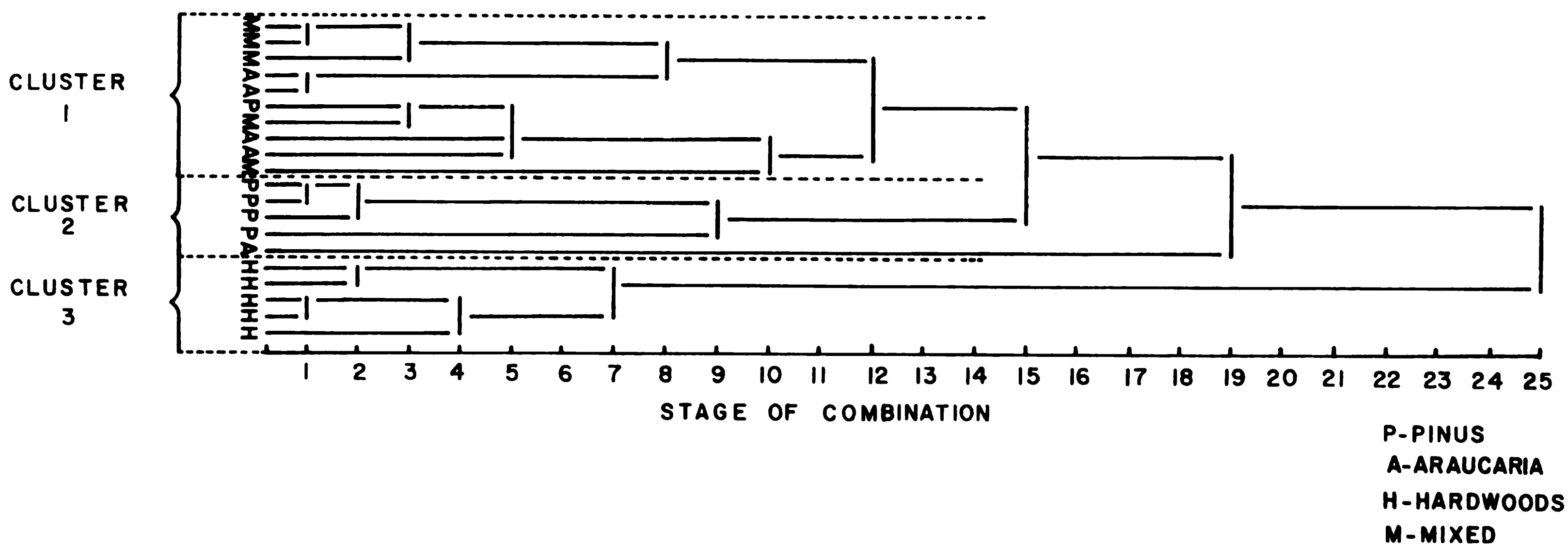


Figure 20: Hierarchical clustering based on Ward's method for ratio 9.

Table 14: Analysis of variance of each feature (dimension in the clustering.

Scale	Dimension (Features)[a]	Squared Error	Between Clusters	F-ratio[b]	Critical Alpha
1:40,000	1	5.65	14.34	13.53**	.075
	2	6.71	13.28	10.54**	.288
	3	12.71	7.28	3.05	.575
	4	5.84	14.15	12.92**	.031
Channel 7	1	2.94	17.05	30.82**	.004
	2	11.76	8.23	3.73*	.947
	3	7.18	12.81	9.50**	.663
	4	7.99	12.00	8.00**	.395
Color Composite	1	1.05	18.94	96.25**	.718
	2	7.02	12.97	9.85**	.258
	3	6.02	13.97	12.36**	.003
	4	2.48	17.51	37.52**	.000
Ratio 6	1	3.09	16.90	29.10**	.001
	2	7.35	12.64	9.17**	.003
	3	6.48	13.51	11.11**	.001
	4	7.11	12.88	9.65**	.009
Ratio 9	1	3.25	16.74	27.46**	.060
	2	4.99	15.00	16.00**	.033
	3	4.15	15.84	20.32**	.062
	4	4.85	15.14	16.64**	.000

[a] Features: 1 - mean
2 - standard deviation
3 - skewness
4 - kurtosis

[b] F-Ratio with 3 and 16 degrees of freedom

* - significance at .05 level (95%)

