AN ABSTRACT OF THE DISSERTATION OF

Jing Huang for the degree of Doctor of Philosophy in Forest Science presented on March 23, 1998. Title: Spatial Characterization and Analysis of Forests in the Mount Bachelor Volcanic Chain, Central Oregon

Signature redacted for privacy.

Abstract Approved: _____
David A. Perry

Forest spatial pattern is a primary interest of landscape ecology due to the relationships between spatial configuration of biotic components and ecological processes. The spatial pattern must be measured in meaningful ways so that relationships between forests and their environment can be analyzed. Aerial and satellite imageries provide ecologists a variety of scale choices at which the spatial information of forests can be presented in levels ranging form individual trees to landscapes.

The high lava plain of central Oregon is characterized by young lava flows of moderate relief interrupted by scattered cinder cones and lava buttes. The regosolic soils developed on pumice support open coniferous forests of nationwide significance. The relationship between forests and the harsh rocky land has not been analyzed, yet large portions of the forest have been logged at various intensities over the last 40-50 years. A better understanding of the relationship between forests and environment is needed for management of healthy forest ecosystems.

It is the intent of this study to use remotely sensed data to measure the spatial variability of forest patterns across the lava landscape in Mount Bachelor volcanic chain, and to analyze relationships between forest structural attributes and environmental variables. First, I used aerial photographs to characterize tree point pattern and measure canopy crown closure and density. A step-wise digital approach based on spatial, spectral,

and topographic characteristics of the photographic data was developed to measure forest spatial patterns. The method provides a fast and accurate, yet low cost way to characterize tree point pattern and measure canopy crown closure and density. Second, I used Landsat TM imagery to estimate leaf area index (LAI). A new approach using multiple regression analysis was developed to overcome the saturation problems of commonly used vegetation indices at high LAI ranges and improve the performance of Landsat TM data in estimating LAI. Finally, I conducted an analysis using the spatial data developed in this study and supplemented with that obtained from Deschutes Nation Forest. The study documents methods of integrating multiple GIS data layers for spatial analysis and parameterizing relationships between forests and environment.

# Spatial Characterization And Analysis of Forests in The Mount Bachelor Volcanic Chain, Central Oregon

by

Jing Huang

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented March 23, 1998
Commencement June, 1998

# Acknowledgements

I would like to express my deepest thanks to Dr. David A. Perry, professor of Ecosystem Analysis, for providing the graduate opportunity at the Oregon State University. His efforts in initializing the project, continuing support during the course of the research, and careful editing of the manuscript are greatly appreciated. The completion of this comprehensive project would not have been possible without his guidance on a variety of ecological and related scientific topics.

I am also deeply indebted to Dr. Andrew Youngblood who I thank for initializing the project and providing assistance in the field. I am extremely grateful for his help in field data collection.

My grateful thanks go to the other members of my graduate program committee. Dr. Chaur-Fong Chen provided many unique ideas and technical supports to this study. Prof. Robert Schultz opened his Lab to me for free and was always available when I needed help. Dr. David Marshall provided advice on the research. The stimulating discussions with Dr. Richard Waring during the past four years benefited not just my Ph.D. study, but my career as well.

Thanks are also due to many fellow graduate students for their generous help in discussing progress of the project, data analysis and interpretations, and especially their friendships that enriched my life. I would like especially thank Nestor Rojas, Kathy Johnson, Marla Gilham, Susan Simard, Meg Ruby, Flint Hughes, NaDene Sorensen,, Martin Hutten, Shelley Church, Hua Chen, Eileen Helmer, Jingfan Duan and Gladwin

Joseph for their assistance either in field data collection or lab work They really make my graduate experience very worthy and memorable.

My gratefulness goes to the staff and faculty in Forest Science Lab and Forest Department. Especially I want to thank Carol Wood, Sandy Lewis, Logan Norris, George Lienkaemper, Maria Fiorella, Warren Cohen, David Turner, Gody Spycher, Susan Stafford, Lisa Ganio, Ching-Yan Li, and Daolan Zheng for their support of the study.

Thanks are also given to Karen Bennett, Larry Chitwood, Lupita Weslor, Kimberly Johnson, and Eric Leavy in Deschutes National Forest, US Forest Service, for their assistance in the study.

My special thanks go to my father Fusheng Huang, and my mother Guiying Huang who fostered my dream on the natural world, and my wife Shari Xu who has been helping me to make the dream into truth. Their love of me makes this study more meaningful.

The research was supported by the USDA Forest Service Grant FS6500-46 awarded to Forest Science Department of Oregon State University.

# TABLE OF CONTENTS

# TABLE OF CONTENTS (Continued)

# TABLE OF CONTENTS (Continued)

# TABLE OF CONTENTS (Continued)

# LIST OF FIGURES

# LIST OF FIGURES (Continued)

# LIST OF TABLES

To my parents and my wife.

# Spatial Characterization And Analysis of Forests In The Mount Bachelor Volcanic Chain, Central Oregon

## Chapter 1. Introduction

### 1.1. Background

The high lava plain of central Oregon is characterized by young lava flows of moderate relief interrupted by scattered cinder cones and lava buttes. It contains forests of nationwide significance (Franklin and Dyrness, 1973). The regosolic soils developed on pumice support open coniferous forest vegetation and apparently enable variety of species such as ponderosa pine (*Pinus ponderosa*), lodgepole pine (*Pinus contorta*), Douglas fir (*Pseudotsuga menziesii*) and true fir (*Abies grandis* and *Abies concolor*) to extend eastward. The frequent coincidence of the boundaries of ponderosa pine forest and pumice soil areas indicates a unique relationship between forests and environment in this area. The response of forests in the harsh rocky land has not been analyzed, yet large portions of these areas have been logged at various intensities over the last 40-50 years (Popovich 1991). Many issues and problems have been identified by resource specialists. Further studies in addition to an inventory of the forest under management are necessary for maximum production of desired forestry commodity while maintaining a healthy forest ecosystem. Problems to be resolved include such questions as what relationships exist between forests and environment, what the driving variables are to shape forest spatial patterns, and what impacts of the change of forest spatial patterns are on the environment

and regeneration of the forests. Answering these questions requires analyses of multiple data layers at multiple dimensions and scales.

Spatial pattern of landscapes is a primary focus of landscape ecology (Forman and Godron 1986), due to the relationships between spatial configuration of biotic components and ecological processes (Turner and Gardner 1989, Waring 1994). Landscape spatial pattern must be measured in meaningful ways so that the interactions between patterns and ecological processes can be understood (Turner and Gardner 1989). Remote sensing is a commonly used tool for acquiring spatial information regarding landscape changes at various scales (Leckie 1990). Aerial and satellite imageries provide ecologists a variety of scale choices at which spatial information can be presented in levels ranging from individual trees (Huang 1997a), to stands (Cohen et al. 1992), and landscapes (Ripple et al. 1991). A step-wise approach is required to study patterns and processes across scales. First, one must have measures to describe spatial patterns, so that criteria can be established for relating patterns to their causes and consequences (Gardner et al. 1987, Milne 1988). Subsequent studies analyze relationships among different measures and test hypotheses. The broad scale nature of landscapes requires creative approaches to experimental design (Turner 1989).

Objectives of this study were to use remote sensing techniques to characterize forest spatial patterns across scales from tree to stand levels in Mount Bachelor volcanic chain area, and then integrate all available data layers into a GIS to analyze the responses of forest spatial patterns to environmental variables. The study is expected to improve the understanding of forest-environment interaction and provide information that will assist resource managers in maintaining healthy forest ecosystems.

## 1.2. Description of the study area

The Mount Bachelor Volcanic Chain is located in the eastern part of the central High Cascades of Oregon, southeast of the clustered composite volcanoes of the Three Sisters and Broken Top mountains (Fig. 1.1). The Mount Bachelor Volcanic Chain provides one example of the type and scale of eruptive activities that have produced the bulk of the

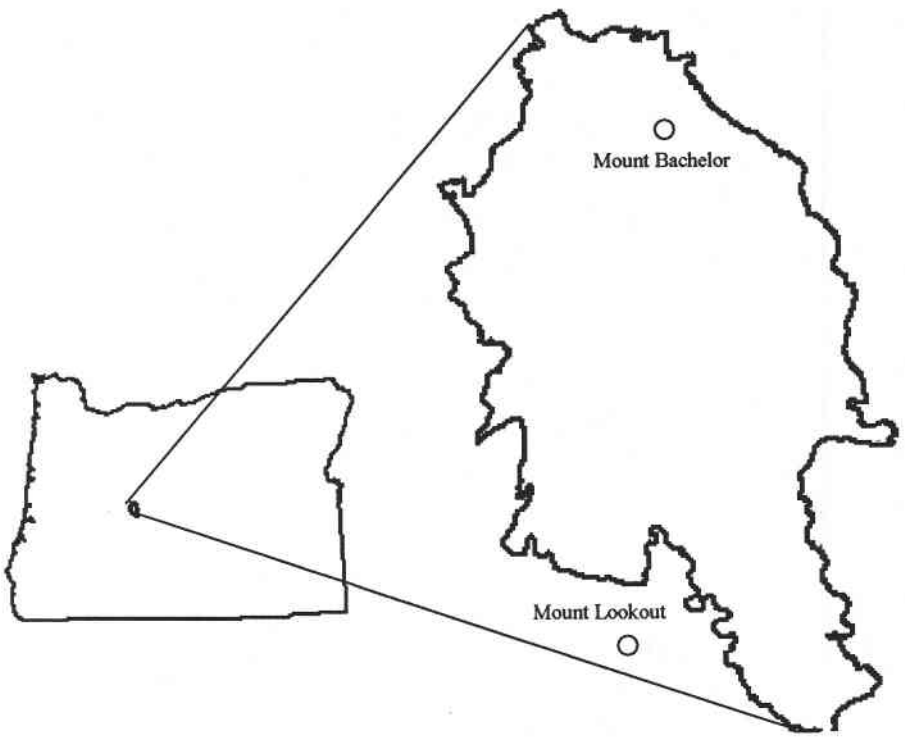



Fig. 1.1, The Study area

High Cascade platform, which consists chiefly of scaria cones and lava flows, shield volcanoes, and a few steep-sided cones of basalt and basaltic andesite (Scott 1992). The

chain is approximately 27 kilometers long and 16 kilometers wide, covering 250 $km^2$. Many vents in the study area lie along north-trending alignments. Eruptions along the chain were not continuous but occurred as discrete pulses from localized segments of the chain. Estimates derived from the stratigraphic framework of volcanic rocks and unconsolidated deposits suggest that the bulk of the chain was formed during a period of 12,000 to 18,000 years before present (Scott 1992).

The topography of the study area reflects the volcanic history. Broad basins from basalt flows are dotted with many cinder cones and volcanic domes. Stepped benches from old basalt flows and fault scarp ridges are also common. Slope ranges from 0 to 48%. Elevation varies from 1297 to 2760 meters with an average of 1646 meters.

The climate of the study area is characterized by a short growing season and minimal summer precipitation. Climatic data from Wickiup Reservior, 3 kilometers south of the study area, indicate average annual precipitation of 520 millimeters, 85% of which falls between October and April as snow. Diurnal summer temperature fluctuates widely, with hot days and cold nights. July, August and September are very dry with total rainfall averaging 58 millimeters. Much of the summer rain comes during brief, high-convection storms.

Soils are formed from deposit materials of the volcanic eruptions, including volcanic ash, pumice, and cinders (Larsen 1976). The most extensive deposit is volcanic ash from the eruption of Mount Mazama (Crater Lake) about 6,000 years ago. The Mazama ash deposit is about 3 feet thick on the south of the study area and extends northerly, becoming thinner and finer with distance from its source (Scott, 1992, Larsen 1976).

Because of young geologic age and the similarity of soil parent materials across the study area, soils show relatively little morphological development.

The rough lava flows support open stands of mixed coniferous forests mainly dominated by ponderosa pine (*Pinus ponderosa*) associated with a variety of other species such as Douglas fir (*Pseudotsuga menziesii*), western white pine (*Pinus monticola* ), lodgepole pine (*Pinus contorta*), white fir (*Abies concolor*), grand fir (*Abies grandis*) and mountain hemlock (*Tsuga mertensiana*). A variety of shrub species present on the study area including snowbrush (*Ceanothus velotinus*), pinemat mazanita (Arctostaphylos nevadensis), greenleaf manzanita (*Arctostaphylos patula*), antelope bitterbrush (*Purshia tridentata*), smallflower penstemon (*Penstemon davidsonii*), and golden chinkapin (*Castanopsis chrysophylla*). Long-stolon sedge (*Carex pensylvanica*), Ross sedge (*Carex rossii*), needlegrass (*Stipa occidentalis*), and squirreltail (*Stitanion hystrix*) are the common herbaceous plants.

### 1.3. Study objectives

The study had two overall objectives. First, aerial photographs and Landsat TM imageries were used to characterize forest structure and spatial patterns, including tree point pattern, canopy crown closure, density, and leaf area index (LAI). I used a photographic scanner to scan the true color aerial photographs into digital formats. A step-wise digital method based on the spatial, spectral and topographic characteristics of the imageries was developed to extract tree point pattern, canopy crown closure and density. The digital image analysis method provided a fast and accurate, yet low cost way

of extracting forest information in the scale of individual tree level. I used Landsat TM imagery to estimate LAI across the lava landscape. A multiple regression model was built to combine all spectral bands for the full power of estimating LAI from Landsat TM imagery. The model considered the multicollinearity problem among bands to improve the accuracy of prediction.

Second, I integrated the spatial data layers developed in this study and supplemented with that obtained from US Forest Service, Deschutes National Forest into a GIS, and analyzed the relationships between forest spatial patterns and environmental variables. Statistical analyses were conducted to detect the spatial coincidence among species occurrence, the spatial patterns of canopy crown closure and LAI, and topographic, soil and geologic data layers. Empirical relationships which describe forest responses to environmental variables were established between forest attributes and topographic, soil and geologic factors. The study was expected to improve the understanding of forest-environment interaction and provide information that will assist resource managers in maintaining healthy forest ecosystems.

# Chapter 2. Characterizing Forest Spatial Patterns Using Digitized Aerial Photographs

## 2.1. Abstract

Aerial photographs have been used extensively to aid forest surveys and are considered to be essential tools for forest research and management. Recent developments in digital photogrammetry have come to the point where tedious visual interpretation and manual measurements can be effectively automated. Digital data can be obtained by either digital camera or digitizing the film prints into machine readable picture elements (pixels). Digital camera is a fairly new technology only available in recent years. However, huge sets of information obtained by aerial photography in the past have been stored as film prints. Film scanners are frequently used to convert these film prints into digital data. The increasing needs for characterizing forest spatial patterns require a thorough understanding of the spatial and spectral characteristics of these scanned data and their potential applications. In this paper we address the problem of measuring forest spatial patterns in analyzing high resolution digitized aerial photographs. The objectives of this study are to: (1) evaluate the data quality of digital conversion using photographic scanners with a view to forest applications, (2) examine the spectral and spatial characteristics of the scanned data, and (3) characterize forest spatial patterns using these scanned data. The data quality of digital conversion was discussed with regard to resolution, scanning noise, file size and features of interest. A step-wise digital approach based on spatial, spectral, and topographic characteristics of the scanned data was developed to extract forest spatial patterns. The method was evaluated for the efficiency

with which desired features can be identified. Spatial data layers of tree point pattern, and crown closure were produced as the results of these operations. Color aerial photographs covering the ponderosa pine forest in south central Oregon (scale 1:15840) were used in the study

## 2.2. Introduction

Spatial pattern is a primary focus of landscape ecology (Forman and Godron 1986), due to the relationship between spatial configuration of biotic components and ecological processes (Turner and Gardner 1989, Waring 1994). Spatial structure must be identified in meaningful ways so that interactions between landscape patterns and ecological processes can be understood (Turner and Gardner 1989). Aerial photographs provide a pictorial record of forests from which spatial data can be obtained and have been used by ecologists and foresters to aid survey and management since 1940 (Spurr 1960).

Forest attributes interpreted from aerial photographs include species composition, height, crown closure, stand density and age structure (USDA Forest Service 1975). Interpretation of these attributes is largely a visual and manual process that requires the use of image characteristics such as tone, texture, pattern, shape, size, shadow and site (Hoppus and Evans 1994, Lillesand and Kiefer 1994). Photo interpreters and photogrammetrists have always dealt with hardcopy photography as the source of information (Saleh, 1996), however, recent developments in digital photogrammetry allow tedious visual interpretation and manual measurements to be automated (Firestone et al. 1996). At present, conversion of film prints into digital forms is achieved by means of

image scanning technology (SaIeh 1996). Important questions remaining to be addressed include: What forest spatial attributes can be obtained from digitized aerial photographs? What accuracy can be achieved from digital image processing? What is the appropriate resolution to obtain desired information with a minimum volume of scanned data?

This paper presents a digital image analysis procedure for processing digitized aerial photographs and extracting forest spatial patterns. Our objectives are to (1) evaluate the data quality of digital conversion using photographic scanners with a view to forest applications, (2) examine the spectral and spatial characteristics of the scanned data, and (3) characterize forest spatial patterns using these scanned data. First, we formulate a procedure to evaluate the data quality of digital conversion with a view to the ground features of interest (in this particular study, the trees). The appropriate resolution for scanning photographs is determined based on that evaluation. Second, we examine the spectral and spatial characteristics of the scanned data at the chosen resolution. We analyze the optical properties of tree crowns which, by in large, determine the scattering and absorptive characteristics of forest canopies. These characteristics provide the basic information for extracting features of interest from digitized photographs. Finally, we formulate a digital image analysis procedure to measure crown closure and characterize tree point patterns. The accuracy of the approach is assessed by GPS survey and visual interpretation of the aerial photographs.

## 2.3. Methods

### 2.3.1. Data acquisition

#### 2.3.1.1 The study area

This research was conducted in the Mount Bachelor volcanic chain in the Deschutes National Forest, central Oregon. This area is approximately 27 km long in a north-south direction and 16 km wide in an east-west direction, covering 250 $km^2$. Elevation ranges from 1297 meters in the south-east up to about 2760 meters at the peak of Mount Bachelor. At lower elevations, forests consist mainly ponderosa pine (*Pinus ponderosa*) and lodgepole pine (*Pinus contorta*). As elevation increases, the forest type changes to ponderosa pine, Douglas fir (*Pseudotsuga menziesii*) and true fir (*Abies grandis* and *Abies concolor*) mixed forest. At higher elevations, the primary tree species are mountain hemlock (*Tsuga mertensiana*) and lodgepole pine. Forests are generally uneven-aged with diverse tree sizes. Tree stocking is relatively low, allowing individual trees to be clearly identified from stereoscopic view of the aerial photos. The background is pumice soil with very high reflectance. Aerial photos used in the study are true color with a scale of 1:15840, acquired in the summer of 1991 by the USDA Forest Service. A HP DeskJet flatbed scanner was used to scan the aerial photos into digital format. PANTONE standard color cards were used to evaluate the data quality of digital conversion.

### 2.3.1.2. Evaluation of scanning resolution and noise

The quality of scanned imagery is an important issue in generating softcopy. Criteria for the quality of scanned data include geometric precision, image resolution, image noise and dynamic range (Kolbl and Bach 1996). These are largely determined by the quality of both the film and scanner. The user, however, does need to determine the image resolution before digitalization. The resolution of a scanner is measured by resolution in dots per inch (dpi). Current available scanners provide a resolution up to 2400 dpi. Image noise is defined as the variation of pixel matrices over a homogeneous area (Kolbl and Bach 1996). Images scanned with smaller pixels generally have more noises than images with larger pixel size, as the larger pixels can be obtained by computing the mean of smaller pixels and the computations of the mean naturally cause a smoothing. It is important to determine the appropriate resolution for minimizing image noise while extracting the information in as much detail as possible during digital conversion.
We used standard color cards to evaluate the quality of digital conversion. Two PANTONE standard color cards, pure magenta and yellow, were aligned with the testing photo for scanning (Fig. 2.1). These two color cards were made from the same materials as the photographic film, and were assumed to have the same reflectance characteristics. We used standard deviation (SD) of the scanned data generated from both color cards and the testing photo to evaluate the gain of increasing resolution over image noise. With perfect film and scanner, the scanned data from these two pure color cards would be homogeneous with zero SD. Therefore, SD indicates the quality of the digital conversion that is related to the quality of both photographic films and scanners.

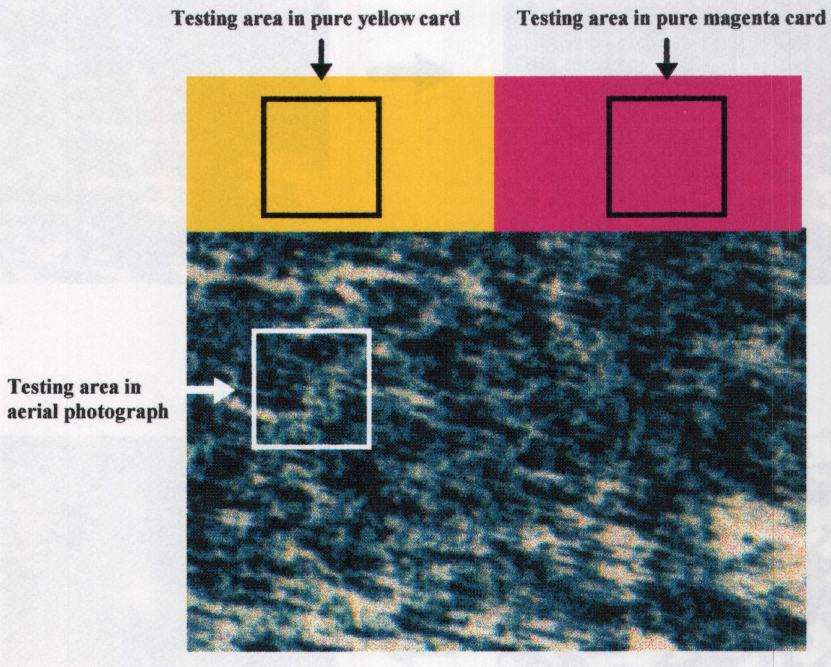



Fig. 2.1, Testing images for scanning noise evaluation. Pure color cards are aligned with aerial photograph for scanning.

Square areas of 5 by 5 mm were chosen from the testing photo as well as each of the two pure color cards, and scanned in unsigned 8-bit data structure in fifteen different resolutions from 100 dpi to 1500 dpi with 100 dpi of increment (Fig. 2.1). Therefore, a total of 45 images were generated from this process. The SD of pixel values in each of the images was calculated and then evaluated against resolution increment. Fig. 2.2 shows the summarized statistics. The SD of the forest image decreases slightly while the scanning resolution increases from 100 dpi to 500 dpi, and becomes relatively constant when scanning resolution is higher than 500 dpi. By contrast, the SD of pure yellow and magenta color cards in red and green bands remains fairly close to zero while the scanning

resolution increases in the range of 100-800 dpi, and increases significantly when the scanning resolution is higher than 800 dpi. The SD in the blue band is higher than zero at the beginning. Decreasing SD with increasing resolution in the forest image is due to the low spectral resolution of the scanned data. With unsigned 8-bit data structure, the spectral resolution is very limited (only 256 spectral classes can be differentiated). The spectral variation over the forest image is very high because of heterogeneous spectral properties of different objects on the image. When scanning resolution increases, the number of pixels increases, but all pixel values are forced into the range between 0 and 255. Therefore, the SD of pixel values in the forest image decreases. With resolutions higher than 800 dpi, noise generated during scanning contributes a large amount of SD, and hence compensates for such a decreasing trend. By contrast, the spectral response of pure color cards is more homogenous than that of the forest image. Ideally, the SD of the pixel values over the pure color images would remain zero regardless of resolution (if no noise was produced during scanning) because all pixel values would be the same. In reality (Fig. 2.2), SDs of pixel values remain fairly constant and near zero when the scanning resolution is lower than 800 dpi, and increase significantly when the scanning resolution increases above 800 dpi. This phenomenon results from noise, which is related to the quality of both photographic films and scanners. The test also revealed that for the particular scanner and color cards we used, data quality in blue band was poorer than in green and red bands.

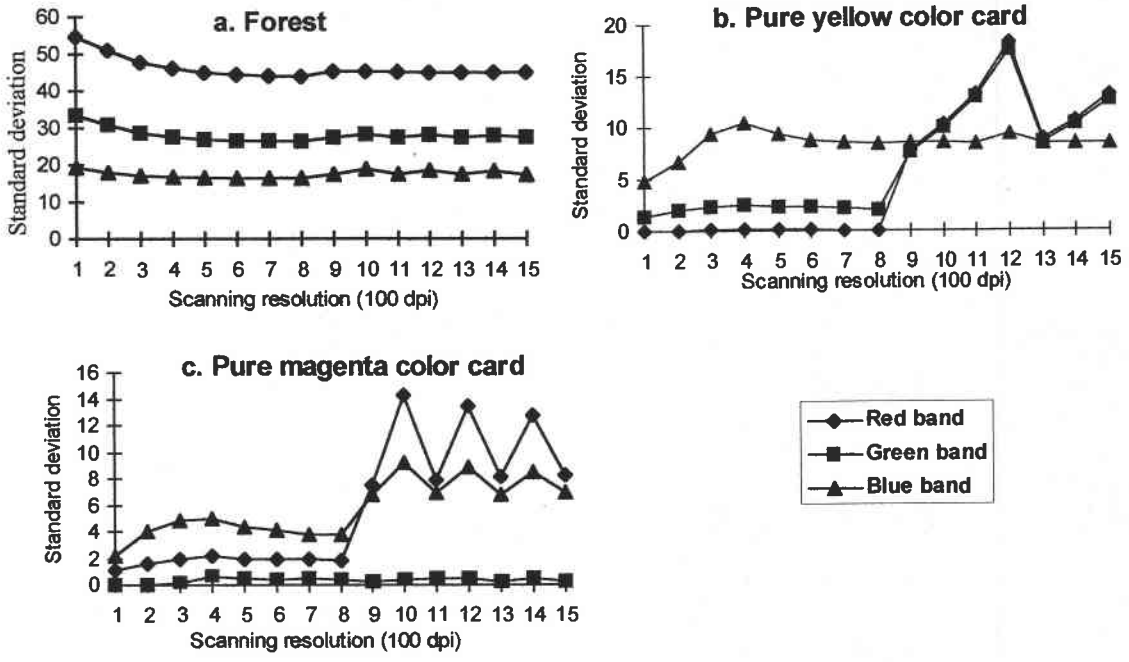



Fig. 2.2. The changes of SD of pixel value in different testing images. Notice that SD increases significantly when scanning resolution is higher than 800 dpi

### 2.3.1.3. Choosing scanning resolution

The final scanning resolution was chosen based on the evaluation of scanning noise, features of interest and file size of digital conversion. The spatial resolution of the scanned data must be fine enough to detect the features of interest (i.e. the pixel size must be smaller than the size of features of interest). File size also imposes constraints on choosing resolution for softcopy production. The file size of a standard 9 by 9 inch aerial photograph can easily reach 200 to 300 megabytes, depending on scanning resolution (Fig. 2.3). As a rule of thumb, the optimum resolution would be the one at which the identification of desired features is maximized while the file size and the noise of the data due to high resolution are minimized.

