An Application of Artificial Intelligence Techniques in Classifying Tree Species with LiDAR and Multi-Spectral Scanner Data

Benedict Kit A Posadas

AN APPLICATION OF ARTIFICIAL INTELLIGENCE TECHNIQUES IN

CLASSIFYING TREE SPECIES WITH LiDAR AND

MULTI-SPECTRAL SCANNER DATA

By

Benedict Kit A. Posadas, Jr.

A Thesis
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in Forestry
in the Department of Forestry

Mississippi State, Mississippi

August 2008

AN APPLICATION OF ARTIFICIAL INTELLIGENCE TECHNIQUES IN

CLASSIFYING TREE SPECIES WITH LiDAR AND

MULTI-SPECTRAL SCANNER DATA

By

Benedict Kit A. Posadas, Jr.

Approved:

Emily B. Schultz
Associate Professor of Forestry
(Major Professor)

Thomas G. Matney
Professor of Forestry
(Committee Member)

David L. Evans
Professor of Forestry
(Committee Member)

James P. Shepard
Head and Graduate Coordinator of
the Department of Forestry

George M. Hopper
Dean of the College of Forest Resources

Name: Benedict Kit A. Posadas, Jr.

Date of Degree: August 9, 2008

Institution: Mississippi State University

Major Field: Forestry

Major Professor: Dr. Emily B. Schultz

Title of Study: AN APPLICATION OF ARTIFICIAL INTELLIGENCE TECHNIQUES IN CLASSIFYING TREE SPECIES WITH LiDAR AND MULTI-SPECTRAL SCANNER DATA

Pages in Study: 79

Candidate for Degree of Master of Science

Tree species identification is an important element in many forest resources applications such as wildlife habitat management, inventory, and forest damage assessment. Field data collection for large or mountainous areas is often cost prohibitive, and good estimates of the number and spatial arrangement of species or species groups cannot be obtained. Knowledge-based and neural network species classification models were constructed for remotely sensed data of conifer stands located in the lower mountain regions near McCall, Idaho, and compared to field data. Analyses for each modeling system were made based on multi-spectral sensor (MSS) data alone and MSS plus LiDAR (light detection and ranging) data. The neural network system produced models identifying five of six species with 41% to 88% producer accuracies and greater overall accuracies than the knowledge-based system. The neural network analysis that included a LiDAR derived elevation variable plus multi-spectral variables gave the best overall accuracy at 63%.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

TABLE		Page

TABLE | Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

CHAPTER I

INTRODUCTION

The potential for monitoring world forest conditions through visual image interpretation has been explored ever since the field of remote sensing began non-military applications. A primary driver for research in this area is lack of sufficient resources, such as time or manpower, to send assessment and management operations into the field. The use of remote sensing in forestry has had an emphasis on "timber management, maintenance and improvement of existing forest stands, and fire control" (Lillesand and Kiefer, 2000) since wood is the principal raw material from forests. Specific applications include "tree species identification, studying harvested areas, timber cruising, and forest damage assessment" (Lillesand and Kiefer, 2000).

1.1 Overview

Tree species identification is an important element in wildlife habitat assessment. Forested wildlife habitats can stretch for hundreds or thousands of acres, and accurately inventorying and mapping them is costly. Remote sensing techniques have been adopted in the hopes of reducing these costs. Aerial photogrammetry was one of the earliest remote sensing fields that investigated remotely identifying tree species (Lillesand and Kiefer, 2000). The extent tree species could be recognized in aerial photographs is dependent upon the scale and quality of the images, the variety and arrangement of species in the image, and the skill and experience of the interpreter. Commercial uses of more advanced remote sensing instruments were also explored as they became available.

Multi-spectral scanners (MSS), able to gather more data at a much greater range of the electromagnetic spectrum, have been used to successfully distinguish between conifers and hardwoods (Rohde and Olson, 1972) and between two Southern pine species (Hughes et al., 1986). Still, the effectiveness of multi-spectral scanners is dependent on several environmental and physiological factors, that are often variable over time. Consistent classification of species under varying conditions has not been demonstrated, but improvements have been made (Conradsen and Gunulf, 1986; Franklin et al., 2001).