** - significance at .01 level (99%)

The analysis of variance tested whether the between-clusters squared error (the part of the total error due to differences in the clusters centers) was significantly larger than the squared error (part of the total error due to effects) - if the variances for all clusters were the same. The null hypothesis was that there was no effect due to the clusters (squared error was due entirely to random effects). From Table 14, it can be seen that the F-Ratio was significant for almost all features in all scales, except for feature 3 (skewness) in scale 1:40,000. Since the F-Ratio is not very robust, the critical value of alpha (CRITICAL ALPHA in Table 14) played an important role in checking the validity that variances for all clusters were the same. A large value (> 0.2) of critical alpha infers that there was no effect due to the clusters. In Table 14, even though we had significant F-ratios for almost all features in all scales, this did not occur with the CRITICAL ALPHA.

In the 1:40,000 scale, standard deviation and skewness had values greater than 0.2 In channel 7, only the mean had a value smaller than 0.2. In color composite, mean and standard deviation had values greater than 0.2. In ratios 6 and 9, all the features had values smaller than 0.2. Table 15 shows the classification information based on the above results. Thus, a listing of the group means from each forest type that ended up in each cluster formed was provided. In 1:40,000 scale, considerable misclassification occurred for Pinus, Araucaria and Hardwoods, although the Mixed group was reasonably well classified.

Table 15: Listing of the group means from each forest type that ended up in each cluster.

Scale	Cluster	Forest Types: Pinus	Araucaria	Hardwoods	Mixed
1:40,000	1	0	2	4	0
	2	4	3	0	0
	3	1	0	0	1
	4	0	0	1	4
Channel 7	1	1	3	0	2
	2	0	0	5	0
	3	2	0	0	3
	4	2	2	0	0
Color Composite	1	0	0	5	0
	2	1	0	0	0
	3	0	0	0	5
	4	4	5	0	0
Ratio 6	1	0	0	5	0
	2	0	1	0	1
	3	0	4	0	4
	4	5	0	0	0
Ratio 9	1	0	0	5	0
	2	0	1	0	0
	3	0	3	0	5
	4	5	1	0	0

In channel 7, only the Hardwoods group was correctly classified. The reason for this misclassification in these two scales was due to the fact that the critical alpha was greater than 0.2 for skewness in 1:40,000 scale and greater than 0.2 for standard deviation, skewness, and kurtosis in channel 7. The two ratios, 6 and 9, both had the same misclassification, grouping Araucaria and Mixed in cluster 3. This was due to the fact that the critical alpha was smaller than 0.2 in all four variables. The best classification was achieved in the color composite. Pinus and Araucaria were classified in cluster 4, Hardwoods in cluster 1, and Mixed in cluster 3. Cluster 2, with one Pinus, was considered as an outlier (same situation as occurred in Figure 17). The reason for this good classification was because of a critical alpha greater than 0.2 for the two first variables, especially for the mean (see Table 14), and a critical alpha smaller than 0.2 for skewness and kurtosis. From this we concluded that the squared errors in color composite were due to random effects.

In comparing Table 15 to Figures 16, 17, 18, 19, and 20, we noticed that classification results were the same only in the color composite. Both the methods comployed in Figure 18 and the one in Table 15 were completely different in structure, but both based on squared error. Therefore, equality of results means that the color composite had a minimum squared error within clusters and a larger squared error between clusters, a most desirable situation. This also means that the clusters are compact and isolated.

From the cluster analyses, we concluded that the mean and standard deviation variables should be used for forest classification based on densitometric measurements. Skewness and kurtosis variables did contribute only for misclassification of the four forest types. Thus, the color composite was again selected for forest classification based on densitometric measurements.

CHAPTER V

SUMMARY AND CONCLUSIONS

This analysis of forest canopy reflectance focused on the State of Parana, located in southern Brazil. Since its forest lands have undergone and are continuing to undergo rather drastic changes, resource planners must have accurate data for effective management. This requires the quantification of available resources (the resource base), current conditions, and changes over time. In this sense a photographic record would provide a rapid and accurate means of acquiring information needed for management planning and operation. Several alternatives were evaluated, from aerial photos through satellite images.