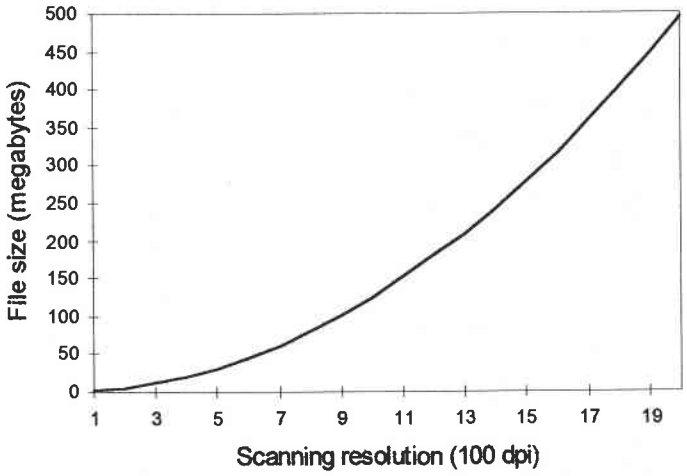



Fig. 2.3, File size as a function of scanning resolution for a standard true color aerial photograph

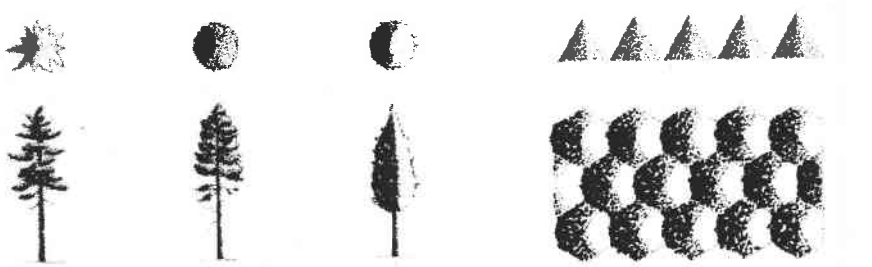
The actual decision in a particular application is always a compromise among scanning resolution, noise and file size. In our study, the field data indicated that a lower limit of 1 meter crown diameter would include all canopy trees. The scanning resolution for final digital conversion was set to 500 dpi, which was lower than the threshold of 800 dpi where scanning noise increased significantly (Fig. 2.2). With this resolution, one pixel is approximately equivalent to 0.8 by 0.8 square meters of ground area, which is smaller than the smallest tree crowns in which we were interested.

### 2.3.2. Spatial representation of tree crowns

Characterization of forest canopy by its spectral and optical properties from aerial photographs depends upon the interaction of solar radiation with its surface, which is determined in turn by the radiation wavelength (color), the incident and viewing angle to the surface, and most important of all, the fundamental optical properties of the canopy

surface (Egan, 1985). The most general way of characterizing a surface is through a model with a family of parameters. The following sections describe the spatial and spectral properties of tree crowns and forest canopy, and develop a model with which tree point pattern and crown closure can be measured.

To illustrate the spatial patterns of a forest canopy, tree crowns are represented as symmetric with a pointed oval or triangular profile (Fig. 2.4a). In dense stands, the pattern of crowns is honeycombed and the profile is saw-toothed (Fig. 2.4b). Due to uneven illumination of sunlight around tree crowns, there is always a light spot on one side of the crown defined on the ground by an imaginary triangulation joining the sun, the airplane and the tree crown.



**a. Tree crowns of different species** **b. Stand profile and canopy surface**

Fig. 2.4. Schematic representation of tree crowns and forest canopy

This confined spot appears strongly illuminated. Areas of the crown surrounding the bright spot are less illuminated depending upon their distance from the bright spot. On the opposite side of the light spot is a dark area called shade. Such characteristics of illumination define the spectral response of tree crowns as an asymmetric bell-shaped surface, skewed toward the bright spot of the crown. Fig. 2.5 shows the spectral profile

and surface of tree crowns over a scanning aerial image in the study area. The peak of the spectral surface in Fig. 2.5a corresponds to the lightest spot of a tree crown from the view of the camera. Three distinctive spectral peaks in Fig. 2.5b. indicate three adjacent tree crowns. The spectral properties of closed stands and open stands differ (Fig. 2.6).

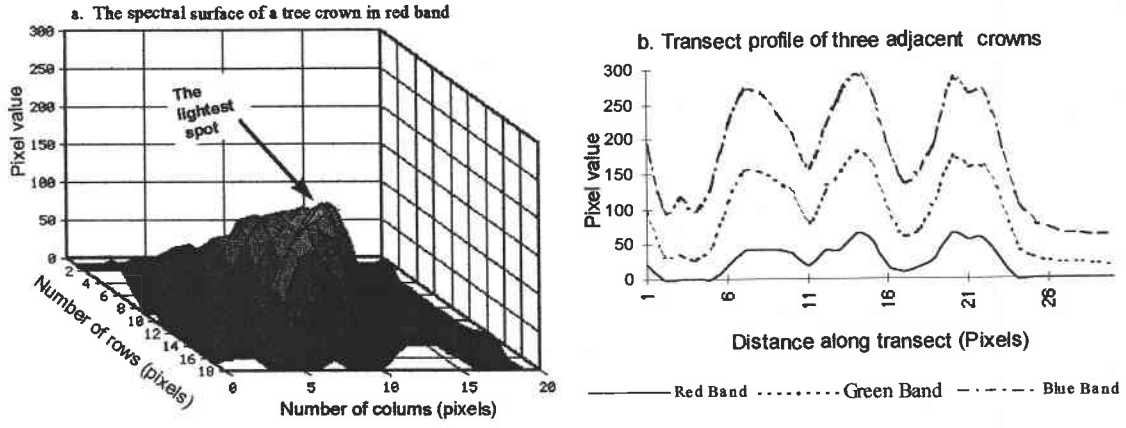



Fig. 2.5. The spectral surface and profiles of ponderosa pine (*Pinus ponderosa*) tree crowns in the study area.

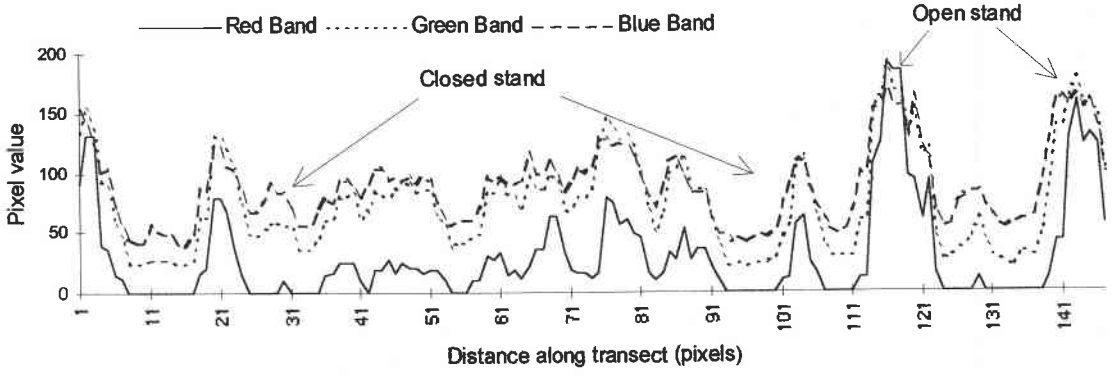



Fig. 2.6 Spectral response of a transect across a forest stand on the location as of Fig. 2.5.

In closed forest, canopy reflectance is less influenced by background reflectance, and the spatial distribution of spectral responses over the stand is rhythmical with the light spot on tree crowns appearing as peaks and dark shades as the valleys. By contrast, canopy spectral properties of open forests with light soils and few shrubs are strongly influenced by background reflectance, and the spatial distribution of spectral reflectance is irregular. Crown spectral responses are saturated by the high reflectance of pumice soil background and are therefore difficult to detect. The influence of background reflectance in open stands requires special treatment in later image processing.

### 2.3.3. Spectral properties of tree crown, shadow and background

In addition to the differences in illumination intensity within a crown, the spectral properties of crowns in red, green and blue bands are much different from shadow and background (non-forest area) due to the selective absorption of sunlight by tree leaves. Because tree crowns absorb red light and reflect green light, reflectance in the red band from tree crowns is much smaller than that in green and blue bands (Fig. 2.7). By contrast, shadows do not have such a selective nature and reflect sunlight proportionally. The non-forest background has very high reflectance and also reflects red, green and blue lights evenly. The reflectance profile across a forested transect is more sensitive in red band than in green and blue band (Fig. 2.6). This implies that the red band preserves most of the information of forest canopies. Based on these observations, we formulated a Band Ratio Index (BRI) defined as:

$$BRI = \sqrt{\frac{(R_g - R_r)^2}{(R_b - R_g)^2 + 1}} \qquad (2.1)$$

Where R is the reflectance value, and r, g, and b represent the red, green and blue bands. One is added to the denominator to avoid dividing by zero. This index was used to differentiate between tree crowns and shadows. The index values of crown pixels are expected to be significantly higher than that of the shadow pixels because of the greater absorption of red light by tree crowns.

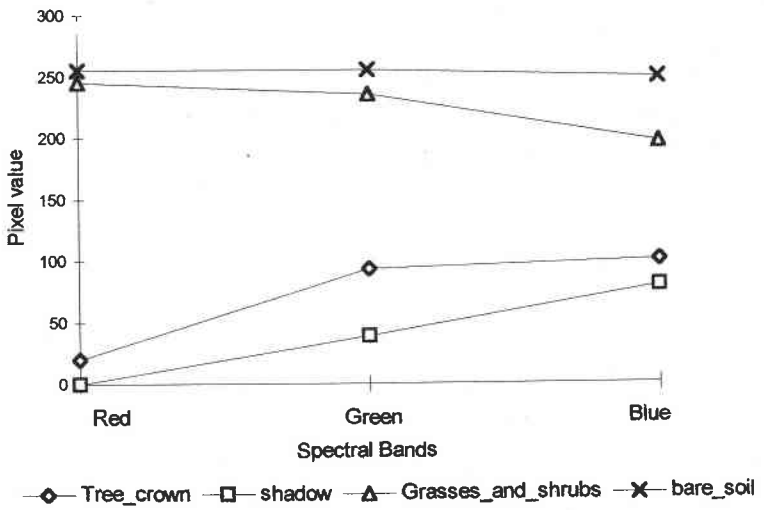



**Fig. 2.7. Spectral response of crown, shadow, shrubs and grasses, and bare soil. Four sample points from the study area.**

### 2.3.4. Filtering feature classes based on spectral and spatial properties

Detecting tree crown can be considered as a segmentation process in which only two types of regions are of interest: tree crowns and non-tree crown background. Recognition of tree crowns implies contrast with the background, but the scattering and absorptive properties of the background also affect the recognition of tree crowns. Irregular luminance due to the strong reflectance of the pumice soil background may saturate the

spectral peaks of tree crowns and blur crown images while the peak responses are important to distinguish trees from background. Shadows further complicate accurate detection of crowns. Therefore, it is necessary to filter out the background and shadow pixels from the scene before mapping tree point pattern and measuring crown closure.

We built a filter to remove background from the image. Based on reflectance intensity and the spectral distribution in RGB space, the filter was set as follows:

```
if(pixel_red>threshold_R)and(pixel_green>threshold_G)and(pixel_blue>threshold_B)
        pixel_red = NODATA
        pixel_green = NODATA
        pixel_blue = NODATA
end if
```

The threshold parameters were set by a statistical analysis that randomly sampled the image and determined the parameters. When a pixel is set to no data by the filtering operation, the pixel will not join the image analysis in later steps and will have no effect on pixels with data.

We built another filter to remove shadows from the scene. The shadow filter was set to make decisions in two steps. It first looked for a possible shadow pixel based on pixel value. Because crown pixel values and shadow pixel values are fairly close (Fig. 2.7), decisions based on absolute pixel values will not, alone, be sufficient to clearly differentiate between crown and shadow pixels. A further step was taken to identify the status of a pixel using its band ratio index (BRI) value. The shadow filter was set as:

```
if(pixel_red<threshold_R)and(pixel_green<threshold_G)and(pixel_blue<threshold_B)
        if( BRI<threshold_BRI)
                pixel_red = NODATA
                pixel_green = NODATA
                pixel_blue = NODATA
        endif
endif
```

After filtering, the spectral responses of the image become more homogenous and are ready for crown closure measurement and point pattern characterization (Fig. 2.8.).

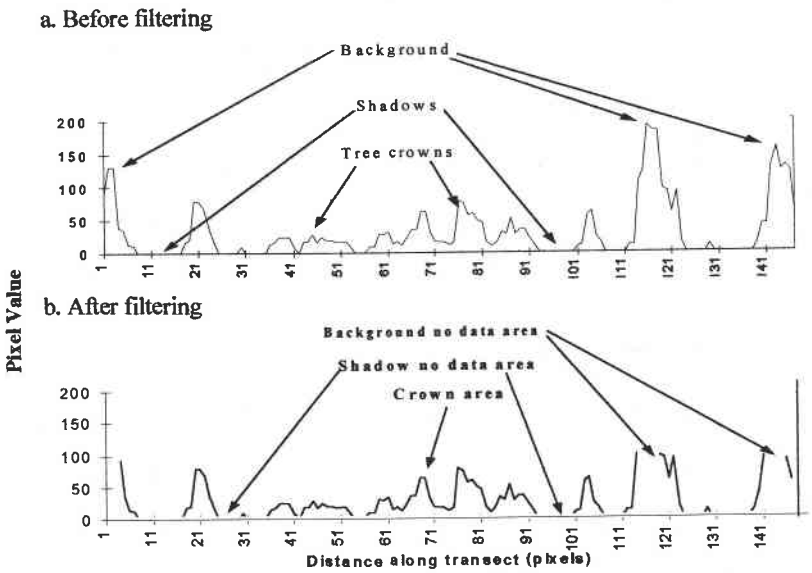



**Fig. 2.8, The spectral response curves before and after filtering shadows and background across the same transect as Fig. 2.6.**

### 2.3.5. Analysis of spectral topography

The final step in detecting tree crowns and measuring crown closure is accomplished by looking for the bell-shaped spectral topography of the filtered image. Spectral topographic features are produced by local spectral and intensity variations across the surface of the scene and correspond mainly to tree crowns. The following section describes a procedure to detect bell-shaped spectral features and utilize the findings to characterize tree point patterns and measure crown closure.

### 2.3.5.1. Moving window filter

A moving window was used to detect the location of a tree and measure crown closure. A moving window is a square filter with an odd number of pixels per side (Fig. 2.9). A pixel is located in the center of the filter and its coordinates are defined as the coordinates of the filter location within an image. The window filter moves row by row all over the image to find the features of interest based on a set of rules describing the spatial domain attributes of bell-shaped spectral responses, such as the size and shape (Fig. 2.9).

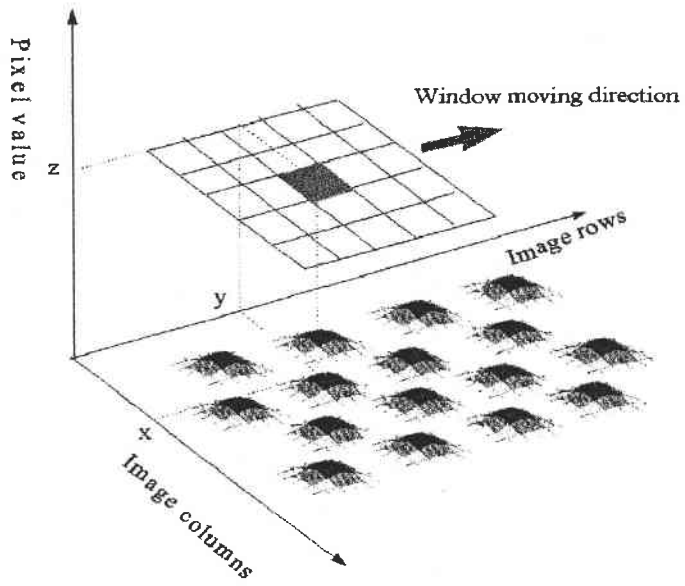



**Fig. 2.9. A moving window with 5x5 pixels. The convex-shaped spectral topography on X and Y plane represents crown spectral response**

During the searching process, the center pixel of the moving window will be overlaid onto each pixel of the image and then the setup rules are checked within the window and correspondent actions are taken. The moving window skips pixels with no data (i.e., areas

in which shadows and background have been filtered) and goes forward until pixels with data are found. The size of the moving window is determined by the filtering purpose. For example, for mapping tree point pattern, the size of the moving window approximates the average size of tree crowns.

#### 2.3.5.2. Characterizing tree point pattern

A common objective of spatial analysis of forest landscapes is the examination of the spatial occurrence of trees (Turner and Gardner 1989). If trees are represented as points, we call the spatial configuration of these points point pattern (Boots and Getis 1988). Forest ecologists examine tree point patterns for variety of reasons, including in particular to learn more about forest establishment and landscape dynamics and the processes responsible for generating them.

The spectral topography of a tree crown is represented as a bell shape with a light spot on its peak as discussed previously. The moving window method can be used to detect the location of the light spots that are assumed to be trees (Fig. 2.10). The window moves forward by one pixel in each time step. When the window filter moves to a new location along its route, the matched pixels on the image are picked up and the largest value is calculated. If the central pixel of the window has the largest value, the peak of a bell-shaped spectral response is detected. Its coordinates are recorded and the window moves to the next location until the entire image is scanned. The size of the moving window is critical in this operation. The window size must be approximately the same as that of an average tree crown. If the window size is too large, it is possible that only one

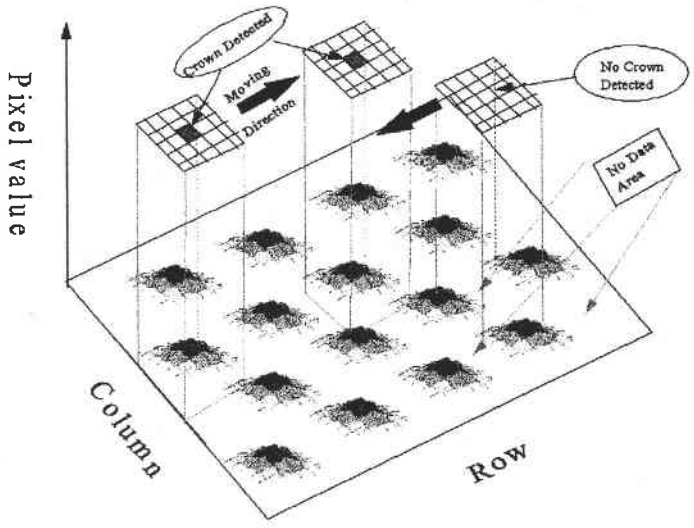



**Fig. 2.10, Tree point detection. The window filter moves row by row to detect the bell-shaped spectral topography and locate peak responses.**

tree crown is detected where actually two adjacent trees fit into the moving window at the same time. If the window size is too small, it is also possible that two peaks could be detected which might be the spectral response of one large irregularly shaped tree crown. With this procedure, an even-aged stand with regular or random tree distribution will be relatively easy to process and hence be characterized with greater accuracy, while an uneven-aged stand with trees distributed in clusters will be more difficult to process and hence be characterized with lower accuracy. As we will discuss in later sections, these are the main sources of error produced in mapping tree point pattern.

### 2.5.3.3. Measuring crown closure

Crown closure is defined as the projected surface area of tree crowns per unit ground area (Spurr 1960). Measuring crown closure is straightforward after background and

shadows are filtered from the image, because the only features left on the image are tree crowns. The standard definition of crown closure is clear enough theoretically but not operable in image analysis. The unit of ground area must be further specified, as the measure of crown closure per hectare could be much different from the same measure per square meter of ground area. The unit of ground area must be larger than the area covered by one image pixel. It does not make any sense to measure crown closure within one pixel of ground area because this pixel would be either 100% or 0% covered by tree crowns. The smallest unit is limited by the spatial resolution of the image. Crown area is defined as the area of tree crown viewable from the airplane at the moment the aerial photograph was taken. This means that any areas in shadow will be identified as non-crown.

The ground unit used to measure crown closure in our study was set to 1962.5 square meters based on the forest survey guide prepared by USDA Forest Service (circular plot with radius length of 25 meters) (USDA Forest Service 1994). This area is approximately equivalent to an image area of 55 by 55 pixels with the 0.8m pixel size used in our study. A moving window of 55 by 55 pixels was used to scan the image row by row (Fig. 2.11). Crown closure was calculated as the number of pixels representing tree crown divided by the total number of pixels in the moving window. The result was assigned to the central pixel of the moving window as crown closure measurement at that point.

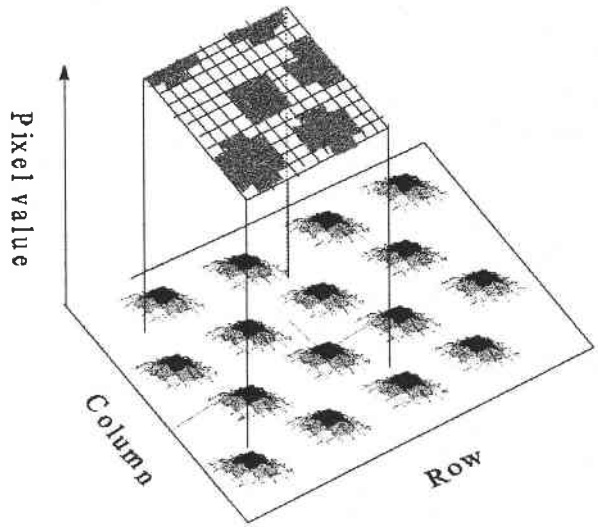



**Fig. 2.11, The simulated topography of a forest stand. The window moves row by row to count the number of pixels representing tree crowns.**

## 2.4. Application verification

To establish the practical validity of the approach described above, we used true color aerial photographs (1:15840 scale) for application verification. The photographs were scanned into digital format in 500 dpi spatial resolution by a HP DeskJet flatbed scanner.

### 2.4.1. Parameter analysis

We first conducted a statistical data analysis to determine the parameters necessary for spectral filtering to better differentiate among tree crowns, shadows and background. Here background means any area without trees. Since the areas with shrubs and grasses have lower reflectance than the areas with bare soil and therefore set the lower limit for the threshold parameters, only areas with shrubs and grasses were sampled as background.

For each photo to be processed, fifty points were selected randomly from each of the three feature classes, tree crowns, shadows and background, based on photo-interpretation and field experience. BRI values were also calculated for these points that represent crowns and shadows using formula (2.1). We analyzed the reflectance data and conducted two-sample tests for significant differences among feature classes. The pixel value distributions in RGB bands for crown, shadow and background are normally distributed and are suitable for standard t-test, but BRI for crown is highly skewed. Therefore, the non-parametric Wilcox test was used in hypothesis testing for BRI data. Tab. 2.1 shows the summarized statistics. Notice that the confidence intervals of red, green and blue bands, and BRI for different feature classes do not overlap in 99% level. The highlighted limits in Tab. 2.1 are used as the threshold parameters for filtering feature classes.