LiDAR (light detection and ranging) is one of the more recently developed spatial technologies and has the ability to add site characteristics (Kraus and Pfeifer, 1998) and canopy attributes (Lefsky et al., 1999; Ritchie et al., 1993) to resource analysis. However, the library of imagery available to the general public is not as extensive geographically nor extends as far back in time as MSS data, and LiDAR data currently command a relatively high price. For these reasons, its use has been primarily limited to research.

For specific applications, remote sensing techniques can adequately measure many forest stand characteristics in a less time-consuming manner than through fieldwork. Reliable and consistent classification of tree species is critical to the future enhancement or replacement of field assessment by spatial technologies. The inclusion of LiDAR data to enhance the effectiveness of classifying forest species in conjunction with multi-spectral scanner data warrants investigation.

1.2 Objectives

The objectives of this research were two-fold. The first was to determine the feasibility of using a computer classification program to identify tree species or species groups using LiDAR and MSS data. The second was to compare and contrast four different approaches of artificially intelligent classification systems in identifying species or species groups. Two approaches used knowledge systems: one to classify a dataset containing MSS data and another for a dataset comprised of MSS plus LiDAR data. The other two approaches used a learning system to classify MSS and MSS plus LiDAR data.

The successful use of learning and knowledge systems to identify species from spatial data would be an important advance in the remote assessment of forest resources. A species GIS-type could be produced for a wide range of forest and urban applications, including inventories, forest health, and wildlife habitat types.

CHAPTER II

LITERATURE REVIEW

2.1 Remote Sensing Technologies and Forest Tree Species Identification

2.1.1 Multi-Spectral Scanners

Aerial photography and multi-spectral scanners display observed data as images. Both technologies record the spectral reflectance patterns of objects, in wavelengths within the visible light spectrum and beyond. However, sensors for MSS access a larger range of the electromagnetic spectrum (0.3 to 14 nanometers), observe in greater detail (very narrow bands of wavelength), and record a greater number of bands simultaneously than an aerial photo. Cameras for aerial photographs can only detect wavelengths of 0.3 to 0.9 nanometers and record three or four wide-bands at a time (Lillesand and Kiefer, 2000).

MSS data have been used to classify forest species composition. Aerial MSS data can be used to distinguish between conifers and hardwoods (Rohde and Olson, 1972) and to separate hardwood competition in pine plantations (Knight, 2003) with as high as 88% overall accuracy. Hughes and others (1986) were successful in using low-level MSS data to distinguish between loblolly pine (*Pinus taeda* L.) and longleaf pine (*Pinus palustris* Mill.). Casey (1999) used high resolution aerial platform digital frame camera data to classify four species groups in a mixed pine-hardwood forest with a 65% overall accuracy.

Since the effectiveness of MSS data is influenced by environmental and physiological factors, whatever affects the spectral reflectance of the observed stand determines how useful the data can be for tree species identification. Experimental conditions such as weather, time of year, and plant-soil relations are variable over time. Thus, the consistent use of MSS data by itself as a means of species classification has had its limitations. These limitations can be partially overcome by the use of *a priori* information (Conradsen and Gunulf, 1986) and spatial co-occurrence texture analysis (Franklin et al., 2001). LiDAR data can also be used in conjunction with airborne MSS data. Collins (2003) succeeded in using object oriented identification of tree crowns to separate seven species in southeastern United States hardwood stands with an overall accuracy of 54%.

2.1.2 LiDAR

LiDAR (light detection and ranging) is a technology conceptually similar to radar except it employs laser light instead of radio waves to map out the three-dimensional distribution of objects in an area. Pulses of light directed towards the ground bounce off solid objects and are reflected back to sensors. The time it takes for the light to travel back is used to calculate the distance between the objects and the sensors. Since modern LiDAR systems automatically georeference their measurements, the data gathered are readily compatible with GIS applications (Lillesand and Kiefer, 2000).

LiDAR has been used since the late 1970s to obtain terrain elevations for both land and water (Lillesand and Kiefer, 2000). The digital elevation models (DEM) constructed from LiDAR data provide a wealth of information on site variables associated with a forest stand. Modern LiDAR systems can also be used to measure forest stand structure characteristics both directly and indirectly from stands of forests. The location of the forest canopy and the bare ground can be determined, along with surfaces in-between (Lillesand and Kiefer, 2000), and used to derive total height at a geographic level as narrow as an individual tree or as broadly as a landscape. From these LiDAR-derived tree heights, an indication of the overall forest structure can be ascertained. Measurements of stand characteristics such as biomass, basal area, stand height, and vertical structure can be calculated from that starting point (Lefsky et al., 1999; Means et al., 1999; Nelson et al., 1988; Nilsson, 1996; Zimble, 2002).