To inventory forest canopies requires a knowledge of canopy structure as well as growth, phenologic characteristics, and how they affect remote sensor signals. One way to investigate such factors is through mathematical models describing the interactions of these various characteristics.

To analyze forest canopy reflectance, analog analysis was employed since it requires a low level of hardware. The densitometer, an instrument that measures density of aerial photos or multispectral scanner imagery, was used in the study.

The objective of this study was to develop a methodology to distinguish major forest types in Brazil using aerial and space imagery. The forest types were Pinus elliottii and Pinus taeda, Araucaria angustifolia (Bert.) O. Ktze., Hardwoods and Mixed. The potential of identifying these forest types from low-level photography through satellite imagery was evaluated, as well as the potential usefulness for forest classification of signatures of various forest canopy components.

The approach used to discriminate between the forest types was based on texture-type separation of density measurements. The model developed uses as variables the four statistical moments: (1) the arithmetic mean of grey density variations; (2) the mean quadratic deviation of grey density variations from the arithmetic mean of grey density; (3) the moment coefficient of skewness of grey density variations from the arithmetic mean of grey density; and (4) the moment coefficient of kurtosis of grey density variations from the arithmetic mean of grey density. The model was used to discriminate between forest types by principal components analysis, multivariate analysis of variance, discriminant analysis, and cluster analysis of canopy reflectance characteristics. Spectral reflectance (density values) for forest types between 0.4 micrometers and 1.1 micrometers were used. The hypotheses for the analysis were that the spectral and spatial variations of the forest canopies signatures (reflectance) would cluster separately for each scale, and that the optical densities of the forest canopies signatures were sufficiently discrete to allow discrimination between them.

Forest types were first photointerpreted in the 1:40,000 scale, and then the same forest types transposed to the other levels of imagery. Density measurements were then taken with the densitometer at each level of imagery and within each forest type using random sampling procedures. Five areas were chosen, and a total of 125 density measurements were taken for each forest type in each scale of imagery. From each set of 25 measurements considered as one sample, the variables were calculated, resulting in 5 mean observation values per forest type per scale.

From the analyses, the following conclusions were reached:

1. Even with significant differences between samples within each forest type, excellent discrimination between the forest types was achieved.

2. LANDSAT MSS imagery proved to be useful for forest discrimination in color composite.

3. Density measurements are very effective for forest discrimination.

In discriminant analysis, only two discriminant functions were used for classification, both functions using the mean and standard deviation as variables. Skewness and kurtosis did not contribute to discrimination. Also, in cluster analysis, these two variables accounted only for misclassification. Even with different dispersion matrices, perfect discrimination was achieved using linear discriminant function. This proved that linear functions are relatively robust for having well-separated group centroids.

In cluster analysis, success was achieved since the groups already defined were obtained when the methodology was applied. A minimum squared error was achieved for the color composite, which showed that the forest types formed compact and isolated clusters. Implications are that texture type separation of density measurements provided excellent discrimination between the forest types, and that the optical densities of the forest canopies were sufficiently discrete to allow discrimination between them.

Since success was achieved with this methodology, a suggested procedure is to begin with principal component analysis, followed by a cluster analysis, a MANOVA, and a discriminant analysis. This sequence provides a procedure for reducing the number of variables to be treated through principal component analysis. Those linear combinations which have small variances are discarded and only those combinations with large variances are studied.

Cluster analysis searches for natural groupings. When an image is scanned, there are no pre-defined groups. Cluster analysis will tell how those groups that occur are formed. From this point, the groups are labelled and discriminant analysis would be done to test for actual group separation. One example, in the study, is related to Figure 18 in cluster 1. Cluster 1 contains Pinus and Araucaria groups. Therefore, after finding the number of clusters, discriminant analysis separated well the Pinus from Araucaria (Figure 13).

From the fifteen images types studied (Table 13), the color composite proved to be the best in densitometric studies of forest canopy reflectance.

It is also suggested that further research in reflectance be directed at individual species, and not concentrate only on forest types. Also, it is highly recommended that digital analysis of LANDSAT MSS be conducted parallel to such a study to see the potential usefulness of these two methods.