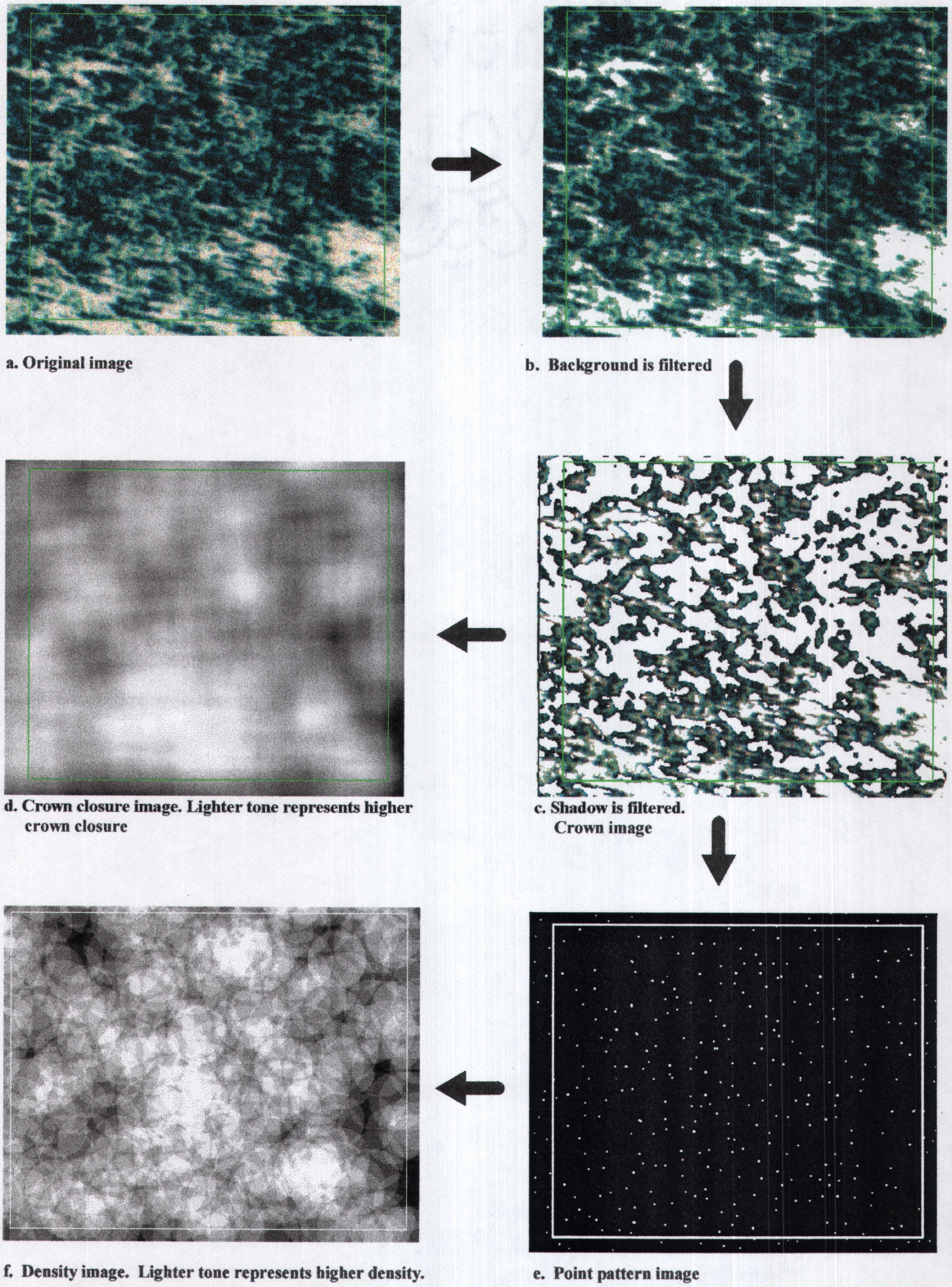
Tab. 2.1. The summarized statistics of parameter analysis

| Data Type | | Mean | Confident interval (99%) | |
|---|---|---|---|---|
| | | | Lower limit | Upper limit |
| Crown | Red | 54.2 | 42.1 | 66.3 |
| | Green | 104.9 | 95.2 | 114.6 |
| | Blue | 103.3 | 96.8 | 109.8 |
| Shadow | Red | 0.4 | 0 | **1.3** |
| | Green | 29.6 | 24.6 | **34.7** |
| | Blue | 58.0 | 52.1 | **64.0** |
| Background | Red | 236.1 | **229.5** | 242.7 |
| | Green | 227.4 | **220.9** | 233.8 |
| | Blue | 203.4 | **196.4** | 210.3 |
| Crown BRI | | 673.6 | 7.1 | 1340.2 |
| Shadow BRI | | 1.1 | 0.9 | **1.3** |

### 2.4.2. Image processing

Image processing includes the step-wise procedures of filtering and topographic feature analysis (Fig. 2.12). First, areas that met the criteria of background were set to no data by the background filter (Fig. 2.12b). Second, areas that met the criteria of shadows were set to no data by the shadow filter (Fig. 2.12c). The threshold parameters were used in making these decisions. After filtering, only tree crowns remained on the image (Fig. 2.12c). Note that portions of some tree crowns were recorded as "no data", thus left a "hole" of no-data on the top of these crowns. One reason for this problem is that a big tree crown with a dead top could show very high reflectance, which was falsely identified as a background feature. These pixel values representing the dead top were set to no data in later filtering processes.

The image after filtering was ready for measuring crown closure and characterizing tree point patterns in spectral topographic analysis. Window filters moved row by row over the filtered image (Fig. 2.10 and 11), detecting peak pixels for tree point pattern characterization, or counting crown pixels for crown closure measurement. Fig. 2.12d shows the gray scale representation of crown closure measurement. The gray scale was defined as the number of pixels representing crowns divided by the total number of pixels on the moving window. High crown closure was represented as a light tone and low crown closure as a dark tone. Fig. 2.12e shows the raster representation of tree point pattern. After a point pattern data layer is developed, it is straightforward to generate a density map by counting the number of dots on a moving window (Fig. 2.12f). The size of



a. Original image

b. Background is filtered

c. Shadow is filtered. Crown image

d. Crown closure image. Lighter tone represents higher crown closure

e. Point pattern image

f. Density image. Lighter tone represents higher density.

**Fig. 2.12. The step-wise procedure of image processing. The squares on the images are linked geographically. They represent the same ground area.**

the window was the same as used in crown closure measurement. Crown closure as calculated here is a conservative measurement because we removed shadow areas from the image that might contain understory trees. However, this should introduce only minor errors. By the definition of crown closure, only trees in the upper canopy layers are considered. These are unlikely to fall in the shadow of another upper canopy tree in an even-aged stand, but in an uneven-aged stand, they might. Some large ponderosa pine trees in our study area have huge tree crowns which could shadow other trees that should be included in a crown closure measurement. Tree point pattern mapping is influenced by the size of the moving window, especially when the forest stand is unevenly aged and distributed in a cluster pattern. A large window may cause small trees to be ignored and a small window may cause a large irregular tree crown to be falsely identified as two points. The background can also affect the accuracy of point pattern mapping. If some parts of the background area are not removed during filtering because the threshold parameter for removing background is too high, the pixels covering these areas are likely located on the peak of the spectral topography (Fig. 2.13). This will not be a problem if the peak is located on the edge of the data area since the central pixel of the moving window can never match the peak pixels on the edge of the image. If the remaining background area is wider than the width of the moving window and the peak pixel is located somewhere away from the edge of data area, a bell-shaped spectral topography could be formed and a false tree point detected (Fig. 2.13). The ideal input data for our proposed methodology are the images taken from even-aged, closed forest stands.

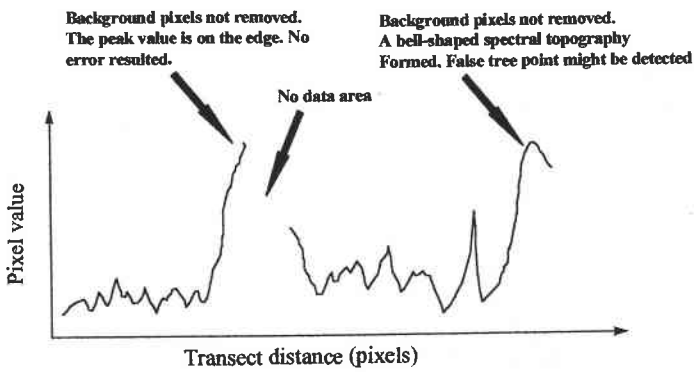



**Fig. 2.13, Simulated spectral response curve of a forest transect. Showing how errors resulted when background area is not filtered.**

### 2.4.3. Accuracy assessment

Only the crown image (Fig. 2.12c) and the point pattern image (Fig. 2.12e) are assessed for accuracy since errors can be introduced only in these processes. Accuracy of the crown image indicates the success of filtering feature classes. Accuracy of the point pattern image indicates the success of analysis of spectral topography.

#### 2.4.3.1. Crown image assessment.

The crown image was geographically linked with the original image. 100 sample points were selected randomly from the original image by visual examination, and the corresponding points on the crown image were evaluated (Tab. 2.2). If the crown and non-crown areas corresponded between the original and filtered images, the test results were accurate. Otherwise, the test results were inaccurate. Our test was 88% accurate in this assessment.

Tab. 2.2, Crown image assessment contingency table.

| Sampled Points \ Evaluated Points | | Crown Image | |
|---|---|---|---|
| | | Crown Area | Non-Crown Area |
| Original Image | Crown Area | True | False |
| | Non-Crown Area | False | True |

### 2.4.3.2. Point pattern image assessment.

The point pattern image was assessed in the same way as the crown image except the assessment was done for the whole image and an open stand as well, as open stands might have a higher error rate. The original image was geographically linked with the point pattern image, and one hundred tree crowns were located randomly on the original image by visual interpretation. If a crown on the original image was detected on the point pattern image, the result was true. Otherwise, it was false. Our test was 78% accurate in this assessment. The open stand assessment was conducted in a harvested area in which 40 trees were left. We produced a tree map using a GPS unit to locate each of the 40 trees. An aerial photograph covering the same area was used to detect these trees and generate tree point pattern. The final point pattern image was geo-referenced and overlaid onto the tree map, and 28 of 40 trees were detected. One point appeared on the point pattern image but the corresponding tree did not exist on the tree map. Therefore, 70% of the trees in this area were successfully detected.

The above applications show the utility of our methodology in characterizing and measuring forest spatial patterns. By successful validation of the approach and accuracy assessment of results, we have demonstrated the robustness of the method.

## 2.5. Conclusion

Accurate and robust measure of forest spatial pattern is a primary requirement in landscape studies. In this paper, we have presented a digital image analysis approach applicable to characterizing and measuring forest spatial patterns using digitized aerial photographs. Aerial photographs are one of the most important media storing remotely sensed information with high spatial resolution, and have shown utility in a variety of application domains (Leckie, 1990). Aerial cameras will remain the principal sensor for applications in medium and large scales for the near future (Mussio and Light 1995). Most of the analytical methods for aerial photos applicable to forest survey have been generally manual and labor intensive. Due to the increasing needs for measuring forest spatial patterns in landscape studies, automated tools and digital approaches are required. The step-wise approach we developed provides a systematic framework for generating digital data from photo prints and analyzing information related to the spectral, spatial and topographic characteristics of tree crowns and forest canopy.

The conversion of analog to digital data via scanner requires decisions on selecting appropriate resolutions and a careful evaluation of data quality. The resolution of a scanner is measured by resolution in dots per inch (dpi). Current available scanners provide a resolution up to 2400 dpi. Image noise is defined as the variation of pixel

matrices over a homogeneous area (Kolbl and Bach 1996). The sources of image noise could be related to the quality of both the films and photographic scanner. Images scanned with smaller pixels generally have a larger amount of noise than images with larger pixel size, as the larger pixels can be obtained by computing the mean of the smaller pixels and the computations of the mean naturally cause a smoothing. The spatial resolution of a photographic scanner is far beyond that needed for forest applications (Mussio and Light 1995). Resolution must be chosen to maximize the ability to identify objects of interest while minimizing image noise and data volume.

The development of a step-wise digital approach considers the spatial, spectral and topographic characteristics of the scanned data with a view to the features of interest (in this particular study, the tree crowns). It basically involves matching image domain features to the attributes of assumed tree crown models, the bell-shaped spectral topography. The initial foreground and background discrimination is based upon spectral response. Further crown and shadow discrimination is based upon both spectral response and band ratio index. Spatial considerations are incorporated by a moving window filter. Crown closure is calculated directly from the crown image after filtering by a moving window method. Tree point patterns are characterized by locating the lightest spots on the spectral topography of tree crowns with a moving window filter. The computational complexity tends to increase from spectral to topographic analysis. However, the computationally simpler operation, the spectral filter (background and BRI filter), is used first to select and eliminate the candidate pixels for background as well as shadows. Only these candidate pixels representing tree crowns are processed by the moving window filter. The data volume is already significantly decreased when topographic analysis is

used to measure crown closure and tree point pattern. Such step-wise analysis makes the entire procedure efficient.

A set of threshold parameters must be determined for filtering. These parameters are spectrally sensitive and are highly dependent on the exposure of each of the photos. Since the exposures for different photos are different due to the solar condition and viewing angle at the moment of photography, parameter analysis is required for processing each digitized aerial photograph. This is especially important if the photos are taken on different dates or by different cameras. Conducting statistical analysis to determine these parameters for each of the photos to be processed could be a burden if a project involves hundreds of photos. Fortunately, we deal with digital data that are formatted so that the statistical analyses can be easily automated. In this study, a set of software tools was developed in S-plus for all the required statistical analyses. As long as the procedure is completely developed with the first photo and necessary tool kits are built, later applications are quite efficient.

The main objective of this research was to evaluate the feasibility of using digitized aerial photographs to characterize and measure forest spatial patterns. Our primary emphasis was on the development of an efficient digital image processing procedure for landscape spatial pattern characterization and measurement. The technique is feasible and reasonably accurate. We believe that digital image processing of scanned aerial photographs offers great potentials of low cost, yet sound techniques with continuously evolving capabilities for landscape analysis in medium and large scales.

# Chapter 3. Estimating Forest LAI Using Landsat TM Imagery

## 3.1. Abstract

Accurate leaf area measurements are critical for estimating fluxes of carbon, solar energy, and water in forest ecosystems. The potential of using Landsat TM data to estimate LAI has yet to be explored. Two families of spectral indices have been commonly used to relate remotely sensed data to LAI, two-band ratio indices and indices derived from Tasseled Cap Transformation. Two-band ratio indices such as SR and NDVI fail to respond to higher LAI because these indices consider only two bands and discard the bulk information from other bands. Indices derived from Tasseled Cap Transformation consider all bands except the thermal, but the transformation only rotates the spectral data space to a direction for which the variance of the data are maximized. The spectral variance is the combined result of soils, moisture, vegetation types and structural characteristics, and is not optimized to any specific biophysical characteristics of the forests such as LAI. In this paper, we looked at the spectral data from another perspective. We found that each band of Landsat TM data except the thermal contains unique information related to LAI, and the total information is captured only by some combination of bands. We used multiple regression analysis to combine the multiple bands and optimize the spectral data directly for estimating LAI. Problems of multicollinearity, nonlinearity and nonadditivity in the spectral data were examined in detail and considered in model building by specifying multiplicative variables in the model. A procedure was developed to formulate composite variables based on both statistical and physical

approaches. The results showed that our model was not saturated within the LAI range of 2 to 11, which is a typical range for the conifer forests in the Pacific Northwest. A comparison between our model and commonly used vegetation indices suggested that our model is superior for the purpose of predicting LAI from Landsat TM data.

### 3.2. Introduction

Accurate leaf area measurements are critical for estimating fluxes of carbon, solar energy, and water in forest ecosystems (Waring and Schlesinger1985). Leaf Area Index (LAI, the projected total surface area per unit ground area) is the most commonly used variable amendable to both remote sensing and field measurement (Running et al. 1986, Waring et al. 1982). Numerous previous studies have demonstrated that remotely sensed data are related to forest LAI (Chen and Guilbeault 1996, Johnson et al. 1994, Peterson et el. 1987, Spanner et al. 1994, Wulder et al. 1996, Yoder and Waring 1994), and a variety of spectral indices have been developed and associated statistically with LAI using remote sensing measurements. These indices generally fall into two families, 1) band ratio indices based mainly on red (or visible) and near infrared (NIR) radiation (Spanner et al. 1990, Chen and Guilbeault 1996), and 2) Tasseled Cap Transformation derived indices (Crist and Cicone 1984). The performance of many of these indices has been simulated using radiative transfer models or tested using data from radiometers. Spectral indices are empirical, and often analyzed by regression to seek relationships between remotely sensed radiance and LAI. The main problems in estimating LAI using these indices are the saturation of the sensor response at higher LAI and the insensitivity of these indices to vegetation changes at the stand scale (Baret and Guyot 1991, Spanner et al. 1994).

Although many studies have shown correlation between spectral indices and LAI, the relation becomes asymptotic as LAI increases (Spanner et al. 1990). The strength of the correlation also decreases significantly at the stand scale as these indices become spatially insensitive (Wulder 1996).

Many efforts have been made to improve the relationship between LAI and spectral indices by either exploring the non-linear relationship between the red and near-infrared bands (Chen 1996), or introducing texture information to supplement the LAI relationship with spectral indices (Wulder 1996). Nemani el al. (1993) used mid-infrared (MIR) correction as a scalar to calibrate NDVI for degrees of canopy closure varying from zero to one. The MIR correction substantially reduced NDVI values over low LAI, but a positive curvilinear relationship still existed between the MIR corrected NDVI and LAI, and reached an asymptote at a LAI of 3-4. Chen (1996) evaluated ten two-band vegetation indices in boreal conifer forests and concluded that SR (Simple Ratio) and NDVI were still better correlated to the field data than indices not expressed as a function of SR. Some of the non-linear two-band indices employing more mathematical operations would amplify the effect of noise. These results indicated the limitations of two-band indices in predicting LAI.

The objective of this study was to explore the use of all TM bands as a means of improving the accuracy of LAI prediction from remote sensing data. The paper presents a methodology for combining all Landsat TM band data to predict LAI. We found that each band contains some unique information related to LAI, and that full information was captured only by some combination of the bands. Multiple regression was used to combine the spectral data and optimize the spectral variance to estimate LAI. We

demonstrate the necessity of dealing with violations of the regression assumptions in spectral data, and develop a procedure for reducing the effects of these on regression analysis. Problems of multicollinearity, nonlinearity and nonadditivity among the Landsat TM bands and LAI are discussed and considered in the analysis by formulating composite spectral variables. We also compare our approach to the commonly used vegetation indices. The Landsat TM image used in the study was converted to radiance, and was corrected atmospherically and topographically before estimating LAI to provide a more normalized basis for comparison among studies.

## 3.3. Methods

### 3.3.1. Study area and field measurement of LAI

The research was conducted in the Mount Bachelor volcanic chain area in Deschutes National Forest, central Oregon. The chain is approximately 27 km long in a north-south direction and 16 km wide in an east-west direction, covering 250 $km^2$. Elevation ranges from 1297 meters at the south-east edge up to about 2760 meters on the peak of Mount Bachelor. Precipitation varies from about 60 cm annually at lower elevation in the south-east to 200 cm on the peak of Mount Bachelor. Forests consist mainly of ponderosa pine (*Pinus ponderosa*) and lodgepole pine (*Pinus contorta*) at low elevation, ponderosa pine, Douglas fir (*Pseudotsuga menziesii*) and true fir (*Abies grandis* and *Abies concolor*) at mid elevation, mountain hemlock (*Tsuga mertensiana*) and lodgepole pine at high

elevation. All forest types are uneven-aged with diverse tree sizes. LAI varies spatially within the study area as a result of topographic and climatic changes.

Leaf area of different species was allometrically estimated assuming a linear relationship with sapwood cross-sectional area at breast height (Waring et al 1982). The linear conversion coefficients were adopted from Waring et al (1982). In the summers of 1994 and 1995, We surveyed sixteen circular plots of 25-meter radius. Plots were selected in different forest types based on field reconnaissance and inspection of color aerial photographs to achieve a range in projected LAI. In each plot, all trees with a diameter at breast height (DBH) exceeding 5 cm were measured. Cores were taken from each tree to measure bark and sapwood width, and then converted to cross sectional area estimates at breast height. Understory trees with DBH less than 5 cm were also measured but not

Tab. 3.1, Plot-averaged overstory biophysical characteristics for the forests

| Plot No. | Species[a] | Average[b] Height (m) | Average[b] DBH (cm) | Average[b] Density ($m^{-2}$) | Sapwood area[b] ($cm^{-2}/m^{-2}$) | LAI[b] ($m^{-2}/m^{-2}$) |
|---|---|---|---|---|---|---|
| 1 | PIPO, PICO | 13.0 | 24.9 | 0.08 | 8.9 | 1.98 |
| 2 | PIPO, PICO | 12.3 | 18.1 | 0.20 | 18.2 | 4.28 |
| 3 | PSME, ABGR, ABCO, PIPO, PICO | 17.4 | 28.6 | 0.07 | 17.1 | 7.54 |
| 4 | PSME, ABGR, ABCO, PIPO, PICO | 11.9 | 17.1 | 0.13 | 12.0 | 5.62 |
| 5 | PIPO, PIMO, PICO, ABGR, ABCO | 13.6 | 18.5 | 0.176 | 21.1 | 4.76 |
| 6 | PIPO, PIMO, PICO, ABGR, ABCO | 11.9 | 15.1 | 0.19 | 21.4 | 5.63 |
| 7 | PIPO, PICO | 14.8 | 22.7 | 0.07 | 15.3 | 3.52 |
| 8 | PIPO, PICO | 19.3 | 32.1 | 0.04 | 13.9 | 3.39 |
| 9 | PIPO, PICO, ABGR, ABCO | 13.4 | 15.1 | 0.25 | 14.1 | 2.67 |
| 10 | PIPO, PICO, ABGR, ABCO | 13.1 | 16.0 | 0.18 | 19.1 | 5.71 |
| 11 | TSME, ABGR, ABCO, PICO | 13.7 | 22.9 | 0.18 | 29.4 | 5.15 |
| 12 | TSME, ABGR, ABCO, PICO | 14.8 | 27.5 | 0.13 | 27.9 | 5.04 |
| 13 | PIPO, PSME, ABGR, ABCO, PICO | 9.7 | 15.6 | 0.13 | 8.5 | 3.31 |
| 14 | PIPO, PSME, ABGR, ABCO, PICO | 10.1 | 17.3 | 0.08 | 6.2 | 2.41 |
| 15 | PIPO, PSME, ABGR, ABCO, PICO | 13.8 | 20.8 | 0.15 | 32.3 | 11.35 |
| 16 | PIPO, PSME, ABGR, ABCO, PICO | 15.8 | 23.6 | 0.10 | 21.6 | 8.35 |
| 17 | bare soil | 0 | 0 | 0 | 0 | 0 |
| 18 | bare soil | 0 | 0 | 0 | 0 | 0 |

a) Species abbreviations are as follow: PIPO = ponderosa pine (*Pinus ponderosa*), PICO = lodgepole pine (*Pinus contorta*), PIMO = western white pine (*Pinus montecola*), ABGR = grand fir (Abies grandis), ABCO = white fir (Abies concolor), TSME = Douglas fir (*Pseudotsuga menziesii*), PSME = mountain hemlock (*Tsuga mertensiana*).

b) These values are plot averages calculated using filed data.

cored. Tab. 3.1 shows the descriptive statistics of each plot and LAI estimates from sapwood measurements. The sample plots were located on the precisely geocoded Landsat TM imagery utilizing a differential GPS unit. The location accuracy of GPS signals was tested in field considering the influence of overstory. Relative positioning error was less than 2 meters. Two additional plots with no vegetation cover (LAI=0) and typical soil conditions were also located to provide contrast.

### 3.3.2. Processing of Landsat TM data

A Landsat-5 TM image acquired on September 2, 1992 (the month in which field data were collected) was used in the study. Although the collection of field data did not coincide with the year of TM data acquisition, year-to-year variations in conifer LAI are considered minor (Spanner 1990). There was no obvious evidence of disturbance in these plots between the time when the satellite image was acquired and the time of field survey. A step-wise image processing procedure was applied to the raw image (Fig. 3.1):

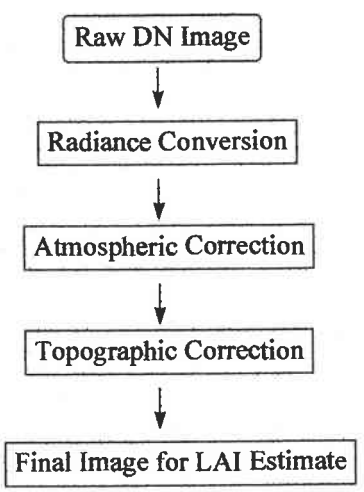



Fig. 3.1, Step-wise image processing procedure

### 3.2.2.1. Conversion of digital numbers (DNs) to at-satellite spectral radiance.

Conversion from DNs to at-satellite spectral radiance is often applied to Landsat TM data before the calculation of vegetation indices (VI) (Nemani and Running 1993, Spanner et al. 1990). The conversion provides a basis for more normalized comparison of data in single scenes, or between images taken on different dates and/or by different sensors. We used the equations published by Markham and Barker (1986) for conversion. A comparison of the radiance data vs. DN data with regard to estimating LAI was conducted to evaluate the gain of the conversion.

### 3.2.2.2. Atmospheric correction

Dark-lake pixel subtractions were applied to all spectral bands in order to eliminate the path radiance term as suggested by many authors (Ekstrand 1996, Spanner 1990), thereby reducing atmospheric influence on the data. Radiance measured by TM 4, 5 and 7 for clear lakes is almost entirely path radiance because there is virtually no upwelling radiance from water at these wavelengths. The assumption of complete absorption of radiation by clear lakes is not valid for visible bands 1, 2 and 3 (Ahern et al. 1987, Gordon 1987, Kowalik et al. 1983). In this study, the water radiance in TM bands 1 through 3 was corrected to 1.23, 0.72, and 0.12 $mW.ster^{-1}.cm^{-2}.\mu m^{-1}$ respectively, and in TM bands 4, 5, and 7 to 0, according to results presented by Bukata el at. (1983) for water with very small concentrations of suspended sediment and phytoplankton. These

characteristics are assumed to be similar to the conditions in the lakes used for calibration in this study.