LiDAR data have also been used to classify species groups. Discriminant analysis can successfully separate pines and mature hardwoods in the southeastern United States, based on the density and intensity of LiDAR returns, with an overall accuracy rate of 72% (Douglas, 2004). However, just like other light-based remote sensing technology, LiDAR is constrained by visibility. Inclement weather, line of sight, and other factors impact the effectiveness of this application which is also still relatively expensive for large area analysis.

2.2 Learning and Knowledge Systems

Learning and knowledge systems are computer programs used to solve complex real-world problems (Weiss and Kulikowski, 1991). Learning systems use computers to extract patterns from previously solved cases to make predictions on new cases. Knowledge, or expert, systems use rules established by human experts to make predictions on new cases.

These two classification systems are complementary in their strengths and weaknesses (Weiss and Kulikowski, 1991). The success of learning systems in correctly predicting data is often overshadowed by the initial success rate of knowledge systems. A knowledge system is already equipped with the relationships that connect the data and can readily process it. A learning system must derive those relationships. However, a learning system is not as bound by human bias and may uncover a previously unknown connection, perhaps making it more effective than a knowledge system.

Learning systems can be thought of as self-correcting programs that constantly alter their decision biases in order to match their processed output with the actual observed data. A neural network learning system (Principe and others, 2000; Mehrotra and others, 1997) simulates a biological nervous system where numerous processing elements occur simultaneously and are interconnected. Knowledge systems, on the other hand, are like digital experts (Stefik, 1995). They start out already capable of processing data and classify them according to known properties and relationships.

Both neural networks and knowledge-based systems have been utilized to limited degrees in modeling and classification problems in natural resource management. Gimblett and Ball (1995) used neural networks to generate management units in the Hoosier National Forest. Gopal and Woodcock (1996) examined the predictability of conifer mortality from changes in spectral data over time. Skirvin and Dryden (1997) used neural networks to classify Landsat TM image data. Moshou and others (2001) used neural networks to classify crops and weeds by their spectral properties. Meador (2002) compared traditional parametric and neural network models in predicting the growth and yield of longleaf pine.

Van Aardt and Wynne (2001) used remote sensing-based classification schemes to quantify the spectral separability of six tree species in the Appomattox Buckingham State Forest. Zhang and others (2000) compared the effectiveness of a neural network against a knowledge-based system in identifying red tides and coastal plumes from Coastal Zone Color Scanner images of the West Florida Shelf.

Natural resource applications in artificial intelligence are in their infancy but appear to have great potential for problem domains where datasets are 1) extremely large, 2) without existing predefined functional relationships, and 3) where complicated interactions are involved (Schultz et al., 1999). Applications of remote sensing data often fall into the above three categories and are logical candidates for analysis using artificially intelligent methods.

CHAPTER III

METHODS – GENERAL

3.1 Study Site Description

LiDAR, MSS, and field data were acquired of conifer stands located in the lower mountain regions near McCall, Idaho, with the funding of the U.S. Forest Service. There were two study sites or blocks (Figure 3.1) encompassing an area of approximately 10,000 acres (4047 ha) with each block containing 5,000 acres (2023 ha). The NE study block had a mean elevation of 7497 ft (2285 m), while the SE study block had a mean elevation of 5115 ft (1559 m). These sites were chosen for remotely classifying individual tree species because of the availability of both LiDAR and MSS data and open stands providing clear separation of tree crowns.

The six tree species common to the area were Douglas-fir (*Pseudotsuga menziessii* [Mirb.] Franco.), grand fir (*Abies grandis* [Dougl.] Lindl.), subalpine fir (*Abies lasiocarpa* [Hook.] Nutt.), Engelmann spruce (*Picea Engelmannii* Parry), Ponderosa pine (*Pinus ponderosa* Laws.), and lodgepole pine (*Pinus contorta* Dougl.).