The techniques will be readily applicable to monitor changes in forest composition in Brazil. Because they do not rely on computer technology, they should be valuable as analytical techniques in other countries without strong computer resources.

APPENDICES

APPENDIX A

INSTRUMENT FOR DATA COLLECTION

The densitometer used in this study basically combined analog computer circuitry and closed circuit television techniques to produce both color enhancement and/or edge enhancement of photographic transparencies and prints. The photographic image is illuminated by a light table or by incident light and scanned by the density scanner. The system then calculates the point density values or gradients over the entire portion of the image being examined. When viewed in black-and-white, the entire range of densities for a particular image can be seen as shades of gray. However, when the transformed image is viewed in color, the system takes the calculated density range and assigns 12 colors to it. In essence, it divides the given density range into 12 equal-interval classes where each color band represents an equal band of film density. From the color display, one should be able to make a reasonably accurate interpretation of the scanned image. The system also has a planimetric capability for area measurement.

The system used is available in the Department of Geography Densitometric Facility located at the Natural Science Building, Michigan State University.

APPENDIX B

ALGORITHMS AND COMPUTER PROGRAMS

The criteria by which independent variables are selected for inclusion in the discriminant analysis are indicated by the following methods. These methods are available in the Statistical Package for the Social Sciences (SPSS) Subprogram DISCRIMINANT.

1. All independent variables are entered into the analysis. The discriminant functions are created directly from the entire set of independent variables, regardless of the discriminatory power of each of the independent variables.

The alternative to this direct method is to use a stepwise selection method. Independent variables are selected for entry into the analysis on the basis of their discriminatory power. The process begins by chosing the single variable which has the highest F-value on the selection criterion. This initial variable is then paired with each of the other available variables, one at a time, and the selection criterion is computed. The new variable which in conjunction with the initial variables produces the best criterion value is selected as the second variable to enter the equation. These two are then combined with each of the remaining variables, one at a time, to form triplets which are evaluated on the criterion. The triplet with the best criterion value determines the third variable to be selected. This procedure of locating the next variable that would yield the best criterion score given the variables already selected, continues until all variables are selected or no additional variables provide a minimum level of improvement.

As variables are selected for inclusion, some variables previously selected may lose their discriminatory power. This occurs because the information that they contain about group differences is now available in some combination of the other included variables. Such variables are redundant and should be eliminated. Thus, at the beginning of each step, each of the previously selected variables is tested to determine if it still makes a sufficient contribution to discrimination. If any are eligible for removal, the least useful is eliminated. A variable which has been removed at one step may re-enter at a later step if it satisfies the selection criterion at that time. The five stepwise selection criteria are:

2. WILKS - The criterion used is the overall multivariate F-ratio for the test of differences among the group centroids. The variable which maximizes the F-ratio also minimizes Wilks' Lambda, a measure of group discrimination. This test takes into consideration the differences between all the centroids and the homogeneity within the groups;

3. MAHAL - This algorithm seeks to maximize the Mahalanobis distance[22] between the two closest groups;

4. MAXMINF - maximizes the smallest F-ratio between pairs of groups;

---

[22]Mahalanobis distance: $D^2 = (\bar{X}_1 - \bar{X}_2)' S^{-1} (\bar{X}_1 - \bar{X}_2)$ where $\bar{X}_1$ and $\bar{X}_2$ are mean samples of N_1 and N_2 and S^{-1} is the sampled covariance matrix from N.

5. MINRESID - criterion which tends to separate groups that are close together, which minimizes R, the residual variations;

6. RAO - is Rao's V, a generalized distance[23] measure. The variable selected is the one which contributes the largest increase in V when added to the previous variables. This amounts to the grestest overall separation of the groups. A variable which contains a large amount of information already included in the previously selected variables may actually cause a decrease in the value of V. This implies a decline in discriminatory power since the groups are being brought more closely together. One would not generally want to include such a variable. When there are a large number of cases, the change in V has a chi-square distribution with one degree of freedom so it can be tested for statistical significance.

In cluster analysis the problem is to find natural groupings in a data set, but in this sense we are obligated to define what we mean by a natural grouping. In what sense are we to say that the samples in one cluster are more like one another than like samples in other clusters? This question actually involves two separate issues - how should one measure the dissimilarity (or similarity) between samples, and how should one evaluate a partitioning of a set of samples into clusters?