### 3.2.2.3. Topographic correction

Topographic variations typical of mountainous terrain cause differences in the amount of solar radiation illuminating stands on different slopes and aspects, and must be corrected. An ideal slope-aspect correction removes all topographically induced spectral variation so that two objects having the same reflectance properties show the same radiance despite their different orientation to the sun. As a visible consequence the three-dimensional relief of a scene gets lost and the image looks flat. We used an empirical approach primarily based on cosines of the incident illumination angles (*i*) to remove the topographic effects (Teillet et al. 1982). A 30 meters DEM available from the USGS was used to generate data sets of slope and aspect. The cosine of surface normal incident angle was calculated by means of slope, aspect, and appropriate solar illumination geometry based on the equation given by Smith et al. (1980). Linear and non-linear regression analyses were used to seek the relationship between *cos i* and spectral radiance for different bands (Fig. 3.2a and Tab. 3.2). *cos i* was significantly linearly correlated with the visible (TM 1, 2 and 3), mid-infrared (TM 5 and 7) and thermal infrared (TM 6) bands (p-value = 0.01). However, the relationship between *cos i* and the near infrared band (TM 4) was significantly non-linear. In the final application, a quadratic polynomial model was applied to the near infrared band and linear models were applied to all other bands to

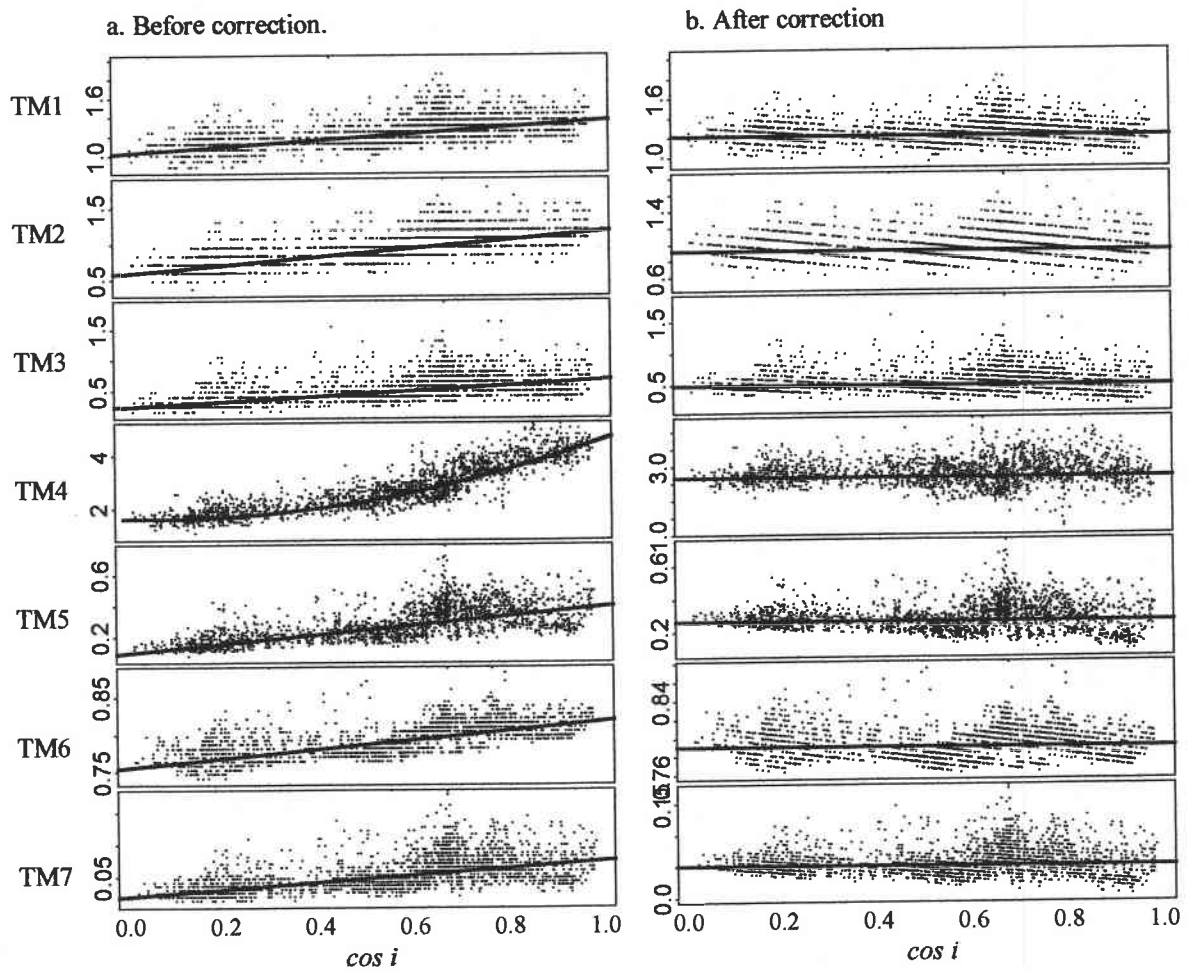



Fig. 3.2. Scatter plot of seven Landsat TM bands versus cosine of sun incident angle for uncorrected and empirically corrected images. Fitted lines in Fig. 2a were used in correction. Data were sampled from the study area to display an extreme range of *cos i*.

remove the trends of spectral radiance against *cos i*. After topographic correction, the spectral radiance for different bands was independent of *cos i*, with correlation close to 0 (Fig. 3.2b and Tab. 3.2). We used the corrected radiance data in all analyses.

Tab. 3.2, Topographic correction of Landsat TM data.

| Spectral Bands | Correlation (*cos i* vs. radiance) Before correction | After correction | Correction models | Multiple R-square | p-value |
|---|---|---|---|---|---|
| TM 1 | 0.562 | 1.71e-09 | y = 0.4122*x + 0.9610 | 0.31 | 0 |
| TM 2 | 0.666 | -6.27e-09 | y = 0.6989*x + 0.4906 | 0.44 | 0 |
| TM 3 | 0.496 | -3.99e-09 | y = 0.5274*x + 0.1652 | 0.25 | 0 |
| TM 4 | 0.864 | 1.20e-09 | $y = 4.8607*x^2 - 1.8901*x + 1.8137$ | 0.80 | 0 |
| TM 5 | 0.651 | -8.33e-12 | y = 0.3603*x + 0.0447 | 0.42 | 0 |
| TM 6 | 0.657 | 5.47e-08 | y = 0.0748*x + 0.7444 | 0.43 | 0 |
| TM 7 | 0.504 | -4.04e-09 | y = 0.0676*x + 0.0094 | 0.25 | 0 |

### 3.3.3. Optimizing multi-band radiance data for estimating LAI

To model pixel level radiance, the total radiance in a pixel can be simplified and represented by a linear process where pixel-level radiance is computed as a linear combination of the reflectance values of individual radiometric elements in that pixel (Gross and Schott 1996, Peddle et al. 1996, Hall 1995). The assumption is reasonable when there are no inter-material interactions. The effective radiance $L_{(\lambda)}$ of a pixel at the sensor can be modeled as a linear sum of radiance due to N radiometric elements in that pixel.

$$L_{(\lambda)} = \sum_{i=1}^{N} r_{(i,\lambda)} f_{(i)} + \varepsilon(\lambda) \qquad (3.1)$$

where $f_{(i)}$ is the fraction of the $i$ th radiometric element in a pixel, $r_{(i,\lambda)}$ is the Landsat spectral radiance of the $i$ th radiometric element detected by the sensor in spectral wavelength $\lambda$, $\varepsilon(\lambda)$ is the error term. $r_{(i,\lambda)}$ can be treated as constant if the assumption of no inter-material interactions is valid. The radiometric elements of a scene can be directly related to biophysical characteristics, such as leaf area in this study. With Landsat

TM data, the dimension of $L_{(\lambda)}$ is seven. *Eq.*3.1 can be solved with a least square solution.

$$F = (R^T R)^{-1} R^T L \tag{3.2}$$

Based on *eq.*3.2, two approaches can be used to seek a library of biophysical characteristics $F$ out of the spectral data $L$ that minimize $\varepsilon(\lambda)$ in *eq.*3.1. First, the spectral matrices of $L$ can be decomposed for a best fit of $F$ given the condition that the physical elements of $F$ are known (Hall 1995 el al., Gross and Schott 1996). Principal component analysis is a pure statistical approach of doing that. Tasseled Cap transformation successfully relates the spectral matrices $L$ to physically meaningful elements (Jackson 1983). However, this approach does not optimize the spectral data directly for estimates of specific biophysical attributes, such as LAI. Another alternative is to use multiple regression to optimize the spectral data directly for estimating a biophysical attribute. As we assumed $R$ to be constant matrix of element $r_{(i,\lambda)}$, LAI, being one element of $F$, can be linearly estimated by radiance in different wavelengths. We used this approach in the study.

Unfortunately, remote sensing data seriously violate the assumption of *eq.* 3.2 for a least square solution because of multicollinearity among bands (Fig. 3.3), nonlinearity between LAI and spectral data (Fig. 3.4), and nonadditivity of the error term $\varepsilon(\lambda)$ associated with all seven bands. Multicollinearity is a problem because the spectral bands of Landsat TM data are highly correlated in visible and mid-infrared bands (Fig. 3.3). Band 4 is independent of all other bands and the correlation between band 6 and the other bands is low. The major effect of multicollinearity is to increase the standard errors of

coefficient estimates for different bands in regression model. Nonlinearity refers to the fact that the relationship between radiance values and LAI varies with the level of LAI.

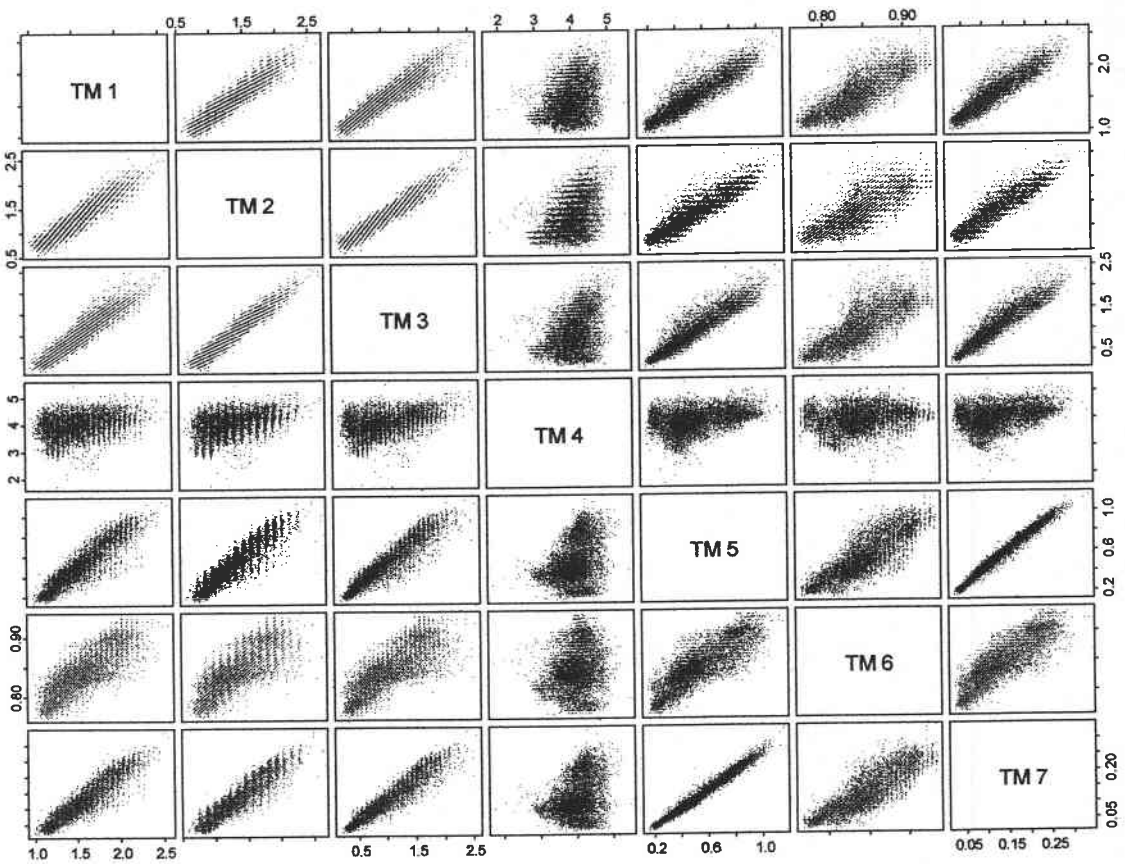



Fig. 3.3, Relationships among the seven spectral bands of Landsat TM image. Linear scatter plots indicates high correlation between bands. Elliptical scatter plots indicate low correlation between bands. Data were from a subset of the scene covering a forested landscape of the study area.

For example, an inverse curvilinear relationship exists between LAI and the radiance of visible bands (Fig. 3.4). Nonadditivity occurs because the relationship between LAI and the radiance of a TM band may vary with the level of other TM bands. Many studies have revealed relationships between LAI and vegetation indices such as the simple ratio (SR) and NDVI (Chen and Guilbeault 1996, Johnson et al. 1994, Peterson et el. 1987, Spanner

et al. 1994, Wulder et al. 1996, Yoder and Waring 1994), which implies that the amount of change in LAI detected in band 4 radiance is associated with the value of band 3 radiance.

Multicollinearity, nonlinearity and nonadditivity are very common problems in remotely sensed multispectral data. One efficient solution is to transform the correlated band data using principal component analysis (Lillesand and Kiffer 1994) or Tasseled Cap transformation (Crist and Cicone 1984). However, these algorithms usually rotate the data space to the direction of maximum variation. Since spectral variance results from the combined variations in soils, moisture, vegetation types and structural characteristics (Crist and Cicone 1986, Warren and Spies. 1992), these algorithms are not an efficient way to improve predictability of a specific forest biophysical attribute such as LAI. In this paper, we use an alternative approach by combining two or more highly correlated bands into single variables, such as a multiplicative term, and then directly relate these composite variables to LAI. In most cases, the nonlinearity is inflated by multicollinearity and nonadditivity of the data and therefore can be dealt with using composite variables (Berry and Feldman 1985). If however, nonlinearity still exists with the specification of multicollinearity and nonadditivity in the model, nonlinear interactive terms, such as multiplicative terms of the exponential and/or quadratic band radiance are necessary. With this approach, a multiple regression analysis can be conducted to extract LAI information from multi-band TM data by specifying band composites as independent variables in the regression model. With seven band TM data, there are many possibilities for formulating such multiplicative variables. We can choose these terms based on both physical and statistical approaches. We can begin by choosing only these composite terms that are

physically meaningful, and dropping terms that do not make any sense physically, such as the multiplicative terms of three or more bands. We can also select variables for the regression model with a statistical approach by following the step-wise procedure of adding or dropping variables one by one based on statistical significance. With a limited data set (direct measurement of LAI is usually expensive), the forward selection is preferred to start by adding the "best" explanatory variables into the model step by step.

### 3.4. Data analysis and results

The data analysis consists of three main parts. First, We begin by examining the performance of individual bands for estimating LAI, and demonstrate the necessity of combining the information from multiple bands for LAI estimates. Second, we describe the process of selecting appropriate multiplicative variables to optimize the spectral data for estimating LAI. Composite terms were chosen based on physical and statistical approaches. Third, we compare the results from the approach used in this study with the commonly used vegetation indices, NDVI, SR and Tasseled Cap Transformation derived indices. Finally, we compare the results obtained using atmospherically and topographically corrected data with those using the original digital number data to evaluate the gains of image correction in predicting LAI from Landsat TM data. We used ERDAS IMAGINE version 8.3 to conduct digital image processing, and used S-Plus Version 4.0 to conduct all data analysis.

### 3.4.1. Examining the performance of individual variables for Estimating LAI

Fig. 3.4 shows the relationship between LAI and individual bands and Tab. 3.3 shows the summarized statistics. The average radiance value in band 4 is the highest. In visible bands, the average radiance value decreases in the order band 1, 2 and 3. The average radiance value in mid-infrared bands is much smaller than in visible bands. The response of individual TM bands to changes in LAI is generally weak. The variance mainly exists in

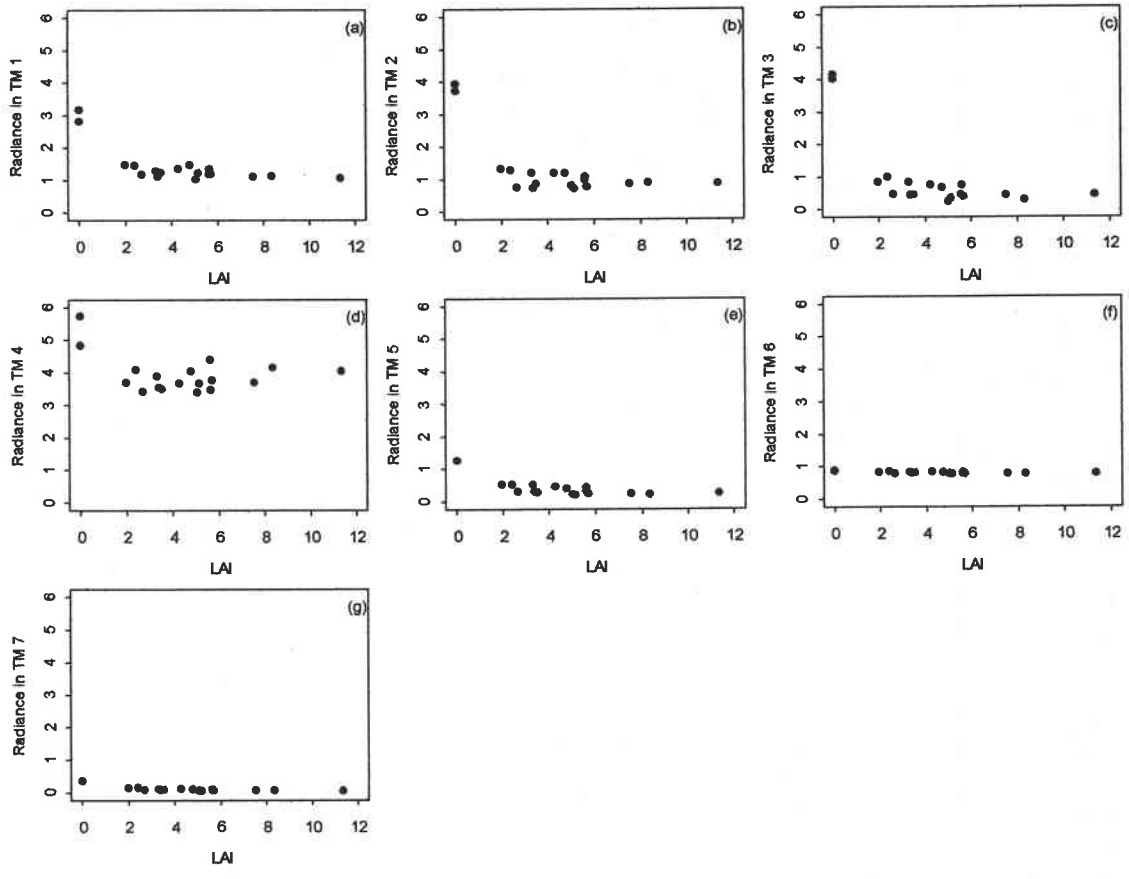



Fig. 3.4, The relationship between LAI and seven landsat TM spectral bands. The radiance axes in all scatter plots are in the same scale for comparison. Notice that band 4 positively responds to the change in LAI at higher level. Band 7 has the lowest value. Band 6 is the least sensitive to change in LAI.

Tab. 3.3, Summarized statistics of band radiance data in response to LAI. The contrast points with LAI=0 were excluded in computations to avoid nonlinearity.

| Band Number | Average Radiance | Variance | Correlation with LAI | Linear regression Between band radiance and LAI | |
|---|---|---|---|---|---|
| | | | | R-square | P-value |
| TM Band 1 | 1.245 | 0.021 | -0.568 | 0.322 | 0.022 |
| TM Band 2 | 0.976 | 0.044 | -0.363 | 0.132 | 0.167 |
| TM Band 3 | 0.558 | 0.049 | -0.486 | 0.236 | 0.056 |
| TM Band 4 | 3.774 | 0.087 | 0.365 | 0.134 | 0.164 |
| TM Band 5 | 0.334 | 0.013 | -0.581 | 0.338 | 0.018 |
| TM Band 6 | 0.800 | 0.000 | -0.059 | 0.003 | 0.224 |
| TM Band 7 | 0.072 | 0.001 | -0.194 | 0.353 | 0.015 |

shorter wavelengths (bands 1, 2, 3, and 4) with band 4 having the largest, and bands 5 and 7 the smallest variances. Band 6 is virtually independent of LAI (Fig. 3.4f). The radiance is negatively correlated to LAI for all bands except band 4, the near infrared. These relationships are consistent with physical measurements (Yoder and Waring 1994, Yoder and Pettigrew-Crosby 1995) and other observations (Fassnacht 1997, Nemani el al. 1993, Spanner 1990). However, previous studies of the relationship between band 4 radiance and LAI have yielded contradictory results. Spanner el al. (1990) reported a negative relationship between band 4 and LAI in the range 1 to 16 $m^2/m^2$. Nemani el al. (1993) reported a negative relationship between band 4 and LAI when LAI was less than 3-4 $m^2/m^2$ and a positive relationship when LAI was larger than 3-4 $m^2/m^2$. Fassnacht (1997) reported a positive relationship between band 4 and LAI in the range 1.4 to 8.4 $m^2/m^2$. In this study, we lack data points within LAI range 0-2, however, with the aid of the two contrast points where LAI is zero, our result is consistent with that reported in Nemani el al. (1993). Yoder and Waring (1994) measured reflectance physically in a lab-controlled environment and reported a very small increase of visible reflectance in green wavelengths (500-600nm) when LAI increased. This response was not detected in our field data.

Generally, a high average radiance with large variation against changes in LAI is best for estimating LAI. Lower radiation values cause the sensor to be insensitive to changes in ground objects. Lower variance of the radiance data also means a lower spectral resolution of the sensor in responding to changes on the ground. The summarized statistics in Tab. 3.3 indicate that visible and near infrared bands give most of the information about forest LAI. Band 4 is the most important in detecting LAI because of its high average and variance. Band 6 has a relatively high radiance value, but the variance is almost zero. Therefore, band 6 has the least information related to LAI. The lower R-square value indicates that the variation of LAI explained by each individual spectral band is very low. Each band contains unique information related to LAI, but the total information will only be captured by some combination of the bands.

### 3.4.2. Optimizing multispectral data for estimating LAI

Multiple regression with a least square fit to LAI can be used to optimize multispectral data for the purpose of estimating LAI. Incorporating all spectral bands into a regression model must be performed considering multicollinearity, nonlinearity and nonadditivity of the spectral data. As we discussed in the previous section, these problems can be solved by combining two or more bands that are highly correlated into multiplicative terms. We formulated multiplicative variables using both physical and statistical approaches. First, individual bands were combined into multiplicative terms based on their response to changes in LAI, i.e. a band correlating negatively with LAI must be combined with another band having the same response. Therefore, bands 1, 2, 3,

5, and 7 can be combined into multiplicative terms with one another, but band 4 can not. As an exception, a multiplicative term was also formulated using band 4 and 2 even though the correlation between these two bands and LAI was opposite. This was done because green leaves reflect both near infrared and green wavelengths. Later in our analysis, we found that this term was significant in the model. Band 6 was not considered in this process because it is independent of changes in LAI (Fig. 3.4f).

If we consider only these multiplicative terms composed of two bands, there will be eleven possible combinations (Tab. 3.4). These multiplicative terms have both statistical and physical meanings. Statistically, they represent the interactions among bands that create nonadditivity, nonlinearity and multicollinearity in spectral data (Berry and Feldman 1985). Physically, they magnify the sensitivity of the spectral data from individual bands to changes in LAI. Combining two bands to magnify the sensitivity of spectral data has been used widely in estimating LAI. The commonly used index, SR, has a more significant correlation with LAI than any individual band because dividing band 4 by band 3 magnifies the spectral response to changes in LAI (compare Fig. 3.4c and 3.4d. also see Spanner et al. 1994, Spanner et al. 1990, Running et al. 1986). NDVI, a modification of SR, was formulated as (NIR-R)/(NIR+R) (Deering el al. 1975). NIR-R is more sensitive to changes in LAI than NIR and R bands alone because the composite variable NIR-R magnifies the spectral response to LAI change (compare Fig. 3.4c, 4d and 3.5a). However, NIR+R is more or less stable when LAI changes because NIR and R have the opposite relation to LAI with NIR the positive and R the negative. Dividing NIR-R by a stable variable of NIR+R will naturally smooth the response of NIR-R to change in

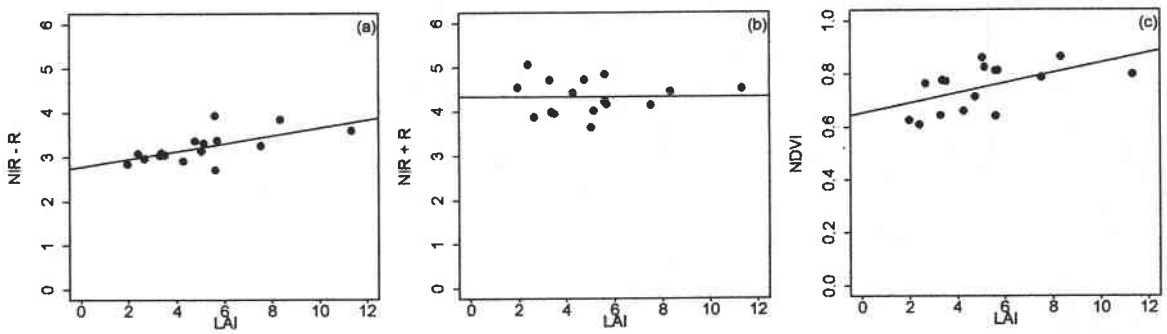



Fig. 3.5, Dissection of index NDVI. Composite variable NIR-R is sensitive to LAI change, but NIR+R is relatively stable. The ratio between NIR-R and NIR+R tends to smooth the response of NIR-R to LAI change. The lines in (a), (b) and © are a least square fit of the data.

LAI. This is the reason why NDVI becomes saturated when LAI is higher than 3-4. The difference between our multiplicative approach and the ratio approach is that we formulated the composite terms by combining bands that have the same responses to LAI (either negative or positive correlation), whereas the ratio approach combined bands that have opposite responses (Nominator positive and denominator negative). Tab. 3.4 lists the summarized statistics of all possible combinations of multiplicative terms.