---

[23] Rao's V statistic: $V = \Sigma^{p}_{i,\, j\, =\, 1} \Sigma\, S^{-1}\, \Sigma\, N_{\mu}(\bar{x}_{i\mu} - \bar{x}_{i})\,(\bar{x}_{j\mu} - \bar{x}_{j})$
where S^{-1} is the common covariance matrix, $N_1, N_2, \ldots, N_{\mu}$, the sample sizes, $\bar{x}_{i1}, \bar{x}_{i2}, \ldots, \bar{x}_{i\mu}$, the mean values of the ith character in the first, second, ..., μth, populations, and $\bar{x}_i = (\Sigma N_{\mu}\, \bar{x}_{i\mu}) / (\Sigma N_{\mu})$.

In the first issue, the most obvious measure of the dissimilarity (or similarity) between two samples is the distance between them. One way to begin a clustering investigation is to define a suitable distance function and compute the matrix of distances between all pairs of samples. If distance is a good measure of dissimilarity, then one would expect the distance between samples in the same cluster to be significantly less than the distance between samples in different clusters. In our case, the Euclidean distance[24] is used as a measure of dissimilarity between the samples because it implies that the feature space be isotropic. Consequently, clusters defined by Euclidean distance will be invariant to translations or rotations-rigid-body motion of the data points. However, they will not be invariant to linear transformations in general, or to other transformations that distort the distance relationships. Therefore, if clusters are to mean anything, they should be invariant to transformations natural to the problem.

One way to achieve invariance is to normalize the data prior to clustering. To obtain invariance to displacement and scale changes, we translate and scale the axes so that all of the features have zero mean and unit variance. Therefore, the Euclidean distance is used as measure of dissimilarity because all of the measurements are of the same type, and that they have been normalized.

---

[24]Given two points $x = (x_1, x_2)$ and $y = (y_1, y_2)$, the ordinary distance between these points is called the Euclidean distance defined as: $d_\varepsilon(x,y) = [(x_1-y_1)^2 + (x_2-y_2)^2]^{\frac{1}{2}}$.

The second issue, evaluation of a partitioning of a set of samples into clusters, the problem is to define a criterion function that measures the clustering quality of any partition of the data. To define hierarchical clustering in terms of the algorithm used to generate the hierarchy is somewhat unfortunate because it equates a method, but can be implemented in several ways, with a specific algorithm. However, this approach to defining a clustering method is direct, simple to understand, and is commonly used in much of the literature. The algorithms used in the study are: 1. - Single-Linkage algorithm which joins together the two closest objects to form a cluster, then the next two closest objects to form a cluster, then the next two closest objects, and so on. If the two objects to be joined lie in different clusters obtained in previous steps, the two clusters are joined instead. The term single linkage is used because two clusters are joined if any of the distances between the objects in the different clusters is sufficiently small - that is, if there is a single link between the clusters. 2. - Complete-Linkage algorithm is related closely to the single-linkage method, only that each cluster is characterized by the longest link needed to connect every member of a cluster to every other member. 3. - Average-Linkage within the new group, that instead of relying on extreme values as in the two cases above, it may be of interest to characterize a cluster by the average of all links within it. 4. - Average-Linkage between merged groups which consists of evaluating the potential merger of clusters i and j in terms of the average similarity for links between two clusters.

5. - Centroid algorithm, that is, in statistical analysis the mean is often the basic summary statistic for a set of data. The familiar t-test and the analysis of variance technique both are used to identify differences between groups by testing for differences between their means. It then should be appealing to cluster hierarchically by merging at each stage those two clusters with the most similar mean vectors or centroids. 6. - The Ward algorithm in which the merges at each stage are chosen so as to maximize an objective function based on the error sum of squares and the objective is to find at each stage those two clusters whose merger gives the minimum increase in the total within group error sum of squares.

In the case of non-hierarchical clustering the algorithm used implements a clustering procedure that labels points (called patterns) in a feature space. The objective is to mark all patterns which are "close" to one another with the same label and to define different labels for any two patterns which are not close. It is based on a squared error criterion and it always searches for clustering with the smallest squared-error.

The specific computer packages utilized (runs made on the CDC 6500 at the Computer Center, Michigan State University, East Lansing, Michigan) are briefly discussed below.

MATRIX COMPUTING SYSTEM

MATRIX is a FORTRAN-based computer system designed for both ease and versatility in implementing any analysis, statistical or otherwise, that can be expressed in terms of albegraic operations on real matrices. MATRIX was used to perform principal components analysis and manova analysis. Documentation is available from the MSU Computer Laboratory.

STATISTICAL PACKAGE FOR THE SOCIAL SCIENCES (SPSS)

SPSS is an integrated system of computer programs designed for the analysis of social science data. The system provides a unified and comprehensive package that enables the user to perform many different types of data analysis in a simple and convenient manner. SPSS allows a great deal of flexibility in the format of data. It provides the user with a comprehensive set of procedures for data transformation and file manipulation, and it offers the researcher a large number of statistical routines commonly used in the social sciences. SPSS Subprogram DISCRIMINANT was used for the discriminant analysis. Documentation is available from the MSU Computer Laboratory.

STORED SIMILARITY MATRIX APPROACH PROGRAM

This program was used to perform the hierarchical clustering analysis. For a description of the program and program documentation, see Anderberg, 1973.

CLUSTER PROGRAM

This program was used to perform the non-hierarchical clustering analysis. For a description of the program and program documentation contact Prof. Richard Dubes, Computer Science Department, MSU.

Figure C1: Sample of 1:40,000 scale showing Pinus and Araucaria forest types.

A - Araucaria
P - Pines


330
40.

Figure C1: Sample of 1:40,000 scale showing Pinus and Araucaria forest types.

A - Araucaria
P - Pines


330
40.
A
A
A
A
A
A
A
A
A
A
A
A
A
A
P
P
P
P
P
P
P
P
P
P

Figure C2: Sample of 1:40,000 scale showing Hardwoods forest type.

H - Hardwoods

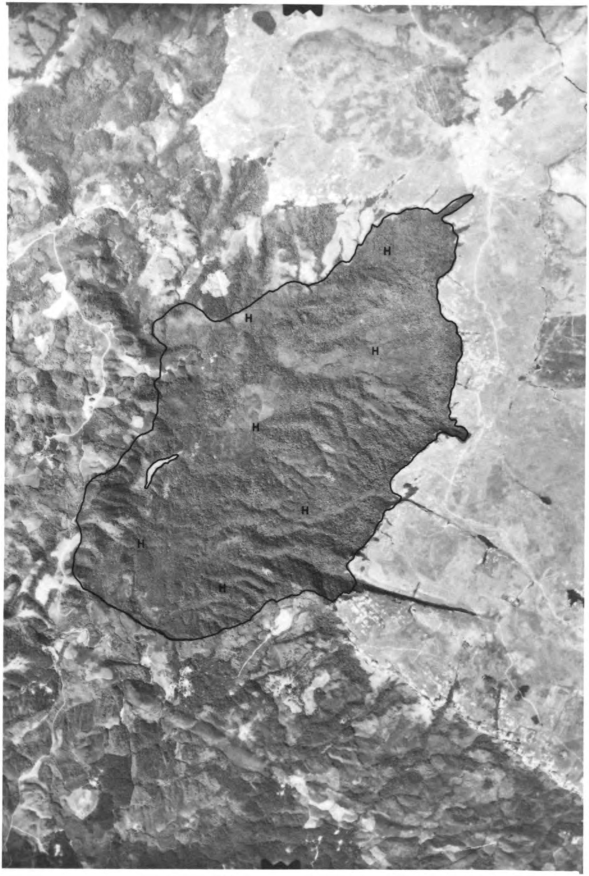

H
H
H
H
H
H
H

Figure C3: Sample of 1:40,000 scale showing Mixed forest type.

M - Mixed


40
M
M
M
M
M
M
M
M
M

Figure C4: Sample of LANDSAT imagery (channel 6: 0.7 um - 0.8 um) in 1:250,000 scale showing the four forest types.