Tab. 3.4, The two-band multiplicative terms formulated to optimize spectral data for LAI estimate

| Composite Variables | Variance | Correlation with LAI | Linear regression Between individual variable and LAI | |
|---|---|---|---|---|
| | | | R-square | P-value |
| $r_1 * r_2$ | 0.163 | -0.458 | 0.210 | 0.074 |
| $r_1 * r_3$ | 0.130 | -0.520 | 0.271 | 0.039 |
| $r_1 * r_5$ | 0.037 | -0.590 | 0.348 | 0.016 |
| $r_1 * r_7$ | 0.002 | -0.595 | 0.354 | 0.015 |
| $r_2 * r_3$ | 0.124 | -0.478 | 0.228 | 0.061 |
| $r_2 * r_4$ | 0.857 | -0.222 | 0.049 | 0.041 |
| $r_2 * r_5$ | 0.037 | -0.523 | 0.274 | 0.037 |
| $r_2 * r_7$ | 0.002 | -0.542 | 0.294 | 0.030 |
| $r_3 * r_5$ | 0.023 | -0.540 | 0.292 | 0.031 |
| $r_3 * r_7$ | 0.001 | -0.551 | 0.304 | 0.027 |
| $r_5 * r_7$ | 0.000 | -0.587 | 0.344 | 0.017 |

After the predictor variables were formulated, a stepwise forward selection procedure was used to identify models that best describe changes in LAI. We started with individual band variables and then added multiplicative terms into the model (18 variables in total, 7 individual bands and 11 composite variables). Residual mean square (MSE), adjusted coefficient of determination ( $R^2_{adj}$ ) and significance level (P-value) were used to guide decisions about adding variables. We also set a termination rule to limit predictor variable numbers because of the small sample size (18 data points). Forward selection continued until all variables were tested in the model. The final model is shown in *eq*. (3.3).

$$LAI = \beta_0 + \beta_1 * r_2 * r_4 + \beta_2 * r_1 * r_3 + \beta_3 * r_1 * r_5 + \beta_4 * r_3 * r_5 + \beta_5 * r_3 * r_7 \qquad (3.3)$$

Summary statistics are listed in Tab. 3.5 and Tab. 3.6. Fig. 3.6 shows the scatter plot of predicted versus measured LAI. The relationship between LAI and spectral data is significantly improved compared to SR and NDVI. No saturation problems occurred within the range of LAI used in this study. Other subset models that give small MSE and high $R^2_{adj}$ are also listed in Tab. 3.6. All variables in these models are significant at

Tab. 3.5, The coefficients of model *eq*.3.3

| Coefficients | Values | P-value |
|---|---|---|
| $\beta_0$ | 2.1613 | 0.1470 |
| $\beta_1$ | 3.0056 | 0.0002 |
| $\beta_2$ | 10.0757 | 0.0082 |
| $\beta_3$ | -41.7499 | 0.0000 |
| $\beta_4$ | 82.1499 | 0.0008 |
| $\beta_5$ | -318.9761 | 0.0016 |

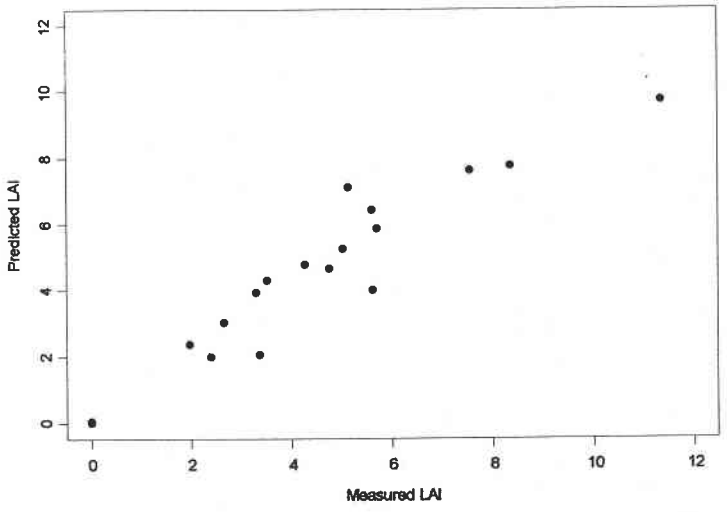



Fig. 3.6, Measured LAI versus LAI predicted by *eq*.3.3, the optimized spectral model to maximally describe the variation of LAI change. No saturation occurred in the range of LAI used in this

p = 0.05. Notice that the individual band variables were excluded in the final models, except bands 4 and 2. Bands 4 occurs independently in models 2, 4, and 5 because it is relatively independent of the other bands (Fig. 3.3). Theoretically, if predictor variables are independent, there must be only one model with no interaction term to best fit the data. Because multicollinearity exists among bands, there are multiple models (model 1, 2,

Tab. 3.6, Best-fit models. Model #1 is *eq.* 3.3. The coefficients of all predictors are significant at p=0.05.

| Model Number | Model specification | Number of Predictors | MSE | $R^2_{adj}$ | P-value |
|---|---|---|---|---|---|
| 1 | $r2*r4, r1*r3, r1*r5, r3*r5, r3*r7$ | 5 | 0.762 | 0.855 | 0.0000 |
| 2 | $r2, r4, r1*r3, r1*r5, r3*r5, r3*r7$ | 6 | 0.732 | 0.848 | 0.0000 |
| 3 | $r2+r4, r1*r3, r1*r5, r3*r5, r3*r7$ | 5 | 0.848 | 0.838 | 0.0000 |
| 4 | $r4, r1*r3, r1*r5, r3*r5, r3*r7$ | 5 | 1.034 | 0.803 | 0.0000 |
| 5 | $r4, r1*r3, r1*r5, r2*r5, r2*r7$ | 5 | 1.096 | 0.791 | 0.0001 |

3 in Tab. 3.6) which can explain similar amount of variance of the data set. Therefore, the chosen multiplicative variables might differ for different studies. The rule of thumb in choosing among models with relatively equal explaining power is that the simpler, the better. Although model 2 has the lowest MSE, it is inferior to model 1 because more variables are used.

### 3.4.3. Comparison with vegetation indices

Fig. 3.7 shows comparisons among our best-fit model and other commonly used vegetation indices. SR is linearly related to LAI, but the variance significantly increases as LAI increases, and finally becomes insensitive to changes in LAI (Fig. 3.7b). NDVI is logarithmically related to LAI and responds less sensitively to change in LAI than SR, but the variance was significantly lower than that of SR (Fig. 3.7c). The relationship between LAI and greenness derived from the Tasseled Cap Transformation is similar to that between NDVI and LAI (Fig. 3.7e). Both brightness and wetness are insignificantly correlated to LAI (Fig. 3.7d and 3.7f). The multiple regression with all three components (B, G and W) gives a better fit than the individual component (Fig. 3.7g). However, this multiple regression is still not sensitive to change in LAI at LAI > 5, and the MSE is very large. The behavior of principal components is generally similar to Tasseled Cap Transformation indices but with larger MSE. Our model provided the best fit and did not reach an asymptote over the range of LAI used in this study (Fig. 3.7a).

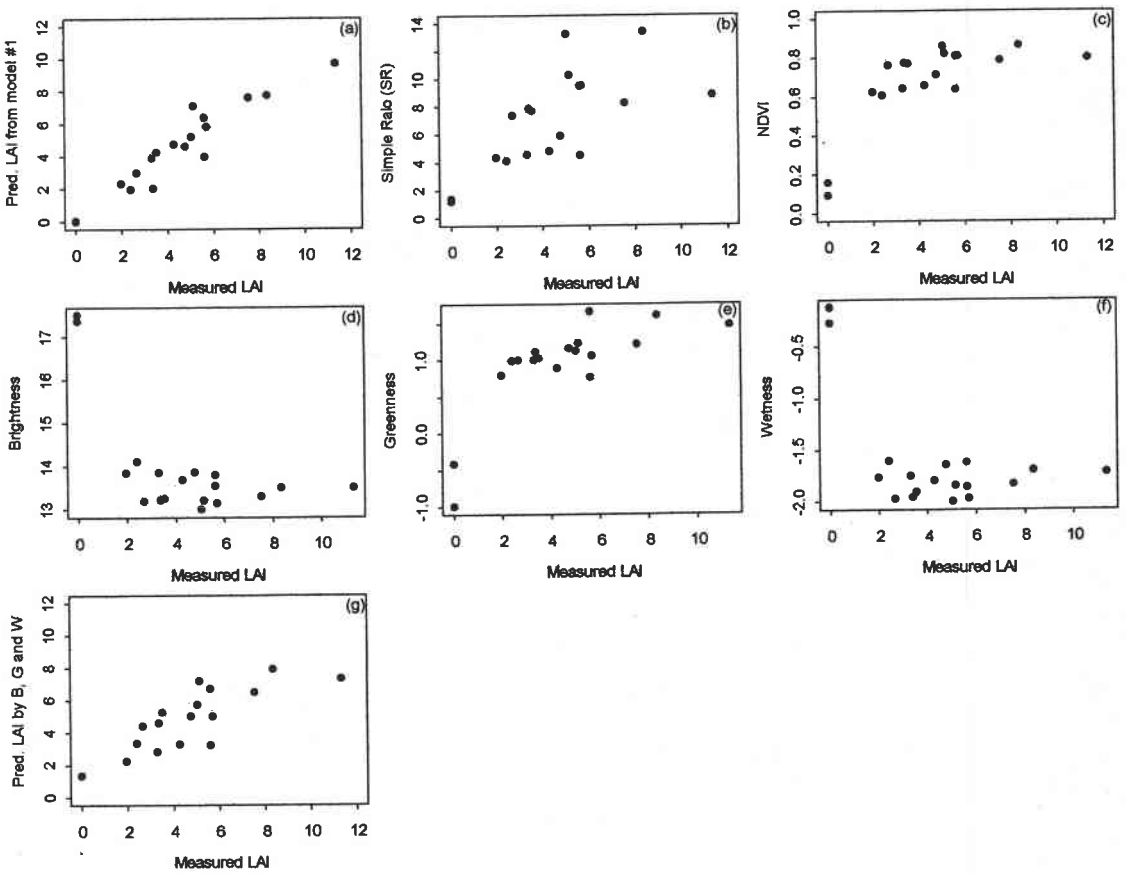



Fig. 3.7, Comparison of *eq.* 3.3 with different indices for estimating LAI. Brightness (B), greenness (G), wetness (W) were computed using the Tasseled Cap Transformation coefficients for Landsat TM5 developed by Crist and Cicone (1986). Principal components analysis was also conducted, but none of the principal components gave meaningful results, hence were not considered in the comparison.

### 3.4.4. Evaluation of the effects of DN conversion and image correction on estimating LAI

Conversion from digital numbers (DNs) to spectral radiance or exo-atmospheric reflectance provides a basis for more normalized comparison of data and methods (Markham and Barker 1986). The data from different bands are not comparable before the conversion from DNs to radiance. The DNs are merely relative numbers stretched to a

specific data structure such as the commonly used 8-bits by a remote sensor. After conversion to radiance, the data have physical meanings and units, and can be compared among bands and different sites. An evaluation was conducted to examine the value of DN conversion for estimating LAI. Fig. 3.8 shows the relationships between individual

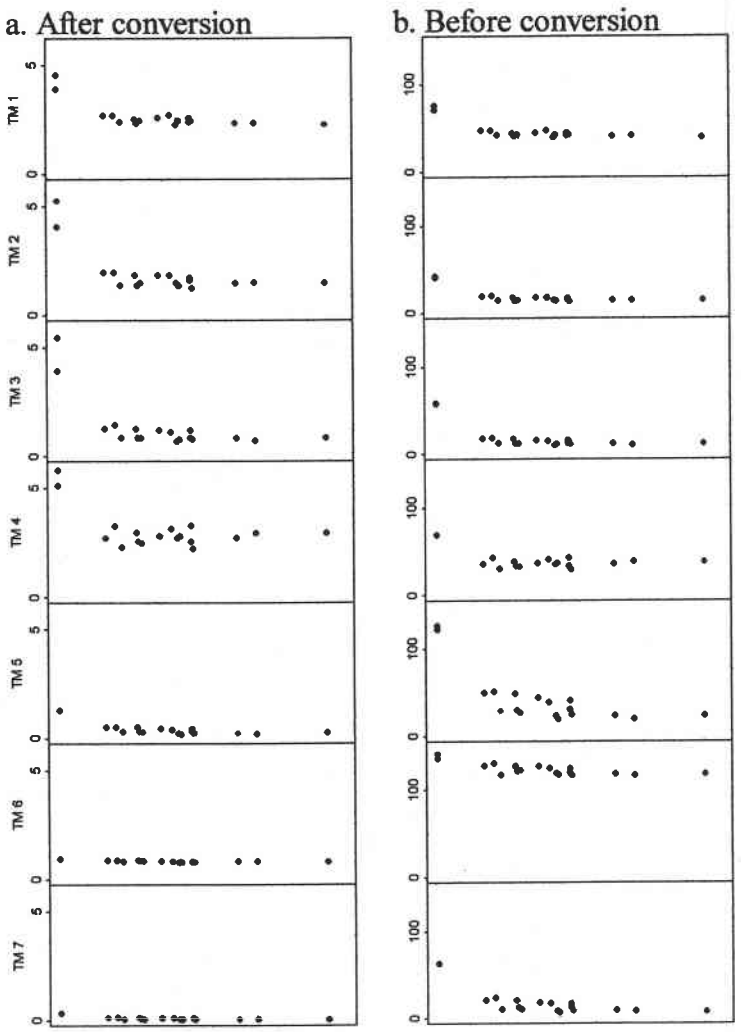



Fig. 3.8, Comparison of raw DN data, and corrected radiance data with regard to the response of individual bands to LAI. The corrections include atmospheric and topographic corrections.

bands and LAI before and after DN conversion. We found that the average and variances of thermal band 6 and mid-infrared bands (5, and 7) were greatly reduced by converting DN to radiance, while the variances in near-infrared and visible bands were increased. The radiance of band 6 become more independent of LAI. These results are expected and agree with physical explanations. Green leaves absorb visible red and blue light, reflect green and near-infrared light, and relate less to mid and thermal infrared light (Yoder et al. 1994, 1995), therefore, the variance in radiance over forest canopies with varying LAI must be larger in the visible and near-infrared bands than in the mid and thermal infrared bands. The signal strength in visible and near-infrared bands is also much stronger that that in mid and thermal infrared bands. Data conversion helps to recover the real at-satellite sensor response to ground objects and eliminate the influence of arbitrary scale with digital number.

For the purpose of accurately estimating LAI, image data must also be corrected so that the radiance will be free from atmospheric and topographic effects (Spanner 1990). However, without an appropriate understanding of the factors underlining these operations, corrections could introduce unexpected noises. A comparison was conducted to study the effects of image correction on estimating LAI. LAI of the study area was estimated using Landsat TM data before and after correction. A difference image was created by subtracting the LAI without image correction from the LAI with image correction. The hypothesis is that if there is no topographic effect on estimating LAI, the difference must be close to zero and independent of sun incident angle. Fig. 3.9 shows the relationship between sun illumination angle and the difference in LAI calculated by the two approaches. The difference in LAI between corrected and uncorrected images is

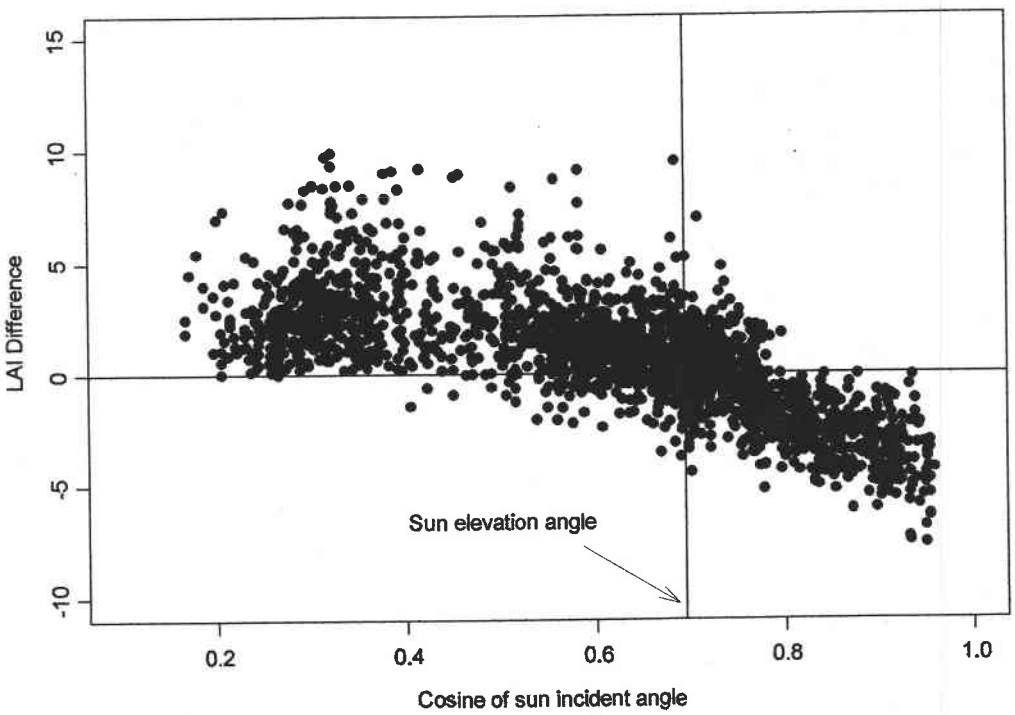



Fig. 3.9, The relationship between cosine of sun incident angle and the difference of LAI estimates after and before image correction, showing the effects of topographic correction on LAI estimates.

significantly correlated to sun incident angle. With the uncorrected image, LAI is estimated lower in pixels where the sun incident angle is greater than 46° (the sun elevation angle for the TM image used in this study), and LAI is estimated larger in pixels where the sun incident angle is smaller than the sun elevation angle. The more the sun incident angle deviates from the normal angle (sun elevation angle), the bigger the difference. LAI estimates of uncorrected and corrected images differ by as much as 5 in pixels where the sun illumination angle is greater than the sun elevation angle, and 10 in pixels where the sun illumination angle is smaller than the sun elevation angle. The LAI of pixels with the sun illumination angle equal to the sun elevation angle is generally unchanged before and after correction.

### 3.5. Conclusions

We have demonstrated the use of multiple regression to optimize Landsat TM data to improve estimates of LAI. Unlike the traditional approaches of formulating a spectral index with band 3 and 4 first and then seeking the relationship between that index and LAI, our approach used all spectral bands to build a multiple regression model, and combined the interrelated bands into a multiplicative term to reduce multicollinearity among bands. A step-wise procedure was developed to apply this approach to estimating forest LAI. Physical explanations were provided to support the model and explain the results. By successful application, we conclude:

1) Each band of Landsat TM data except the thermal contains unique information related to LAI. Visible and near-infrared bands are more related to LAI than mid-infrared bands. The thermal band is independent of LAI at the scale of this study. The best estimate of LAI is provided only by some combination of the bands. Multiple regression analysis can be used to develop the best-fit model.

2) Multicollinearity, nonlinearity and nonadditivity commonly exist in Landsat TM data, and must be considered in regression analysis and interpretation of results. Highest correlation is found among visible and mid-infrared bands. Correlation between the thermal band and all other bands is lower. Correlation between the near infrared band and other bands is the lowest. The near-infrared band is relatively independent of all other bands and most sensitive to changes in LAI, therefore, it is the most important single band in estimating LAI.

3) Introducing Multiplicative terms into a regression model is an efficient way of combining multiple bands for estimating LAI and dealing with multicollinearity in Landsat TM data. Our model shows no asymptotic trend between spectral data and LAI within the range of 1.98 to 11.35. The chosen multiplicative variables might differ for different sites, but the procedure of formulating multiplicative variables and selecting the best-fit model should be broadly applicable.

4) The multiple regression model used in this study is best for predicting LAI from Landsat TM spectral data rather than interpreting the relationship between LAI and spectral bands. With a limited data set, the interpretation of these variables and their coefficients are not reliable. They are subject to change with different specifications of the model. Therefore, any inference about the performance of these variables with regard to LAI must be carefully conducted with the aid of physical measurements.

5) Image correction is required for estimating LAI in mountainous area. The difference between the results using corrected and uncorrected image data is significant. In this study, LAI was over-estimated by as much as 10 $m^2/m^2$ in pixels with small sun incident angle and underestimated by as much as 5 $m^2/m^2$ in pixels with large sun incident angle. Radiance data are also better for estimating LAI than digital numbers. Converting data from digital numbers to radiance helps to recover the real at-satellite sensor response and eliminate the influence of arbitrary scaling with digital numbers.

The main objective of this research was to seek better estimates of LAI using Landsat TM data. We believe using information from all bands achieves the full power of the Landsat Thematic Mapper for estimating LAI, and improves the relationship between spectral data and changes in LAI. We demonstrated the necessity of dealing with

multicollinearity by combining multiple bands into multiplicative terms in the model. We emphasize the use of all band information from Landsat TM data to *predict* forest LAI rather than to *explain* the relationships between LAI and individual bands. Our procedure as proposed in this study optimizes the use of multispectral data for estimating LAI. Further studies are recommended to test this approach with larger data sets in larger scales.

# Chapter 4. Analysis of Relationships Between Forest Spatial Patterns And Environment

## 4.1. Abstract

This study used GIS to integrate multiple spatial data layers for analyzing and parameterizing relationships between forest and environment in the Mount Bachelor Volcanic Chain of central Oregon. The forest spatial data layers include tree species occurrence, canopy crown closure, and leaf area index (LAI). Environmental data layers include topography (slope, aspect, and elevation), soil and geology. The objectives were to (1) analyze the spatial coincidence among forest inventory data, structural characteristics, and topographic, soil and geologic variables, (2) develop empirical relationships between forest characteristics and edaphic and topographic factors, and (3) identify the possible driving variables that shape the forest structure and spatial patterns. A logistic regression model was used to analyze species occurrence in response to topographic, soil and geologic changes. A classification tree model was developed to predict the relationship between crown closure and topographic, soil and geologic variables. A multiple regression model was used to analyze the responses of LAI to these same environmental variables. The results clearly indicated that forest composition, crown closure and LAI differ in their response to physiographic variables. Species occurrence over the lava landscape was best correlated with lava types. Crown closure and LAI were most influenced by elevation. Among the topographic variables, slope had the least correlation with species occurrence and forest structural attributes within the range of 0-

30%. The study demonstrated the efficiency of integrating multiple GIS data layers to perform spatial analysis and parameterize relationships between forests and environment.

## 4.2. Introduction

The high lava plain of central Oregon is characterized by young lava flows of moderate relief interrupted by scattered cinder cones and lava buttes. It contains forests of nationwide significance (Franklin and Dyrness, 1973). The regosolic soils developed on pumice support open coniferous forest vegetation and apparently enable variety of species such as ponderosa pine (*Pinus ponderosa*), lodgepole pine (*Pinus contorta*), Douglas-fir (*Pseudotsuga menziesii*) and true fir (*Abies grandis* and *Abies concolor*) to extend eastward. The frequent coincidence of the boundaries of ponderosa pine forest and pumice soil areas indicates a unique relationship between forests and physiography in this area. The ecology of forests growing on lava flows of this area has not been studied, yet large portions of the forest have been logged at various intensities over the last 40-50 years (Popovich 1991). Resource specialists have identified many issues associated with managing these forests. The need for more information in addition to an inventory of the forest under management arises because maximum production of desired commodities can be attained only with an understanding of the productive capacity and management limitations of the land. Problems to be resolved include such questions as what relationships exists between forest structures and physiographic variables, what driving variables shape forest structures and spatial patterns, and how the changes of forest spatial

patterns affect the environment and forest regeneration. Answering these questions requires analyses of multiple data layers at multiple dimensions and scales.

Understanding landscape surface processes and attempting to quantify relationships has been a dominant research theme over the past decade (Steyaert et al. 1997). Traditional ecological studies have assumed what occurs at one point is independent of what occurs at others (Johnson, 1990), therefore, providing no base for extrapolating over a landscape. The ability to array spatial data in different formats and use these as tools for studying spatial process is at the root of current interests in remote sensing and GIS (Leckie 1990, Goodchild 1987). Remote sensing provides the primary means of describing forest structures at various scales within the full extent of a study area (Strahler and Woodcock, 1986). Detailed structural characteristics including crown closure, crown density and leaf area index can be estimated by analyzing digital aerial and Landsat TM imagery (Huang 1997a and 1997b, Spanner et. al. 1990). These structural attributes have been correlated with many ecosystem processes such as site water balance (Grier and Running 1977), canopy light interception (Yoder and Waring 1994), and net primary production (Gower et al., 1992). The analyses of spatial processes usually involve integrating multiple data layers of forest structures with environment. A GIS can be used to manipulate the spatially distributed data, and to determine spatial coincidence of physical and biological features over a broad range of spatial and temporal scales (Johnson 1990).

In this study, we use GIS techniques to integrate all available data layers and analyze the responses of forest spatial patterns to physiographic variables in the Mount Bachelor volcanic chain, central Oregon. Our objectives are to (1) detect the spatial coincidence

among forest structural characteristics, and topographic, soil and geologic variables across the volcanic landscape, (2) develop empirical relationships between forest characteristics and physiographic factors, and (3) identify the important physiographic features that shape the forest structure and spatial patterns. A better understanding of forest-environment interactions contributes to better decision making on maintaining healthy forest ecosystems.

## 4.3. Methods

### 4.3.1. Study area

The research was conducted in the Mount Bachelor volcanic chain, Deschutes National Forest, central Oregon. The area is a geologically young and complex volcanic region of mostly Pleistocene age lying on the east flank of the Cascade Range (Scott and Gardner, 1992). A series of uneroded lava cones and cinder cones form a chain of 27 miles long from Mount Bachelor to Lookout Mountain. Most of the area is covered by immature regosolic soils developed in deposits of dacitic and rhyolitic pumice erupted from Mount Mazama (Crater Lake) in 6,700 years ago. The coarse-textured pumice soils apparently enable Ponderosa pine (*Pinus ponderosa*) and mixed coniferous forests to extend their range eastward toward the high desert. Ponderosa pine forest is frequently observed to be coincident with lava flows. Elevation within the study area varies from 1297 to 2760 meters. Decreasing elevation accompanied by declining precipitation resulting in sharp transitions in vegetation types from the peak of Mount Bachelor

eastward to dry Lapine basin. At lower elevations, forests consist mainly of ponderosa pine (*Pinus ponderosa*) (PIPO) and lodgepole pine (*Pinus contorta*) (PICO). As elevation increases, forest type changes to ponderosa pine, Douglas-fir (*Pseudotsuga menziesii*) (PSME), grand fir (*Abies grandis*) (ABGR) and white fir (*Abies concolor*) (ABCO) mixed forest. At higher elevations, the primary tree species are mountain hemlock (*Tsuga mertensiana*) (TSME) and lodgepole pine.