A - Araucaria
P - Pines
H - Hardwoods
M - Mixed

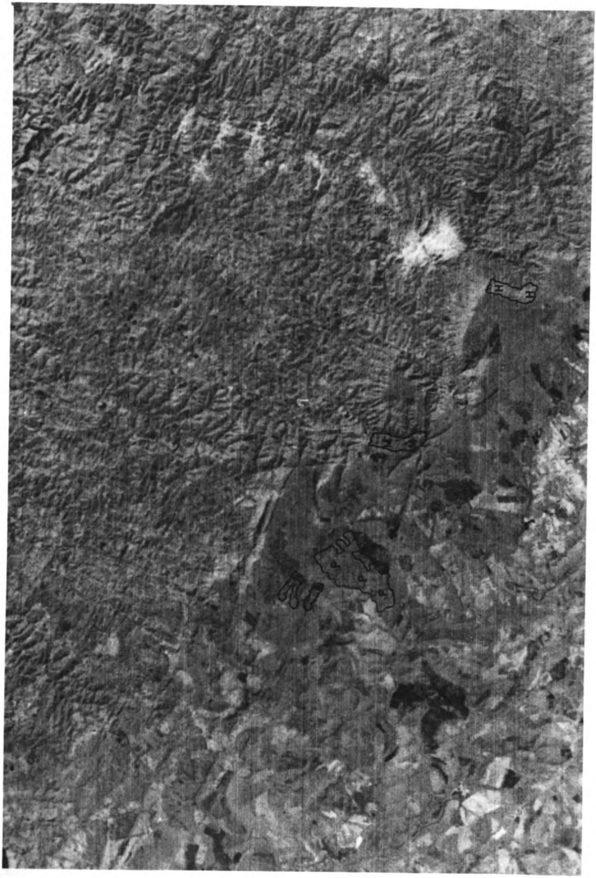


Figure C5: Sample of LANDSAT imagery (color composite reproduction in black and white) in 1:250,000 scale showing the four forest types.

A - Araucaria
P - Pines
H - Hardwoods
M - Mixed

LITERATURE CITED

LITERATURE CITED

Akca, A. 1971. Identification of land use classes and forest types by means of microdensitometer and discriminant analysis. In: Applications of Remote Sensors in Forestry. IUFRO. Freiburg, Federal Republic of Germany.

Anderberg, M. R. 1973. Cluster analysis for applications. Academic Press. New York.

Anderson, T. W. 1958. An introduction to multivariate statistical analysis. John Wiley. New York.

Bartlett, M. S. 1950. Tests of significance in factor analysis. British Journal of Psychology (Statistical Section), 3, 77-85.

Bartlett, M. S. 1973. Properties of sufficiency and statistical tests. Proceeding of the Royal Society, A 160, 268-282.

Box, G. E. P. 1949. A general distribution theory for a class likelihood criteria. Biometrika, 36, 317-346.

Doverspike, G. E., Flynn, F. M., Heller, R. C. 1965. Microdensitometer applied to land use classification. Photogrammetric Engineering. Vol. XXXI, No. 2, 294-306.

Driscoll, R. S., Francis, R. E. 1975. Range inventory: Classification of plant communities. In: Evaluation of ERTS-1 data for forest and rangeland surveys. USDA Forest Serv. Res. Paper PSW-112. Pacific Southwest Forest and Range Exp. Stn., Berkeley, California.

Driscoll, R. S., Reppert, J. N., Heller, R. C. 1974. Microdensitometer to identify plant communities and components on color aerial photos. Journal of Range Management, 27: 66-70.

Dubes, R. C. 1974. The cluster program. Department of Computer Science. Michigan State University. Mimeo.

Dubes, R. C., Jain, A. K. 1979. Clustering methodologies in exploratory data analysis. (For Volume 19, Advances in Computers, Marshall C. Yovits, Ed.). Department of Computer Science. Michigan State University, East Lansing, Michigan.

Fisher, R. A. 1936. The use of multiple measurements in taxonomic problems. Annals of Eugenics 7, 179-188.

Francis, R. E., Driscoll, R. S. 1976. Range inventory: Classification and mapping of plant communities. In: Evaluation of SKYLAB (EREP) data for forest and rangeland surveys. USDA Forest Serv. Res. Paper PSW-113. Pacific Southwest Forest and Range Exp. Stn., Berkeley, California.