### 4.3.2. Data sources

Tab. 4.1 lists characteristics of all data layers used in the study. Forest data layers include species occurrence, crown closure and leaf area index (LAI). Species occurrence

Tab. 4.1, Data layers used in the study. All layers were in raster format with 25 meters pixel size.

| Theme | Data Type | Attribute Summary | Source |
|---|---|---|---|
| Species occurrence | Binary | 0 - Absence, 1- Presence | Developed from Deschutes National Forest inventory data (DNF 1994) |
| Crown Closure | Thematic | Coded in six levels with "1" as 0-10%, "2" as 11-25%, "3" as 26-40%, "4" as 41-55%, "5" as 56-70%, "6" as 70-100% | Huang et al. (1997a) |
| Leaf Area Index | Continuous | Min=0, mean=2.33, max=12.61. | Huang et al. (1997b) |
| Aspect | Thematic | Coded into eight directions. 1=NE [0, 45), 2=EN [45, 90), 3=ES [90, 135), 4=SE [135, 180), 5=SW [180, 225), 6=WS [225, 270), 7=WN [270, 315), 8=NW [315, 360). | Developed from USGS 7.5 minute DEM data |
| Slope | Continuous | Min=1.01%, mean=7.00%, max=37.24% | Developed from USGS 7.5 minute DEM data |
| Elevation | Continuous | Min=1297 m, mean=1700 m, max=2760 m. | Developed from USGS 7.5 minute DEM data |
| Soil Type | Thematic | Total of 47 units aggregated into 7 groups (see Tab. 4.2 for detail). | Soil resource inventory, US Forest Service, Deschutes National Forest (Larsen 1976) |
| Lava Type | Thematic | Total of 42 types aggregated into 10 groups (see Tab. 4.3 for detail) | Scott and Gardner (1992). |
| Harvest | Binary | 0 - No Harvest, 1 - Harvest | Deschutes National Forest, GIS department |
| Roads | Binary | 0 - Non - roads, 1 - Roads | Deschutes National Forest, GIS department |

data were developed from inventory data available from US Forest Service, Deschutes National Forest (Deschutes National Forest, 1994). The inventory data were developed by Pacific Meridian Resource from Landsat TM image acquired in 1988 with supervised classification, consisting of spatial distributions of the five major species found in the study area (Ponderosa pine, lodgepole pine, Douglas-fir, white fir, and mountain hemlock). The data were transformed into binary format with 1 representing the presence of a species and 0 as absence . Crown closure data were developed using digitized aerial photographs with 1 meter pixel resolution (Huang et al. 1997a), aggregated into 25 meters pixel resolution with categorical attributes to minimize file size and conform with other data layers used in the study. Crown closure was coded into six levels: <10%, 11- 25%, 26 - 40%, 41 - 55%, 56 - 70%, 71 - 100%. LAI data were estimated from 1991 Landsat TM scene (Huang et al. 1997b).

Environmental data layers include topographic data developed from USGS 7.5 minute DEM, a soil resource inventory available from Deschutes National Forest (Larsen 1976), and geologic data developed by Scott and Gardner (1992). There are total of 47 different soil mapping units in the study area. We aggregated the 47 units into 7 groups according to soil material, texture and morphology (Tab. 4.2). Forty two lava types were mapped based on composition, age, and origin. Lava type were further aggregated into 10 categories based mainly on lava age and origin (Tab. 4.3). Road and past harvest data were obtained form the Deschutes National Forest. Precipitation data with appropriate resolution was not available for the study area. Since the study area is fairly small, we believe the local distribution of precipitation was mainly influenced by topography.

Tab. 4.2, Soil type stratification (see detail in Larsen, 1976)

| ID | Name | Characteristics |
|---|---|---|
| 1 | Miscellaneous types | Areas where soils are too variables to be described by a typical profile description or are variable in respect to other soil features. |
| 2 | Sandy to loamy soils over glacial till | Sandy volcanic ash and pumice on buried soils over glacial till. Dendritic drainage pattern. Rapid infiltration. Sandy loam texture |
| 3 | Sandy to loamy soils over glacial outwash and lake deposits | Sandy pumiceous volcanic ash and pumice lapilli over a buried soil on glacial outwash. Dendritic drainage pattern. Rapid infiltration. Loamy sand or sandy loam texture. |
| 4 | Loamy to sandy pumice soils over volcanic tuffs and brecias | Sandy volcanic ash and pumice lapilli over buried soils on compact pumice. Poor drainage. Rapid infiltration. Loamy sand texture. |
| 5 | Loamy to sandy pumice soils over hard basalts, andesites, and rhyolites | Sandy, pumiceous volcanic ash and pumice lapilli mixed with highly fractured lavas over sandy to loamy buried soils. Dendritic drainage pattern. Rapid infiltration. Loamy sand texture. |
| 6 | Loamy to sandy pumice soils on volcanic bedrock materials | Sandy pumiceous volcanic ash and pumice lapilli over sandy buried soil. Sandy loam texture. Rapid infiltration. Poor radial and parallel drainage pattern. Sandy loam texture. |
| 7 | Sandy loam to sandy pumice soils on cinder cones | Sandy, pumiceous volcanic ash and pumice lapilli over cindery buried soil. Radial drainage pattern. Rapid infiltration. Loamy sand texture. |

Tab. 4.3, Lava type stratification (See detail in Scott and Gardner, 1992)

| ID | Geologic Type | Characteristics |
|---|---|---|
| 1 | Glacial deposits | Upper Holocene. Till is matrix-supported boulder to pebble gravel in silty sand sandy silt matrix; nonsorted; gravel clasts angular to subround; nonbedded. |
| 2 | Young rhyolite of Rock Mesa and Devils Hill eruptive episodes. | Upper Holocene. Porphyritic, dense to pumiceous rhyolite ($SiO_2$, 72.5-73.5%). Thick, blocky lava domes and flows. Unconsolidated materials including tephra-fall, pyroclastic-flow, pyroclastic-surge and lahar deposits. |
| 3 | Products of Egan cone | Lower Holocene to upper Pleistocene. Non- porphyritic to sparsely and finely porphyritic basalt and basaltic andesite ($SiO_2$, 51-54%). |
| 4 | Products of summit cone of Mount Bachelor | Lower Holocene to upper Pleistocene. Abundantly porphyriticand glomeroporphyritic basaltic andesite lava flow ($SiO_2$, 55-57%). |
| 5 | Products of shield volcano of Mount Bachelor | Upper Pleistocene. Basalt and basaltic andesite lava flows and typhra deposits ($SiO_2$, 50-55%). Including younger lava flows and Olivine basalt lava flows. |
| 6 | Products of shield volcano of Kwolh butte | Upper Pleistocene. Basalt and basaltic andesite lava flows and typhra deposits ($SiO_2$, 50-55%). Including younger lava flows, Olivine basalt lava flows, Plagioclase basaltic andesite lava flows and basalt lava flows of Lava Lake. |
| 7 | Products of Siah chain of vents | Upper Pleistocene. Basalt lava flows and typhra deposits ($SiO_2$, 49.5-51.5%). |
| 8 | Products of shield volcano of Sheridan Mountain | Upper Pleistocene. Lava flows and typhra deposits of basalt and basaltic andesite ($SiO_2$, 52-55%). Including young, intermediate and old age basaltic andesite lava flow tephra deposits. |
| 9 | Basalt and basaltic andesite lava flow from volcanic centers in surrounding areas. | Other basalt and basaltic andesite lava flow from volcanic centers in surrounding areas that are broadly similar in age to the Mount Bachelor volcanic chain. The contents of $SiO_2$ vary from 49% to 63%. |
| 10 | Alluvial, lacustrine, and colluvial deposits | Lower Holocene to upper Pleistocene. |

### 4.3.3. Sampling design

Road coverage was buffered by 15 meters distance on either side of the roads and then overlaid with harvest and other data layers. Any pixel that fell into a road buffer or harvest area was filtered out to reduce the variance associated with human disturbances. Because there were no catastrophic fires in the study area in this century (according to the fire records from Deschutes National Forest), no fire data were considered in the analysis.

A random grid image covering the study area was created and used as a sampling template (Fig. 4.1). All pixels touched by the template grids were sampled, others were filtered out. Five percent of the sampled pixels were randomly selected to produce the final data sets for statistical analysis. There are two reasons for a random sampling. First, the spatial data were highly autocorrelated and were not suitable for parametric analysis. Random sampling can reduce spatial autocorrelation (Legendre and Fortin, 1989). Second, random sampling reduced the size of the data set from 25.5 megabytes to 1.3 megabytes, making it a size appropriate for matrix algebra and statistical analysis with most present day computers.

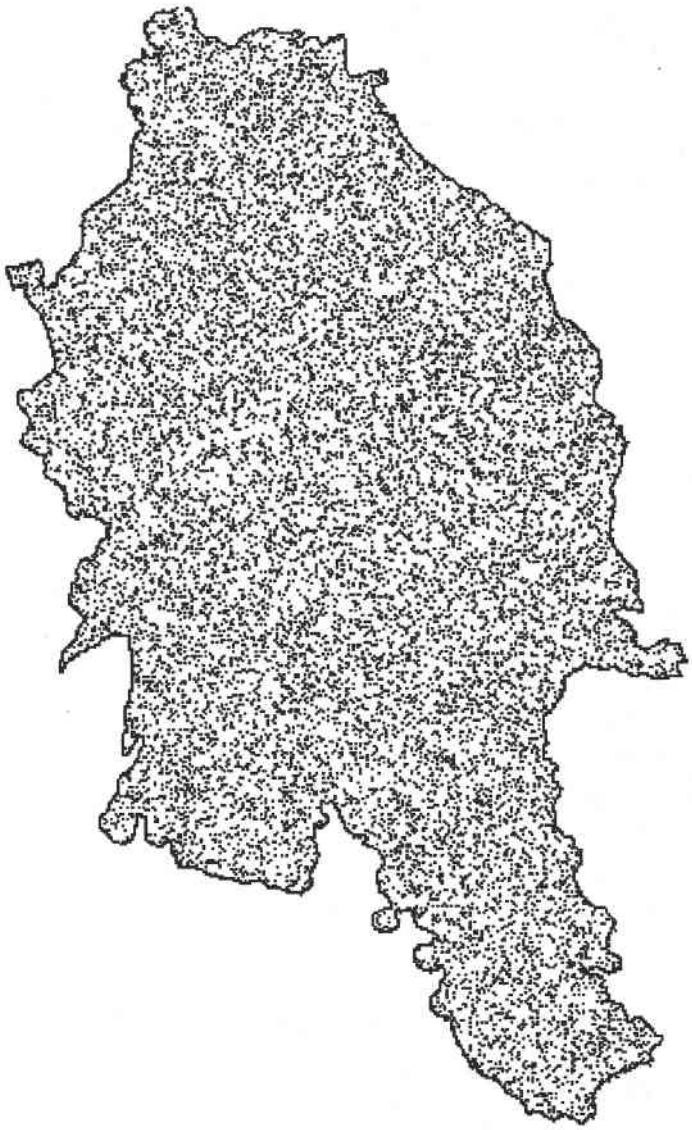


Fig.4.1, Template grid for random sampling of image data

### 4.3.4. Statistical analysis

We used S-Plus Version 4.0 (MathSoft Inc. 1997) to analyze the co-occurrence between forest variables (species occurrence, crown closure and LAI) and physiographic

variables (slope, aspect, elevation, soil type and lava type). The latter were treated as independent variables. We first explored bivariate relationships between each forest variables and the environmental variables. We then conducted parametric analyses based on the exploratory results.

#### 4.3.4.1 Analysis of species occurrence and LAI vs. environmental variables

We used a logistic regression model to relate the probability of the occurrence of ponderosa pine, lodgepole pine, white fir, Douglas-fir, and mountain hemlock to slope, aspect, elevation, soil type and lava type. Symbolically the model is

$$\log\left(\frac{prob(event)}{prob(no\ event)}\right) = \alpha + \beta_{slope}x_{slope} + \beta_{aspect}x_{aspect} + \beta_{elev}x_{elev} + \beta_{soil}x_{soil} + \beta_{lava}x_{lava} \qquad (4.1)$$

where *event* is the occurrence of a given tree species. The model states that "the occurrence of a tree species is modeled as a linear function of slope, aspect, elevation, soil type and lava type". Partial t-tests were used to evaluate the significance of each predictor in the model. We used ANOVA to analyze the deviance for sequential additions of each variable to the model, specifying chi-square to test for difference between models. We further examined the bivariate relationships between the probability of species occurrence and each of the physiographic variables by fitting a "null" model and then adding each of the predictors, one at a time. $C_p$ statistics were used to compare models that were not nested. A small $C_p$ corresponds to a better model in the sense of smaller residual deviance, which can be inferred such that the predictor is a significant variable influencing the occurrence of the species (MathSoft 1997). A multiple regression analysis was employed

to explore the relationships between LAI and environmental variables in a similar approach.

#### 4.3.4.2 Analysis of crown closure vs. environmental variables

We used a Tree-based model to analyze species occurrence in relation to topography, soil, and geology. Tree-based model is a newly developed multivariate method (Chambers and Hastie 1992) that uses binary recursive partitioning to split a data set into increasingly homogeneous subsets until it is infeasible to continue (see Appendix). Assuming crown closure is associated with aspect, slope, elevation, lava type and soil type variables:

$$y = f(x_{slope}, x_{aspect}, x_{elev}, x_{soil}, x_{lava}) \tag{4.2}$$

Where y is crown closure. The algorithm of tree-based model attempts to partition the space of predictor variables ($x$) into homogenous regions, such that within each region the conditional distribution of $y$ given $x$, $f(y|x)$, does not depend on $x$ (Chambers and Hastie, 1992). The idea is that in order to estimate crown closure from the predictors, we followed the path from the top node of the *tree*, called *root*, to a terminal node, called a *leaf*, according to the rules called *splits* (Fig. 4.2). A tree model is the collection of many such rules determined by a procedure called recursive partitioning (Appendix). We used 1% of the root node deviance as the decision criterion. If the node deviance was less than 1% of the root node deviance, the split was stopped. Unlike cluster analysis, in which the objective is to minimize within-group variance of the sampling points, the objective of a Tree-model is to maximize the probability of prediction in multivariate regression. Each split in the Tree represents a statistically significant prediction. Variables with the greatest

predictive power are split at the top nodes (the top is called the "root"). Fig. 4.2 shows an example based on a subset of our data. The split at the first (highest) node is based on elevation, indicating that variable has the greatest power to predict crown closure (in this data subset). Elevations greater than 2206 m are partitioned to the right-hand branch, and elevations lower than that to the left-hand branch. For elevations below 2206 m, the next split is based on aspect, the third splits on lava type, and so forth, with lower nodes having less explaining power than upper nodes. The numbers at the bottom correspond to levels of crown closure as given in Table 4.1.

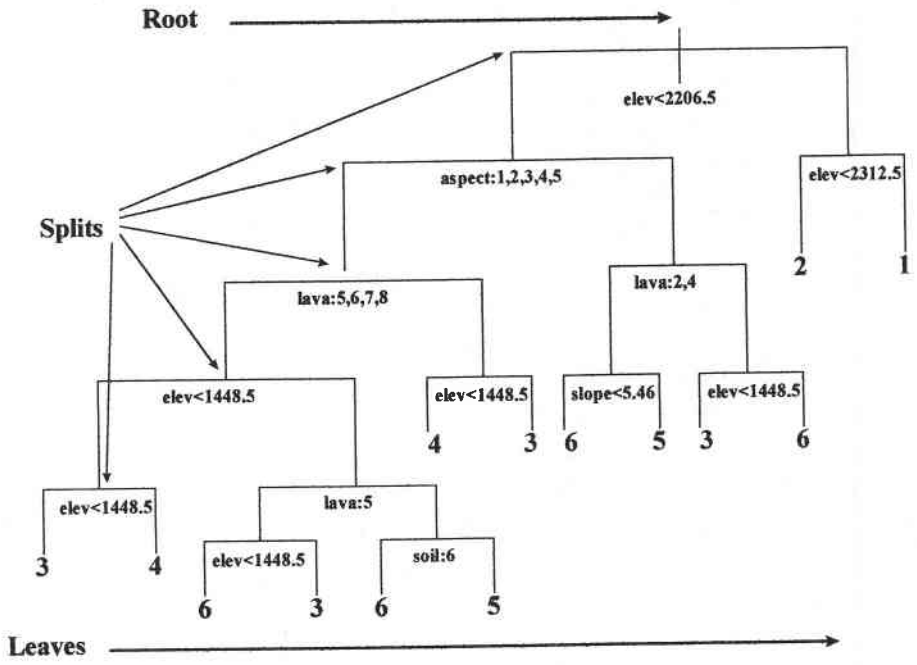



Fig. 4.2, An illustrated tree model relating crown closure to topographic, soil and geologic variables. The data used to produce this graph was a subset for illustration.

## 4.4. Analysis and results

### 4.4.1. Species occurrence and site variables

We used occurrence ratio to evaluate the distribution of different tree species against the five environmental variables. An occurrence ratio was defined as the ratio between the number of pixels on which a tree species occurred and the total number of pixels sampled in a range of the environmental variables. For example, if ponderosa pine occurred in 5 pixels on lava type 1 and the total number of pixels in lava type 1 is 10, the occurrence ratio of ponderosa pine on lava type 1 is 0.5. The change of occurrence ratio for a specific species against an environmental variable shows the influence of that environmental variable on the species. Comparison of the occurrence ratio among different species shows the relative abundance of species in the forest.

#### 4.4.1.1. Species occurrence vs. aspect

The five species responded to aspect in different patterns (Fig. 4.3). The distribution patterns of ponderosa pine was opposite to those of Douglas-fir. A larger proportion of ponderosa pine occurred on EN, ES, SE and SW aspects while a larger proportion of Douglas-fir occurred on NE and NW aspects. Higher proportion of mountain hemlock occurred on WN and WS aspect. White fir and lodgepole pine occurred more or less evenly on all aspects except that the highest proportion of white fir occurred in northern aspects while the highest for lodgepole pine is on southern aspects. These species distribution patterns shows that Douglas-fir, ponderosa pine and mountain hemlock are

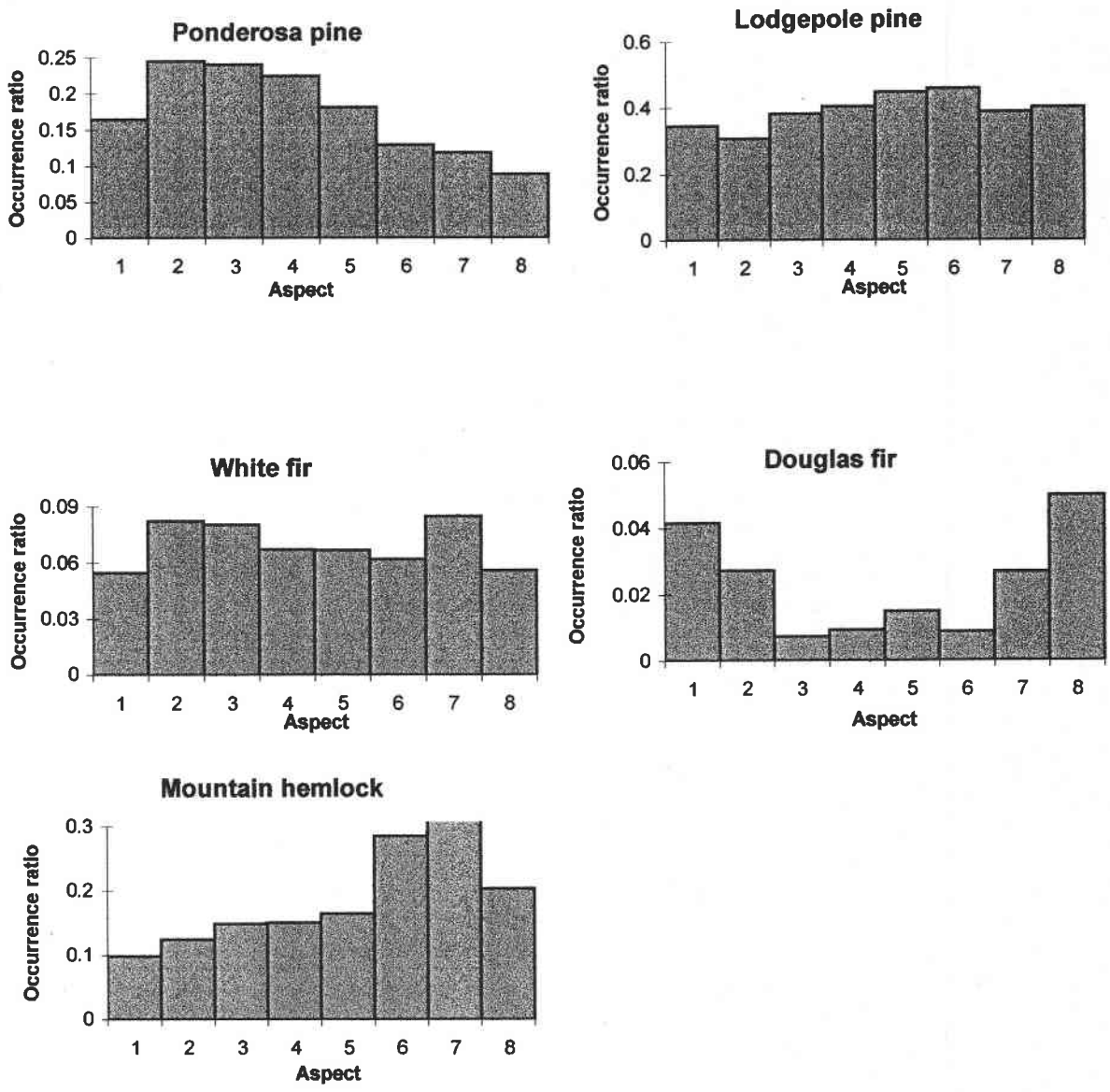



Fig. 4.3, Changes of species occurrence vs. aspect. Aspect was coded as NE=1, EN=2, ES=3, SE=4, SW=5, WS=6, WN=7, NW=8.

strongly associated with aspect while lodgepole pine and white fir are not. Comparison of the frequency of occurrence among different species indicated that lodgepole pine is most widely distributed in study area while Douglas-fir is most limited.

#### 4.4.1.2. Species occurrence vs. slope.

The occurrence of ponderosa pine and lodgepole pine declined as slope increased (Fig. 4.4). However, the occurrence of Douglas-fir, white fir and mountain hemlock all

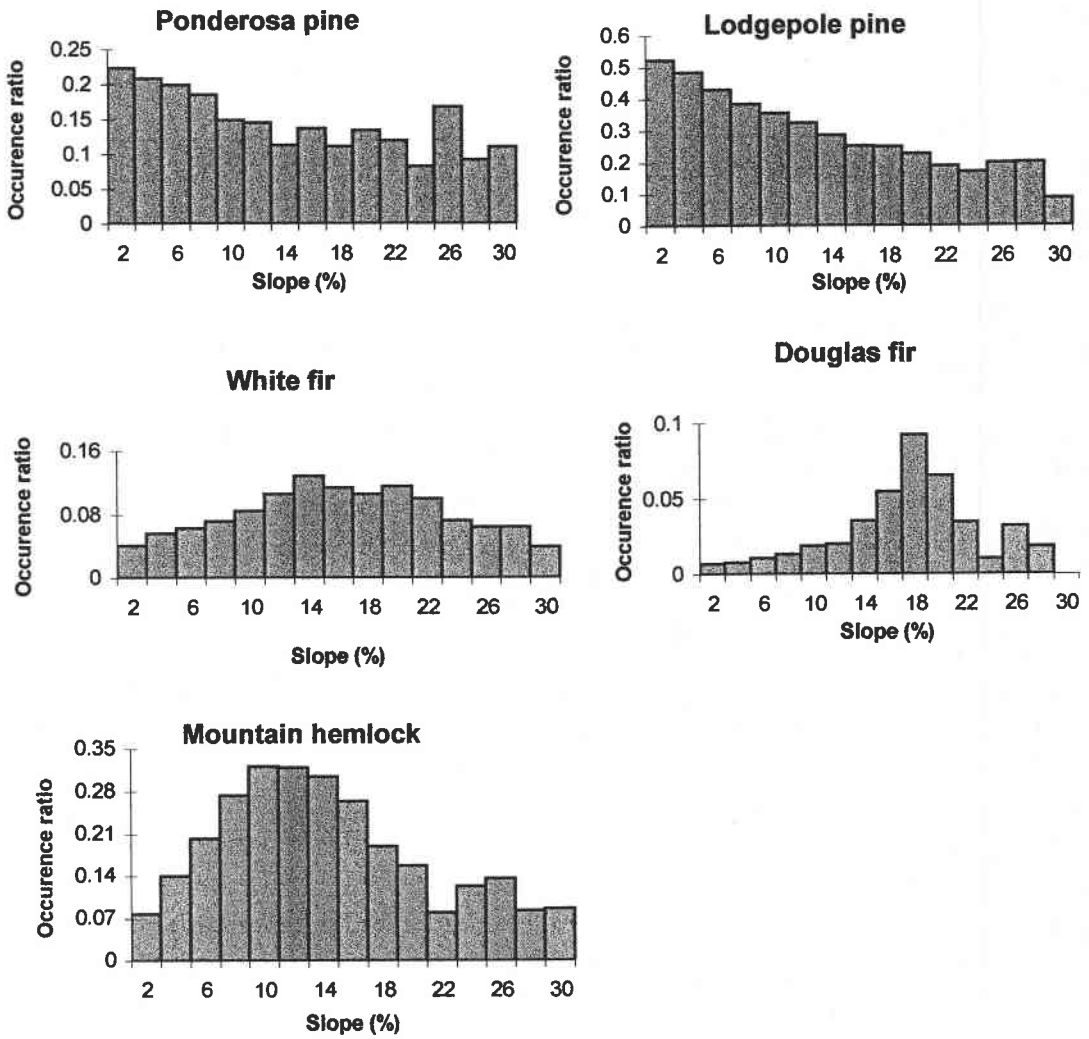



Fig. 4.4, Changes of species occurrence vs. slope.

showed a bell-shaped distribution, i.e. the highest proportion of these three species occurred at certain slope ranges. The peak for Douglas-fir, white fir and mountain

hemlock occurred respectively at slope ranges of 16-18%, 12-14%, and 8-12%. Douglas-fir showed stronger sensitivity of a certain range of slope than white fir and mountain hemlock (Fig 4.4).

#### 4.4.1.3. Species occurrence vs. elevation.

The distribution of ponderosa pine and Douglas-fir along elevation gradient again showed opposite patterns (Fig. 4.5). Most ponderosa pine occurred at lower elevation between 1400 to 1600 meters. The occurrence ratio dropped significantly as elevation increased. Most Douglas-fir occurred in a very limited range of elevation between 2100 to 2300 meters. The occurrence ratio for Douglas-fir dropped significantly when elevation either declined or increased from the peak. White fir and mountain hemlock showed a bell-shaped distribution pattern, while white fir was skewed toward lower elevations and mountain hemlock to higher elevations. Lodgepole pine distributed more or less evenly along lower elevation range and declined when elevation was higher than 2300 meters. Lower elevation of the study area extended southeastward into the Lapine Basin, an area of low rainfall (Larsen 1976), which would be expected to favor drought tolerant ponderosa pine. Douglas-fir, which is less drought tolerant and had a patchy distribution within the study area, occurred primarily above 1700 meters.

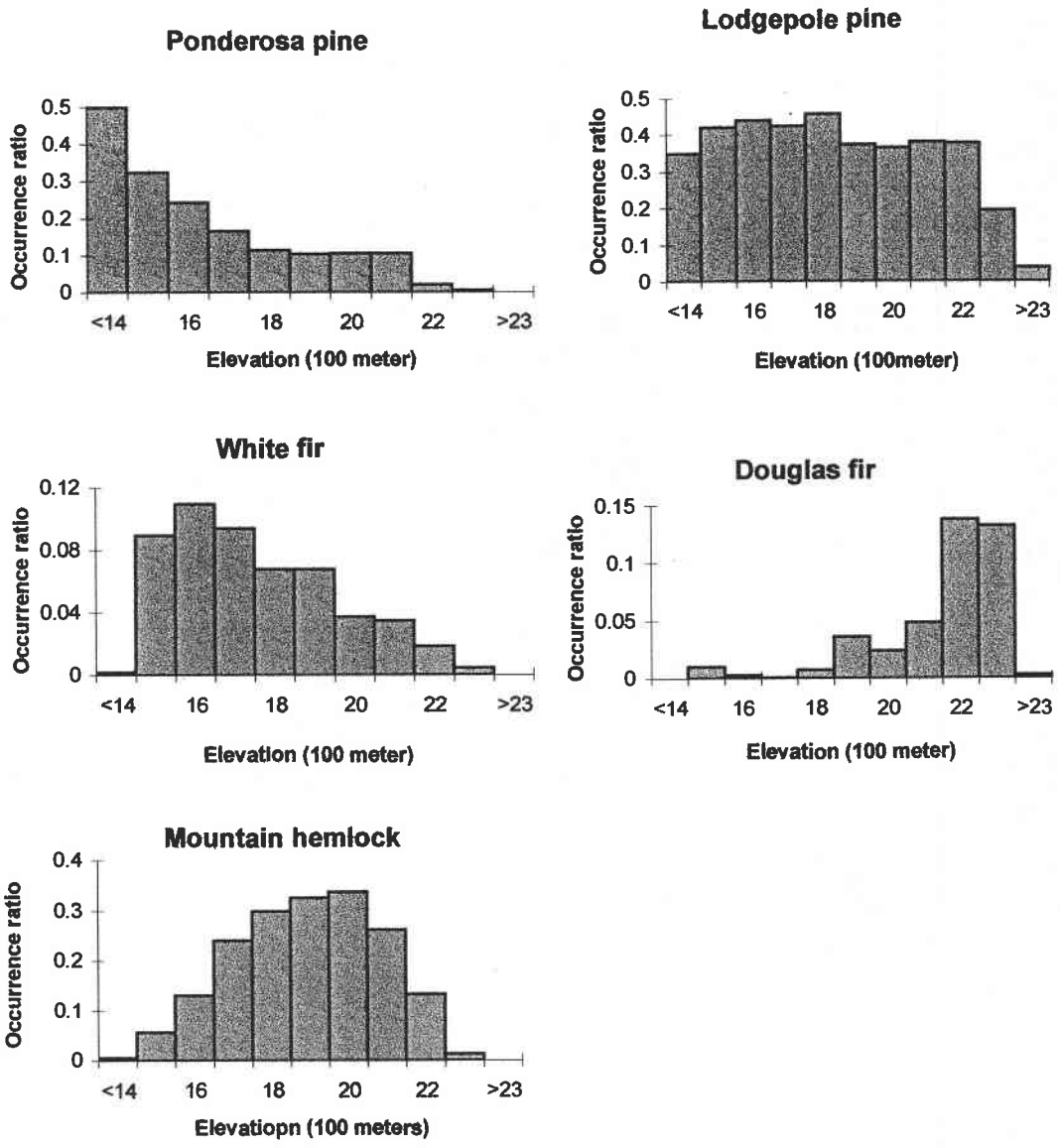



Fig. 4.5, Changes of species occurrence vs. elevation.

### 4.4.1.4. Species occurrence vs. soil type.

Tree species differed in preferred soil types (Fig. 4.6). Ponderosa pine was most abundant on loamy to sandy pumice soils over volcanic tuffs and brecias (type 4) and loamy to sand pumice soils over hard basalts, andesites, and rhyolites (type 5). Lodgepole

pine was most abundant on sandy to loamy soils over glacial till (type 2). White fir was most abundant on sandy loam to sandy pumice soils on cinder cones (type 7). Mountain hemlock was most abundant on loamy to sandy soils on volcanic bedrock materials (type 6). Douglas-fir was most abundant on variable soils (type 1). Fig. 4.6 also indicates that

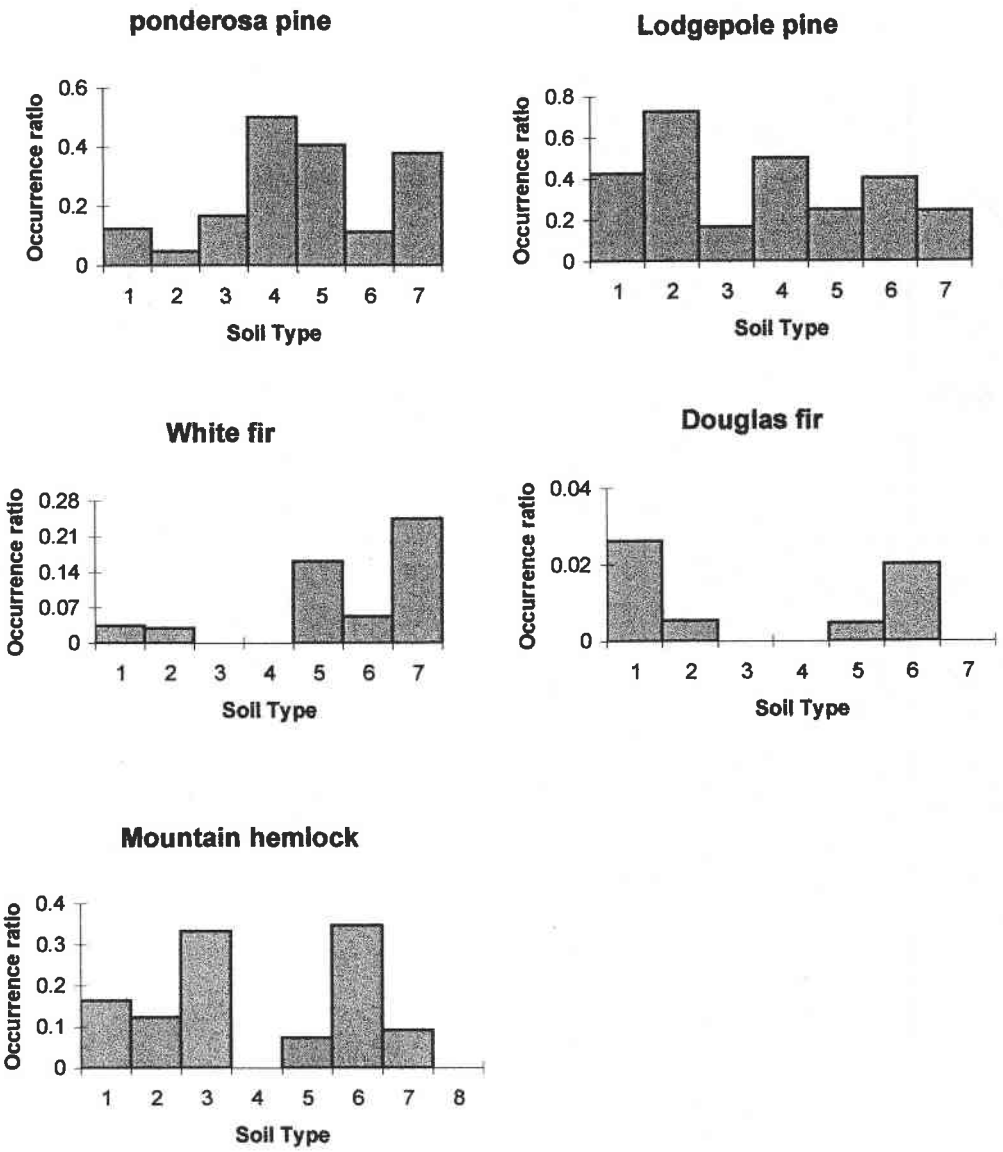



Fig. 4.6, Changes of species occurrence vs. soil types

ponderosa pine and lodgepole pine were less restricted by soil types and occurred on all soil types while the other three species only occurred on limited soil types.

#### 4.4.1.5. Species occurrence vs. lava type.

As with soil types, patterns of species distribution over various lava types differed (Fig. 4.7). A large proportion of ponderosa pine occurred on lava flows from the Siah chain of vents (type 7) and from basalt and basaltic andesite lava flows from volcanic centers in surrounding area (type 9). Lodgepole pine occurred in a variety of different lava types with the highest proportion on lava flows from shield volcano of Mount Bachelor (type 5). This indicates that lodgepole pine is a species well adapted to various landform types of different geologic origins. A high proportion of mountain hemlock occurred on the lava flow from shield volcano of Sheridan Mountain (type 8) and Kwolh Butte (type 6). Most Douglas-fir occurred on lava flows from Egan Cone (type 3) and glacial deposits (type 1), which extend northwest of Mount Bachelor toward the summit of Cascade Range. White fir mainly occurred on lava flows from shield volcano of Sheridan Mountain (type 8) and Siah Chain of vents (type 7).

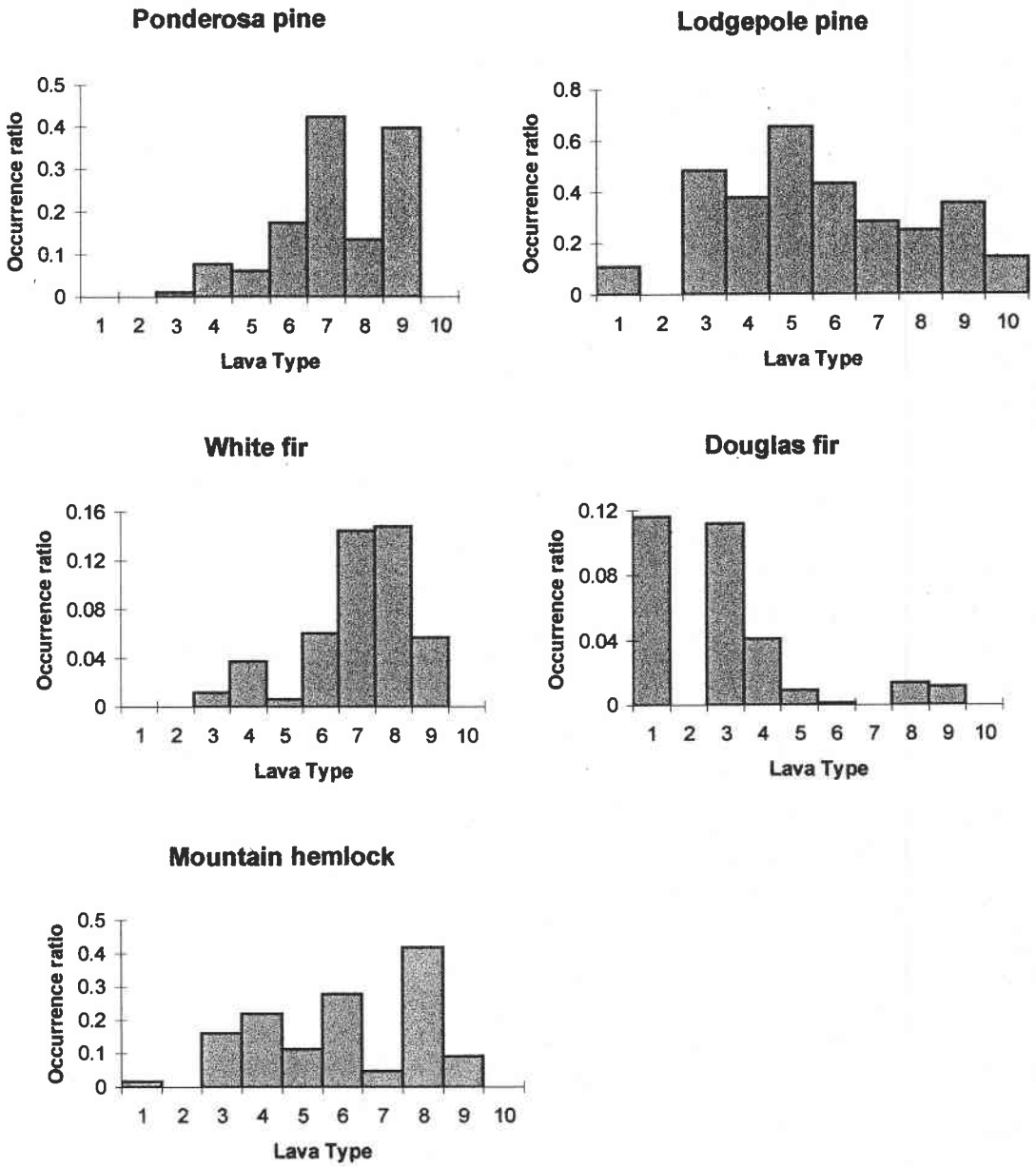



Fig. 4.7, Changes of species occurrence vs. lava types

### 4.4.1.6. Logistic regression analysis of the effect of environmental variables on distribution of ponderosa pine

Species distribution is determined by interactions among the physical matrix of landform and soils, disturbances and the relationships among species that compose the

community (Perry 1994). Bivariate relationships between tree species and physiographic factors can not, by themselves, explain patterns of species occurrence. We built a logistic regression model to analyze the comprehensive influences of topographic, soil and geologic variables on the distribution of ponderosa pine, the most abundant and widespread species in the study area. We used forward stepwise selection to find the best set of explanatory variables. The basic formula was:

$$logit(\mu) = \alpha + \beta_{slope}x_{slope} + \beta_{aspect}x_{aspect} + \beta_{1elev}x_{elev} + \beta_{2elev}x^2_{elev} + \beta_{soil}x_{soil} + \beta_{lava}x_{lava} \qquad (4.3)$$

where $\mu$ is the probability of the occurrence of a tree species. A quadratic term for elevation variable was included in the model because the occurrence of ponderosa pine did not respond to elevation change linearly. Tab. 4.4 lists the summary statistics. The partial t statistics indicated that NE, ES and SE aspects were significant indicator variables as well as slope and elevation after adjusting for aspect. Of the various soils, only those classed as sandy to loamy soils over glacial till (type 2) and loamy to sandy pumice soils over hard basalts, andesites, and rhyolites (type 5) was significantly correlated with the occurrence of ponderosa pine. Type 2 had a negative influence and type 5 had a positive influence. Three lava types (6, 7, 9) were significant indicators of the occurrence of ponderosa pine. ANOVA was conducted to analyze the difference among models with sequential additions of variables. A Chi-square was specified in ANOVA to test the difference between models during adding process. The analysis indicated that each of the five variables, aspect, slope, elevation, soil and lava, contributed significantly to the model in sequential addition. To further examine the contribution of each variables to the occurrence of ponderosa pine, we analyzed bivariate relationships between the probability of tree occurrence and each of the predictors by fitting a null model and then adding each of the

Tab. 4.4. Logistic regression analysis of ponderosa pine distribution in response to topographic, soil and geologic variables.

| *Variables* | *Coefficients* | *t Statistics* |
|---|---|---|
| Intercept | -5.986250440 | -3.084223120 |
| Aspect Variables: | | |
| aspect2 | 0.254407142 | 1.775940413 |
| aspect3 | 0.343522222 | 2.532191758 |
| aspect4 | 0.376643843 | 2.694063700 |
| aspect5 | 0.260678125 | 1.764861237 |
| aspect6 | 0.018983933 | 0.133972602 |
| aspect7 | -0.053762001 | -0.362760611 |
| aspect8 | -0.220497571 | -1.196132909 |
| Slope | 0.040569973 | 5.879158639 |
| Elevation Variables: | | |
| Linear | -108.548391487 | -13.613196138 |
| Quadratic | -20.857772609 | -3.503094354 |
| Soil Variables: | | |
| soil2 | -1.568578438 | -10.280381347 |
| soil3 | -0.598293949 | -0.524873855 |
| soil4 | 0.280966367 | 0.198225787 |
| soil5 | 0.245935640 | 3.225048671 |
| soil6 | -0.062376646 | -0.806069651 |
| soil7 | -0.001211538 | -0.007593702 |
| Lava Variables: | | |
| lava3 | 1.334318553 | 0.675739916 |
| lava4 | 3.770225747 | 1.945432993 |
| lava5 | 2.775112873 | 1.429152066 |
| lava6 | 3.859580637 | 1.990090565 |
| lava7 | 4.371377117 | 2.253630033 |
| lava8 | 3.634340899 | 1.872920529 |
| lava9 | 4.429310762 | 2.265723930 |
| lava10 | -0.154775885 | -0.021006407 |

*Note: The codes for aspect, soil and lava type refer to Tab. 4.1, 4.2 and 4.3

terms, one at a time. The $C_p$ statistics was used to compare models that were not nested (Tab.4.5). A small $C_p$ corresponds a better model in the sense of smaller residual deviance. Clearly lava type was the best single predictor of ponderosa pine occurrence, while soil type was the second best. Of the topographic variables, elevation had the most power. Slope had the least influence over ponderosa pine occurrence. The results can be inferred such that the lava type is the most significant variable influencing the occurrence of ponderosa pine over the study area.

Tab. 4.5. Comparison of fits for each predicting variables to the occurrence of ponderosa pine.

| Variable | $C_p$ statistics |
|---|---|
| Null | 13954.00 |
| Slope | 13852.80 |
| Aspect | 13683.12 |
| Elevation | 13048.99 |
| Soil | 12457.62 |
| Lava | 12209.70 |

In summary, the distribution of ponderosa pine within the study area depended to one degree or another on all physiographic variables, especially lava and soil type. The species tended to occur on products of shield volcano of Kwolh butte ( type 6), products of Siah chain of vents (type 7), and basalt and basaltic andesite lava flow from volcanic centers in surrounding areas (type 9), and not occur on alluvial, lacustrine, and colluvial deposits (type 10). Higher occurrence was also found on soils classified as loamy to sandy pumice soils over hard basalts, andesites, and rhyolites (type 5), which were essentially all those with pumiceous volcanic ash and pumice lapilli mixed with highly fractured lava over sandy to loamy buried soils. The occurrence of ponderosa pine was found less frequently on sandy to loamy soils over glacial till. These results are consistent with observations that the distribution of ponderosa pine is coincident with these fractured lava flows, and is less frequent on glacial and alluvial deposits (Franklin and Dyrness, 1973). Ponderosa pine declined rapidly in abundance when elevation increased, though the effect was nonlinear. The species was more likely to occur on easterly, southeasterly and southerly aspects and less likely to occur on westerly, northerly and northwesterly aspects. The contribution of

slope to predict ponderosa pine occurrence was also significant, although it is least important of all five physiographic variables.

### 4.4.2. Canopy closure and site variables

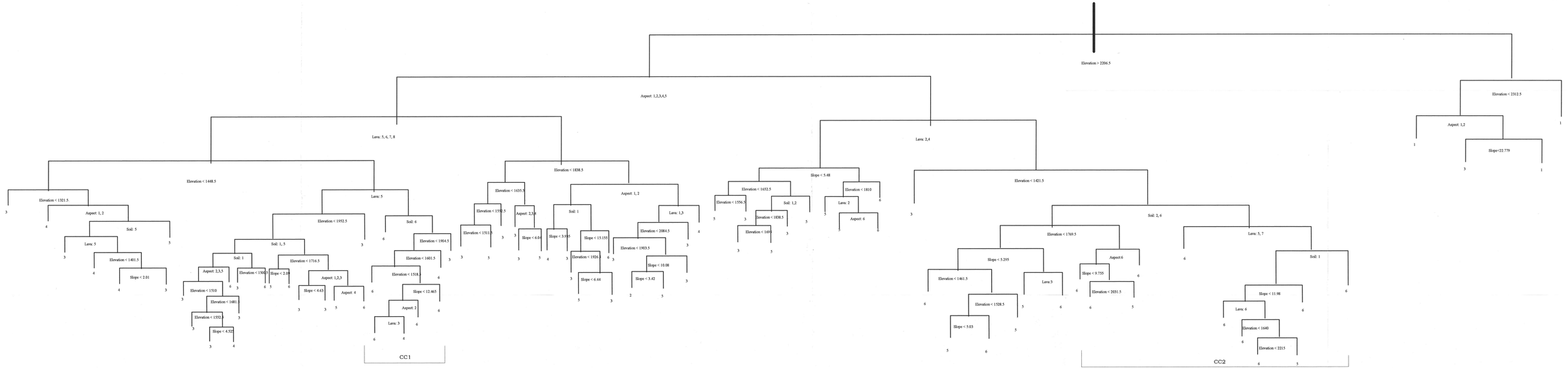
We built a tree-based model to predict the probability of developing different crown closure levels using aspect, slope, elevation, lava type and soil type as predictor variables. Fig. 4.8. Shows the full binary tree. The decision rules were labeled at the nodes of the tree. Crown closure levels were labeled at the bottom of the leaves. A sampling point that met these decision rules was classified into the left split, otherwise it was lumped into the right split. The misclassification rate of the tree model for Fig.4.8 was 16.84%. An interpretation of Fig. 4.8 follows:

1) Elevation is the primary variable partitioning sampling points into two levels. Points with crown closure > 40% (levels 3,4,5,and 6) were predicted to fall at elevations below 2206.5 m, while points higher than that were predicted to have crown closures < 40% (levels 1,2,3).

2) For elevations > 2206.5 m, no points were predicted to have crown closure > 10%, except points falling on slopes less than 21% and aspects other than NE and EN.

3) For elevations < 2206.5 m, the second level partitioning was due to aspect. Sampling points on NE, EN, ES, SE, and SW aspects (labeled as 1,2,3,4, and 5) had a relatively high probability of crown closures between 25% and 70% (labeled 3,4,and 5), while points on WS, WN, and NW aspects (labeled as 6, 7 and 8) had a relatively high

Fig. 4.8, Display of the Tree-based model with canopy crown closure as response variable, slope, aspect, elevation, soil and lava type as predictors.



probability of crown closures >70% (labeled as 6). However, crown closures on all aspects varied widely depending primarily on lava type and elevation.

4) Below 2206.5 m, lava type accounted for the third level of discrimination on both aspect branches. On WS, WN, and NW aspects, points falling on lava types other than 4 and 6 tended to have higher crown closures than those on 4 and 6. The same was true on the other aspects, points falling on lava types 7, 8, 9, and 10 having a higher probability of crown closure > 40% than points falling on other lava types.

5) Highest crown closure occurred in two major partitioning groups labeled as CC1 and CC2 in Fig. 4.8. The variable values corresponding to these two partitions were:

CC1: Elevation < 1904.5 meters and > 1448.5 meters, lava type = 8, 9, and 10, aspect =NE, EN, ES, SE, SW (Labeled as 1, 2, 3, 4, and 5).

CC2: Elevation < 2206.5 meters and > 1421.5 meters, lava types other than 4 and 6, aspect = WS, WN and NW.

6) Slope did not significantly influence crown closure levels in the study area.

It should be noted that the accuracy of prediction in a tree-model decreases from the root (top) to the leaves (bottom). Therefore, results at the leaf levels should be interpreted cautiously

### 4.4.3 Forest LAI and site variables

Fig. 4.9 shows the bivariate relationships between LAI and topographic, soil, and geologic variables. The change in LAI with elevation was nonlinear and bell-shaped. Both the peak and variance of LAI were maximum between 1600 and 1800 m, and declined

significantly has elevation increased or decreased. At its peak, LAI varied between approximately 0.5 and 12 $m^2/m^2$, which could reflect heterogeneous site conditions, disturbance patterns, or both. Lava type also significantly influenced LAI. Highest values occurred on lava flow from the shield volcanoes of Sheridan Mountain (lava group 8) and Siah Chain of vents (lava group 7). The distribution of LAI within most lava types was skewed toward higher LAI values. LAI also showed strong dependence on soil, reaching its highest levels on soil types 5, and 7, which are characterized by volcanic ash and pumice lapilli mixed with highly fractured lava over sandy to loamy buried soils ( type 5)

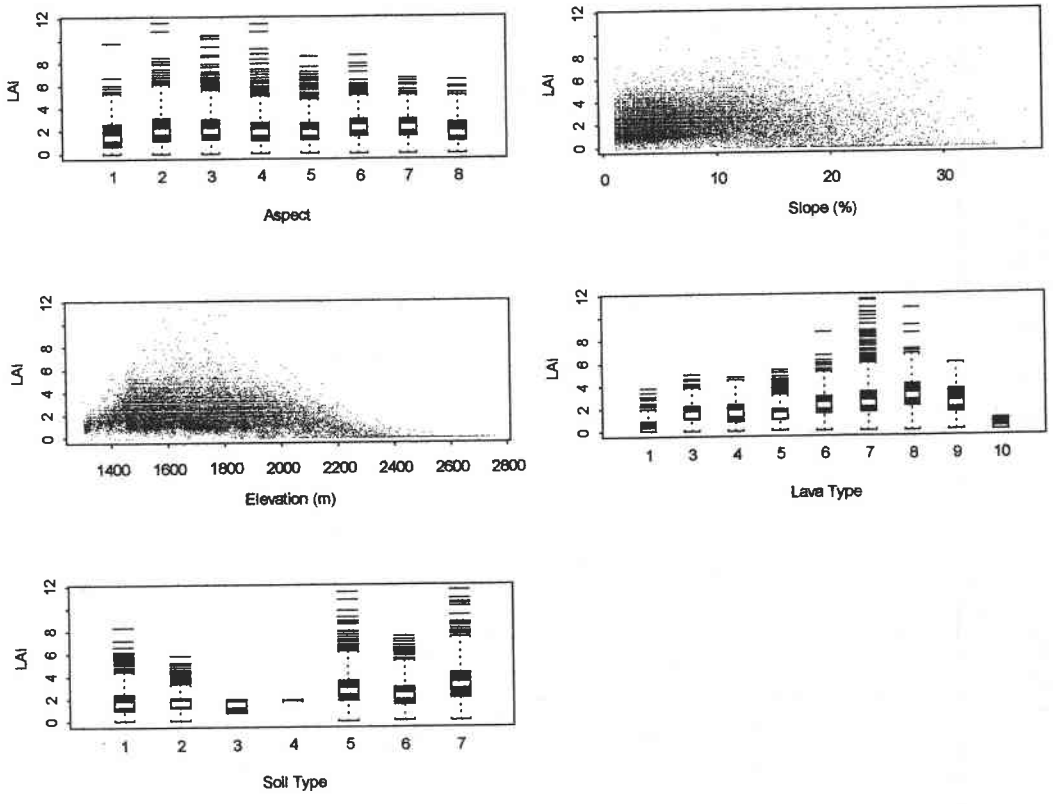



Fig. 4.9, Bivariate relationship between LAI and environmental variables

and cinder buried soil. As on lava, variation of LAI within a given soil types was skewed toward higher values. LAI did not differ significantly among aspects nor on slope below 30%.