Goldstein, A., Rosenfeld, A. 1964. Optical correlation for terrain type discrimination. Photogrammetric Engineering. Vol. XXX. No. 4, 639-646.

Green, P. E., Carroll, J. D. 1976. Mathematical tools for applied multivariant analysis. Academic Press. New York.

Hartigan, J. A. 1975. Clustering algorithms. John Wiley, New York.

Heller, R. C. 1975. Evaluation of ERTS-1 data for forest and rangeland surveys. USDA Forest Serv. Res. Paper PSW-112. Pacific Southwest Forest and Range Experiment Station, Berkeley, California.

Higgins, G. C. 1964. Methods for engineering photographic systems. Applied Optics. Vol. 3, No. 1, 1-10.

Jackson, H. R., Hodgson, W. A., Wallen, V. R., Philpotts, L. E., Hunter, J. 1971. Potato late blight intensity levels as determined by microdensitometer studies of false color aerial photographs. Journal of the Biological Photographic Association. Vol. 39. No. 2. 101-106.

Jordan, D. C., Graves, D. H., Hammetter, M. C. 1978. Use of manual densitometry in land cover classification. Photogrammetric Engineering and Remote Sensing. Vol. XLIV, No. 8 - 1053-1059.

Lillesand, T. M. 1976. Fundamentals of electromagnetic remote sensing. State University of New York. College of Environmental Science and Forestry. Syracuse, New York.

Lillesand, T. M., Eav, B. B., Manion, P. D. 1979. Quantifying urban trees stress through microdensitometric analysis of aerial photography. Photogrammetric Engineering and Remote Sensing. Vol. XLV, No. 10, 1401-1410.

Lillesand, T. M., Kiefer, R. W. 1979. Remote sensing and image interpretation. John Wiley. New York.

Maack, R. 1968. Geografia fisica do Estado do Parana. Banco do Desenvolvimento Economico do Parana. Curitiba, Parana.

Morrison, D. R. 1976. Multivariate statistical methods. McGraw-Hill. New York.

Rao, C. R. 1952. Advanced statistical methods in biometric research. John Wiley. New York.

Ray, R. G., Fisher, W. A. 1960. Quantitative photography - A geologic research tool. Photogrammetric Engineering. Vol. XXVI, No. 1, 143-150.

Rib, H. T., Miles, R. D. 1969. Automatic interpretation of terrain features. Photogrammetric Engineering. Vol. XXXV. No. 2. 153-164.

Rosenfeld, A. 1962. Automatic recognition of basic terrain types from aerial photographs. Photogrammetric Engineering. Vol. 38, No. 1, 115-132.

Rosenfeld, A. 1962a. An approach to automatic photographic interpretation. Photogrammetric Engineering. Vol. XXVII, 660-665.

Sadowsky, F. G., Malila, W. A. 1977. Investigation of techniques for inventorying forested regions. Volume I: Reflectance modeling and empirical multispectral analysis of forest canopy components. ERIM, Ann Arbor, Michigan.

Sayn-Wittengenstein, L. 1978. Recognition of tree species on aerial photographs. For. Manag. Inst. Inf. Res. FRM-X-118. Canadian Forest Service.

Schanda. E. 1976. Remote sensing for environmental sciences. Springer-Verlag. New York.

Sneath, P. H. A., Sokal, R. R. 1973. Numerical Taxonomy, W. H. Freeman and Company. San Francisco.

Sokal, R. R., Rohlf, F. J. 1960. Biometry. W. H. Freeman and Company. San Francisco.

S.P.S.S. 1975. Statistical Package for the Social Sciences. McGraw-Hill. New York.

U.F.P. 1974. Estudos das alternativas tecnicas economicas e sociais do setor florestal do Parana, "Sub-programa Materia-Prima." Universidade Federal do Parana. Centro de Pesquisas Florestais. Convenio Sudesul-Governo do Estado do Parana - I.B.D.F. Curitiba - Parana.

Von Steen, D. H., Learner, R. W., Gebermann, A. H. 1969. Relationships of film optical density to yield indicators. Proceedings of the Sixth International Symposium of Remote Sensing of Environment. Vol. 2, Ann Arbor, Michigan. 1115-1122.

Wilks, S. S. 1932. Certain generalizations in the analysis of variance. Biometrika, 24, 471-474.