We built a multiple regression model to analyze the response of LAI to the combination of topographic, soil, and geologic variables. A step-wise procedure was used to select the model with best least-square fit. The model was:

$$log(LAI) = \alpha + \beta_{slope}x_{slope} + \beta_{aspect}x_{aspect} + \beta_{elev}x_{elev} + \beta_{elev}x^2_{elev} + \beta_{soil}x_{soil} + \beta_{lava}x_{lava} \quad (4.4)$$

Aspect, elevation, soil, lava, and slope contributed significantly to the model (Tab. 4.6). However, although highly significant ($p < 0.00$), multiple $R^2$ was only o.3881. The model

Tab. 4.6 Summarized statistics of *eq.* 4.4.

| *Variables* | *Coefficients* | *p-value* |
|---|---|---|
| Intercept | -0.1737 | 0.0207 |
| Aspect Variables: | | |
| aspect2 | 0.1668 | 0.0000 |
| aspect3 | 0.1519 | 0.0000 |
| aspect4 | 0.1413 | 0.0000 |
| aspect5 | 0.2339 | 0.0000 |
| aspect6 | 0.3290 | 0.0000 |
| aspect7 | 0.3118 | 0.0000 |
| aspect8 | 0.2119 | 0.0000 |
| Slope | 0.0177 | 0.0000 |
| Elevation Variables: | | |
| Linear | -16.8579 | 0.0000 |
| Quadratic | -33.3784 | 0.0000 |
| Soil Variables: | | |
| soil2 | -0.1061 | 0.0000 |
| soil3 | -0.0476 | 0.8223 |
| soil5 | 0.3078 | 0.0000 |
| soil6 | 0.0399 | 0.0019 |
| soil7 | 0.0826 | 0.0213 |
| Lava variables: | | |
| lava3 | 0.0098 | 0.8959 |
| lava4 | 0.3341 | 0.0000 |
| lava5 | 0.0082 | 0.9111 |
| lava6 | 0.4548 | 0.0000 |
| lava7 | 0.5399 | 0.0000 |
| lava8 | 0.6640 | 0.0000 |
| lava9 | 0.5276 | 0.0000 |
| lava10 | 0.1963 | 0.5997 |
| F Statistics p-value | 0.00 | |
| Multiple $R^2$ | 0.3881 | |

did not explain the variance of the data efficiently. Disturbance history, differences in precipitation not accounted for by topography, and variability within soil and lava types might all account for unexplained variance. The soil mapping in particular was not highly accurate (Karen Bennett* 1996, personal communication).

We used $C_p$ statistics to compare the relative importance of different variables in the model. A "null" model was fit and then used as a base to which the site variables was added singly, giving six separate models for comparison (Table 4.7). A smaller $C_p$ corresponds to a better model in the sense of smaller residual deviance. Of the variables we measured, elevation and lava type were most closely associated with variation in LAI, followed by soil type. Neither slope nor aspect contributed much explaining power.

Tab. 4.7. Comparison of fits for each predicting variables to forest LAI.

| Variable | $C_p$ statistics |
|---|---|
| Null | 5861.37 |
| Slope | 5851.53 |
| Aspect | 5747.88 |
| Soil | 5208.44 |
| Lava | 4893.55 |
| Elevation | 4799.79 |

## 4.5. Discussion

How does a forest respond to its physiographic environment? A forest is usually described and measured by varieties of indices and variables, such as the variables of

* Karen Bennett was the soil scientist of Deschutes National Forest, 1645 Highway 20 East, Band, OR 97701.

species occurrence, crown closure and LAI chosen in this paper. Each of these variables describes one aspect of the forest characteristics. Together they contain an important complement of the structural information of the forest. However, These variables are biologically correlated. Fig. 4.10 shows the relationship between LAI and crown closure. Clearly, crown closure was positively correlated with LAI. As crown closure increased, so does the average of LAI. The statistical analyses showed a similar order of environmental variables that contribute to the predictions of crown closure and LAI. No surprisingly, a large number of high LAI pixels also contained the highest levels of crown closure. The distribution of LAI at high crown closure however was extremely skewed and the variance was increased. Forest LAI differed significantly at high crown closure. Crown closure and LAI were developed with totally different techniques from different sources of data in this study. Comparing both sets of measurements is helpful in understanding the relationship between different forest structural attributes, and calibrating both measuring results.

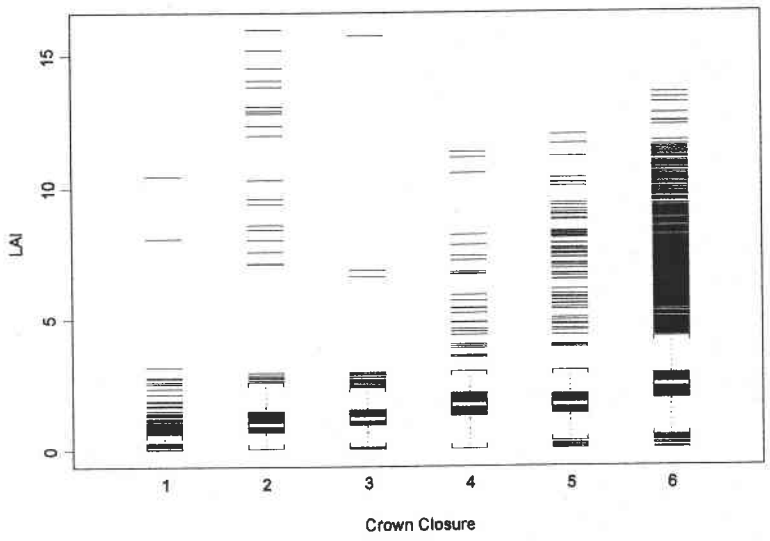



Fig. 4.10, Barplots showing the relationship between crown closure and LAI

We have introduce a strategy for the diagnosis of the physiographic factors that are associated with forest occurrence and structural development. We analyzed each of these structural variables in response to topographical, soil and geologic variables. We found that each structural variables respond differently. Species occurrence was more influenced by lava and soil types, while crown closure and LAI were more influenced by elevation variable. Our analysis of ponderosa pine occurrence and environmental variables showed that higher occurrence was found on soils classified as loamy to sandy pumice soils over hard basalts, andesites, and rhyolites (type 5, see Tab. 4.2), which were essentially all those with pumiceous volcanic ash and pumice lapilli mixed with highly fractured lava over sandy to loamy buried soils. It has been observed that forest regeneration on lava landscape in Deschutes National Forest is related to lava morphology (Larry Chitwood* 1995, personal communications). Two general morphologies develop in these basalt flows that are strictly different: rubble lava and inflated lava (Chitwood, 1992). These stiff flows shear and abrade internally as they move creating a thick layer of hummocky rubble that accumulates on their top. Below the rubble is the dense, fractured lava. It is in these fractured lava flows where soil accumulates through the cracks, trenches and underground opening space of lava. Soil is carried into the interior of the flow by gravity and water. This soil, although undeveloped, is immediately available to maintain water and nutrients, and support vegetation. Tree seeds which fall onto the fractured lava flows have a higher probability to establish than those falling elsewhere. We grouped the fractured

---

* Larry Chitwood was the geologist of Deschutes National Forest, 1645 Highway 20 East, Band, OR 97701.

lava morphology as soil type 5 in soil stratification, although we were unable to relate the dense lava layer structure in a soil stratification. The ponderosa pine occurrence was significantly higher on soil type 5 than on the others. Tree occurrence pattern on a landscape in part represents the establishment history of the species. The statistical analysis of ponderosa pine occurrence seems to support this hypothesis that ponderosa pine has higher probability of establishment on fractured lava structure.

Different species differs in response to different environmental variables. Lodgepole pine was distributed across a wide range of different topography, soil and lava types. Douglas-fir was found to be most selective on environment among the five species. Compared to Douglas-fir, ponderosa pine was less selective, but was distributed in more or less an opposite range of environmental variables. Mountain hemlock and whit fir seemed to adapt to a wider range of environment settings than Douglas-fir, but less than that of logepole pine. These distribution patterns showed that ponderosa pine and Douglas-fir are dominant species on some stands but not on the others, while lodgepole pine coexists in most stands. The negative influence of sandy to loamy soils over glacial till (soil type 2) on ponderosa pine occurrence statistically supports the observation that the distribution of ponderosa pine coincides with the boundary of lava flows as these glacial and alluvial deposits surround the volcanic chain (Scott and Gardner 1992).

Elevation became a significant contributor to predict crown closure and LAI. The highest LAI occurred only between 1600 - 1800m elevation in the study area, whereas the highest values of crown closure occurred between 1448 - 2206m. We believe that water availability restricts crown closure and LAI development at low elevation. At low elevation the study area extends southeasterly toward dry Lapine Basin where annual

precipitation is only 60 cm. When elevation increases, the precipitation increases rapidly in this area (200 cm at the peak of Mount Bachelor, some local areas could receive higher precipitation), resulting in an increase of LAI and crown closure. At high elevation, however, temperature decreases and becomes more and more a limited factor to leaf area and crown closure development due to the shortened growing season, snow damage and slow nutrient circling rate (Gholz 1982). Crown closure and LAI decreased correspondingly (Fig. 4.8 and Fig. 4.9). There is an elevation range where the positive influence of precipitation and negative influence of temperature on crown closure and LAI reach their equilibrium and that allows forest stands to reach the highest crown closure and LAI. Specifically, the elevation range was between 1421 to 2206 meters for highest crown closure, and 1600 to 1800 meters for LAI in this study. Soil was not a significant variable influencing crown closure and LAI of the forests, although this may reflect inadequacy of the data.

### 4.6. Conclusions and recommendations

The studies of this paper successfully demonstrated the method for integrating multiple GIS data layers in spatial analysis to develop relationships between forests and environment. The results clearly indicated that different variables of forest characteristics respond to environmental variables differently. Species occurrence over the lava landscape was most significantly influenced by lava types, whereas crown closure and LAI were most significantly influenced by elevation. Among the topographical variables, slope was least significantly correlated to species occurrence and forest structural attributes.

Different species also show a different preference of establishment in the values of the environmental variables. Lodgepole pine was adapted to a wide range of topography, lava and soil types, whereas Douglas-fir was more selective.

More detailed and accurate data layers are needed to improve the analysis, hence provide more confident conclusions. The inclusion of disturbance and precipitation data would be particularly valuable for further analyses. Improvement in soil data would also be recommended. Based on the analysis of this paper, an improved sampling design could be implemented to analyze and parameterize the response of each species to environmental variables. This paper provided a frame work and starting point.

# Chapter 5. Summary and Conclusions

Paralleling a common concern in landscape ecology, spatial patterns, this study focused on measuring forest spatial patterns and analyzing the relationships between the forest spatial patterns and environmental settings in a lava landscape. There are several stages in the study of patterns and ecological processes. First, one must have measures to describe patterns, so that the criteria can be established for relating patterns to environment. Second, statistical methods can be used to quantify forest spatial patterns in response to environmental changes. With the measures of spatial patterns and the understanding of correlation, a third step can be taken to develop models to suggest possible mechanisms and experiments in exploring the interactions of patterns and ecological processes (Levin, 1992). This study was designed with the awareness of these logistics, however, the study objectives were limited to the first two steps. Using aerial photographs and Landsat TM imageries, we measured forest spatial patterns in different scales across the lava landscape. A GIS was applied to integrate multiple spatial data layers. Statistical analyses were used to bring insights of relationships between forests and environment. The methods and results presented in the study disclosed several fundamental points:

1). Forest spatial patterns can be measured at different scales from individual tree to stand level with reasonable accuracy utilizing remote sensing techniques. The digital and statistical analyses presented in this study serve as instructive methods for characterizing and measuring forest attributes in similar sites.

2). A GIS is necessary to integrate multiple forest and environmental data layers into a common data space for statistical analysis to explore the relationships among them. Different sampling designs can be implemented to meet the needs of statistical analyses. Such an analysis is also a step-wise process. As new and improved data layers are available, they can be added to the GIS to improve the analysis.

3). Species occurrence, canopy crown closure and density, and LAI, as different attributes of the forest that relate to seedling establishments, forest structure, and productivity, describe different aspects of the forest at different scales. These attributes respond to soil, geological history, and topographic characteristics differently. No single attribute can fully describe the response of forest to its environment. A study of analyzing multiple forest attributes in response to environmental variables is necessary to achieve the full understanding of the relationships between forests and environment.

4). The next stage of this study is to use the data layers and parameters developed in this study to build process models to simulate possible mechanisms in exploring the interactions of forest spatial patterns and ecological processes.

# References

Ahern, F. J., Brown, R. J., Cihlar, J., Gauthier, R., Murphy, J., Neville, R. A. and Teillet, P. M., 1987. Radiometric correction of visible and infrared remote sensing data in the Canadian Center for Remote Sensing. International Journal of Remote Sensing, 8, 1349-1376.

Baret, F. and G. Guyot, 1991. Potential and limits of vegetation indeices for LAI and APAR assessment. Remote Sensing of Environment. Vol. 35, pp. 161-173.

Bauer, M. E., L. L. Biehl, and B. F. Robinson, 1980. Final report volume 1: Field research on the spectral properties of crops and soils. NASA Rep. SR-PO-04022, Laboratory of Applications of Remote Sensing, Purdue University, West Lafayette, IN.

Bauer, M. E. et al., 1986. Field spectroscopy of agricultural crops. IEEE Transaction on Geoscience and Remote Sensing, vol. GE-24, No. 1, pp 65-75.

Berry, W. D. and S. Feldman, 1985. Multiple regression in practice. Sage University Paper series on quantitative applications in the social science, 07-050. Beverly Hill and London: Sage Pubns.

Boots, B.N. and A. Getis. 1988. Point pattern analysis. In: Thrall, G. I. ed., Scientific geography series, pp 8-93. Sage publication.

Bukata, R. P., J. E. Bruton, and J. H. Jerome, 1983. Use of chromaticity in remote measurements of water quality. Remote Sensing of Environment, 13:161-177.

Chambers, J. M. and Hastie T. J., 1992. Statistical models in S. AT&T Bell Laboratory. Pp 377-419.

Chen, J. M. and M. A. Guilbeault, 1996. Evaluation of vegetation indeices for retrieving biophysical parameters of boreal conifer forests. In proceedings of 26th International Symposium on Remote Sensing of Environment and 18th Annual Symposium of the Canadian Remote Sensing Society, March 25-29, 1996, Vacouver, BC, Canada.

Chitwood, Larry, 1992. Geologic aspects of soil development on young lava flows of Deschutes National Forest. In Popovich, S., Forest lava study: transect data and result. Deschutes National Forest, USDA Forest Servcie, technical report.

Crist, E. P., R. Laurin and R. C. Cicone, 1986. Vegetation and soils information contained in the transformed Thematic Mapper data. In proceedings of IGARSS' 86 symposium, ESA publication division, ESA sp-254.

Crist, E. P. and R. C. Cicone, 1984. A physically-based transformation of Thematic Mapper data - The TM Tasseled Cap. IEEE Transactions on Geoscience and Remote Sensing, vol. GE-22, No. 3.

Deering D. W., J. W. Rouse, R. H. Haas and J. A. Schell, 1975. Measuring forest production of grazing units from Landsat MSS data. In proceedings, 10th international symposium of remote sensing of environment. Vol. 2, pp1169-1178.

Deschutes national Forest, 1994. Landsat vegetation mapping project. US Forest Servuce, Deschutes National Forest, technical report.

Egan, G. W.. 1985. Photogrammetry and polarization in remote sensing. Elsevies Science Publisher, New York.

Ekstrand, S, 1996. Landsat TM-based forest damage assessment: Correction for topographic effects. Photogrammetric Engineering and & Remote Sensing, vol. 62, No. 2, pp. 151-161

Firestone, L., S. Rupert, J. Olson and W. Mueller. 1996. Automated feature extraction: the key to future productivity. PE & RS, 62(6): 671-675.

Forman, R.T.T and M. Godron. 1986. Landscape Ecology. Willey, New York, New York, USA.

Franklin, J. F and C. T. Dyrness, 1973. Natural vegetation of Oregon and Washington. USDA Forest Service, PNW Forest and range Experimental Station, General techical report PNW-8.

Gholz, H. L., 1982. Environmental limits on aboveground net primary production, leaf area and biomass in vegetation zones of the Pacific Northwest. Ecology 63, 469-481.

Goodchild, M. F., 1987. A spatial analytical perspective on geographic information systems. Int. J. Geographic information system. Vol. 1 No. 4: 327-334

Gordon, H. R., 1987. Calibration requirement and methodology for remote sensors viewing the ocean in the visible. Remote Sensing of Environment, 22:103-126.

Gross, H. N. and J. R. Schott, 1996. Application of spectral mixing to image fusion. In proceedings of 26th International Symposium on Remote Sensing of Environment and 18th Annual Symposium of the Canadian Remote Sensing Society, March 25-29, 1996, Vacouver, BC, Canada.

Hall, F. G., Y. E. Shimabukuro, and K. F. Huemmerich, 1995. Remote sensing of forest biophysical structure using mixture decomposition and geometric reflectance models. Ecological Applications, 5(4), pp. 993-1013.

Hoppus, M. L. and D. T. Evans. 1994. The role of texture in interpreting images of forest land. in Proceedings 5th Forest Services Remote Sensing Application Conference. Am. Soc. Photogramm. Remote Sensing, pp. 316-324.

Huang, J., Chen C. and Perry D. A., 1997a. Characterizing forest spatial patterns using digitized aerial photographs. Remote sensing of environment (Submitted)

Huang, J., Cohen, W. and Perry, D. A., 1997b. Optimizing multispectral data for LAI estimates. Remote sensing of environment (In review).

Jackson, R. D., 1983. Spectral indices in n-space. Remote sensing of Environment, 13:409-421.

Johnson, 1990. Analyzing spatial and temporal phenomena using geographic information systems, a review of ecological applications. Landscape Ecology, Vol. 4, No. 1, Pp 31-43

Kowalik, W. s., R.J.P. Lyon, and P. Switzer, 1983. The effects if additiveradiance terms on rations of Landsat data. PE&RS, 49(5):659-669.

Kolbl, O. and U. Bach. 1996. Tone reproduction of photographic scanner. PE & RS, 62(6): 687-694.

Larsen, D. M., 1976. Soil resource inventory, Deschutes National Forest. Pacific Northwest Region, US Forest Service. Technical report.

Leckie, D. G., 1990. Advances in remote sensing technologies for forest surveys and management. Can. J. For. Res. 20:464-483.

Legendre, P. and Fortin, M. J., 1989. Spatial patterns and ecological analysis. Vegetatio 80: 107-138.

Levin, S. A., 1992. The problem of pattern and scale in ecology. Ecology, 73(6), pp. 1943-1967.

Lillesand, T. M. and R. W. Kiefer. 1994. Remote sensing and image interpretation. 3rd Ed. John Wiley & Sons. New York.

Markham, B. L. and J. L. Barker, 1986. Landsat MSS and TM post-calibration dynamic ranges, exoatmospheric reflectances and at-satellite tempretures. EOSAT Landsat Technical Notes, No.1, pp. 3-8.

Marshall, J. D. and R. H. Waring, 1986. Comparison of methods of estimating leaf-area index in old-growth Douglas fir. Ecology: 67(4), pp. 975-979.

MathSoft Inc., 1997. S-Plus 4, Guide to statistics. Data analysis production division, MathSoft, Seattle.

Mussio, L. and D. L. Light, 1995. ISPRS Commission I: Sensors, Platforms, and Imagery Symposium. PE&RS, ISPRS Report, pp 1339-1344

Nemani, R., L. Pierce and S. Running, 1993. Forest ecoystem processes at teh watershed scale: sensitivity to remotely-sensed Leaf Area Index estimates. Int. J. Remote Sensing, vol. 14, No. 13, 2519-2534.

Peddle, D. R., G. H. Forrest, and E. F. LeDrew, 1996. Remote sensing f boreal forest biophysical variables using spectral mixture analysis and a canopy reflectance model. In proceedings of 26th International Symposium on Remote Sensing of Environment and 18th Annual Symposium of the Canadian Remote Sensing Society, March 25-29, 1996, Vacouver, BC, Canada.

Perry, D. A., 1994. Forest Ecosystems. The John Hopkins University Press. Pp 65-70.

Popovich, S., 1991. Final status report of forested lava study - accomplishments and future direction. USDA Forest Service, Deschutes National Forest, technical report.

Ripple, W. J., Bradshw, G. A. and Spies, T. A., 1991. Measuring forest landscape patterns in the Cascade Range of Oregon, USA. Biological Conservation. 57:73-88.

Running, S. W., D. L. Peterson, M. A. Spanner, and K. B. Teuber, 1986. Remote sensing of coniferous forest leaf area. Ecology: 67(1), pp. 273-276.

Saleh, R. A, 1996. Photogrammetry and the quest for digitalization. PE & RS, 62(6): 675-678.

Scott, W. E. and Gardner, C. A., 1992. Geologic map of the Mount Bachelor volcanic chain and surrounding area, Cascade Range, Oregon. U. S. Geologic Survey, map I-1967.

Smith, J. A., T. L. Lin, and K. J. Ranson, 1980. The Lambertian assumption and Landsat data. PE&RS, vol. 46, No. 9, pp. 1183-1189.

Spanner, M. A., L. Johnson, J. Miller, R. McCreight, J. Freemantle, J. Runyon, and P. Gong, 1994. Remote sensing of seasonal leaf area index across Oregon Transect. Ecological Applications, vol. 4, No. 2, pp. 258-271.

Spanner, M. A., L. L. Pierce, D. L. Peterson and S. W. Running, 1990. Remote sensing of temperate coniferous forest leaf area index, the influence of canopy closure, understory vegetation and background reflectance. INT. J. Remote Sensing, vol. 11, No. 1, 95-111.

Spanner, M. A., Pierce, L., Peterson D. L. and Running, S. W., 1990. Remote sensing of temperate coniferous forest leaf area index - the influence of canopy closure, understory and background reflectance. Int. J. Remote Sensing. Vol. 11, No. 1, 95-111.

Spurr, S.H. 1960. Photogrammetry and photo-interpretation, with a section of applications to forestry. Second edition of aerial photographs in forestry. The Ronald Press Company. New York.

Steyaert, L. T., Loveland, T. R. and Parton W. J., 1997. Landscape cover characterization and land surface parameterization research. Ecological Applications, vol. 7, No. 1, pp1-2.

Strahler A. H., Woodcock C. E. and Smith J. A., 1986. On the nature of models in remote snesing. Remote sensing of environment. 20:121-139.

Teillet, P. M., B. Guindon and D. G. Goodenough, 1982. On the slope-aspect correction of multispectral scanner data. Canadian Journal of Remote Sensing, vol.8, No. 2, pp. 84-106.

Turner M.G. and R.H.Gardner. 1989. Quantitative methods in Landscape Ecology. Springer-verlag. New York,

USDA Forest Service, 1975. Photointerpretation guide for forest inventories. NAS 9-12200, technical report.

USDA Forest Service, 1994. Current vegetation survey. Resource inventory. Version 1.4r. Pacific Northwest Region. Portland, Oregon.

Waring, R., 1994. Overview of the Oregon Transect Ecosystem Research project. Ecological Applications. 4 (2): 211-225.

Waring, R. H. and Schlesinger, W. H., 1985. Forest ecosystem: concepts and management. Academic Press Inc., pp 39-70

Waring, R. H., P. E. Schroeder, and R. Oren, 1982. Application of the pipe model theory to predict canopy leaf area. Can. J. For. Res. 12: 556-560.

Wulder, Mike, S. Franklin, and Mike Lavigne, 1996. Statistical texture properties of forest structure for improved LAI estimates from casi. In proceedings of 26th International Symposium on Remote Sensing of Environment and 18th Annual Symposium of the Canadian Remote Sensing Society, March 25-29, 1996, Vacouver, BC, Canada.

Yoder, B. J. and R. E. Pettigrow-Crosby, 1995. Predicting nitrogen and chlorophyll content and concentrations from reflectance spectra (400-2500 nm) at leaf and canopy scales. Remote Sensing of Environment, 53:199-211.

Yoder, B. J. and Waring, R. H., 1994. The normalized difference vegetation index of small Douglas fir canopies with varying chorophyll concentrations. Remote sensing of environment, 49:81-91.

# Appendix

**Appendix. Numerical method for tree-based model**

The algorithm of tree-based model attempts to partitioning the space of predictor variables ( $X$ ) into homogenous regions, such that within each region the conditional distribution of $y$ given $x$, $f(y|x)$, does not depend on $x$. The numerical algorithm was presented here. Details refer to Chambers and Hastie (1992)

**Initialize:** current node = root = $\{y_i, i = 1,\ldots, n\}$

stack = NULL

**Recourse:** for current node <> NULL

**Loop:** for each $x_j$, partition $x$ into two sets $X_{left}$ and $X_{right}$ such that $f(y|X_{left})$ and $f(y|X_{right})$ are most different.

**Split node:** split current node into $X_{left}$ and $X_{right}$ according to the $x_j$ and the associated split that is best among all $x$'s.

**Test:** if *ok* to split $Y_{right}$

push $Y_{right}$ to stack

if *ok* to split $Y_{left}$

current node = $Y_{left}$

else *pop* stack