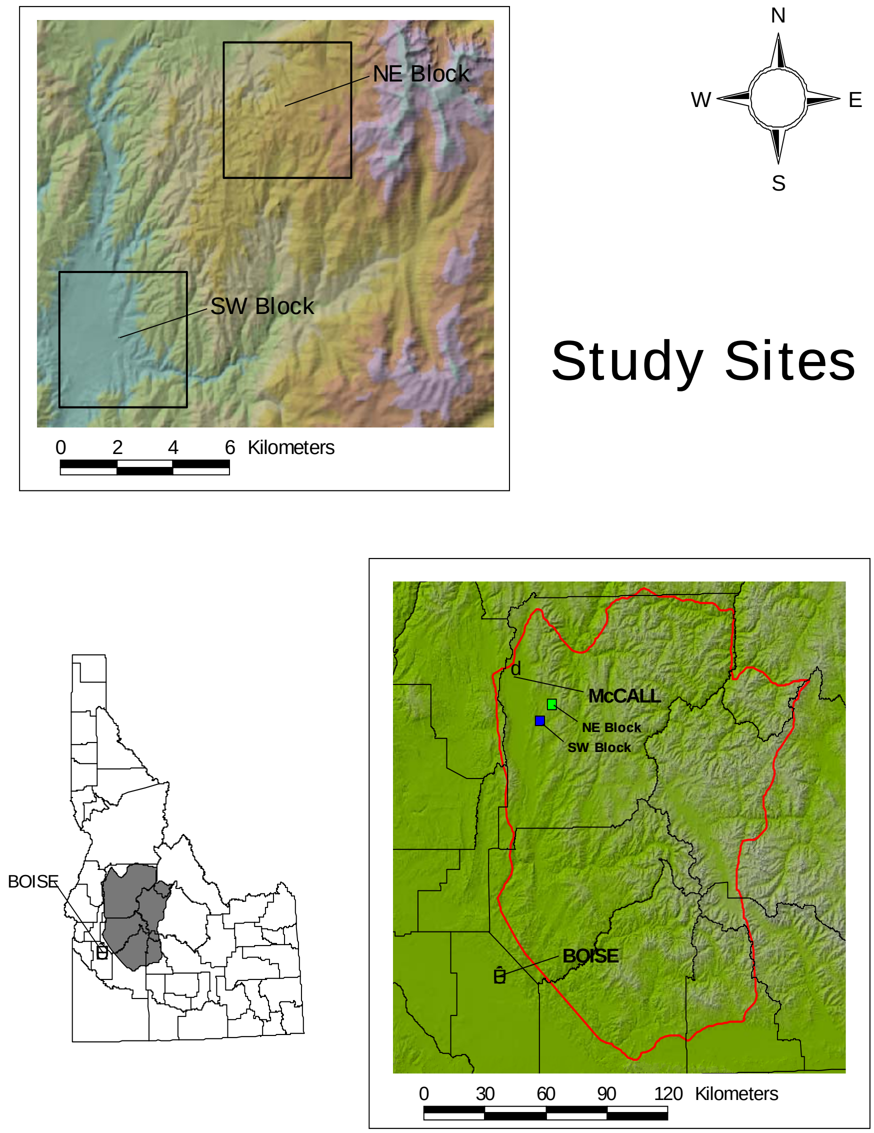

Figure 3.1 Location of the two study sites near McCall, Idaho (Zimble, 2002)

3.2 Data Collection and Preparation

A small-footprint, multi-return LiDAR mission was flown by EarthData Technologies on October 12, 1998. A custom-built Azimuth Aeroscan LiDAR system acquired complete coverage of the two study sites (Zimble, 2002). Airborne MSS 0.32m resolution data were collected in October of 1999 by Space Imaging, Inc. using a DAIS sensor (Bands 1: 0.45 – 0.52μ, 2: 0.52 – 0.60μ, 3: 0.63 – 0.69μ, and 4: 0.76 – 0.90μ).

A total of 49 plots within the two study sites were available for analysis from previous work by Zimble (2002). In the summer of 2001, field data from these two study sites were collected as part of research to investigate vertical stand structure. The inventory of these 49 plots included field measurements taken on trees, snags, and coarse woody debris within each observation plot. Data included survey counts, height measurements, and global positioning recordings. Images derived from the LiDAR and MSS data were also available for analysis. Latitude and longitude measurements were taken on 134 trees across the 49 plots but were not evenly distributed among the six species.

In order to minimize bias in classification procedures, data from 30 additional field plots were collected in the summer of 2002 to more evenly represent species and to provide a greater number of observations for analysis. Recorded field data for all 79 plots included tree species, GPS location, total tree height, height to base of crown height, crown radii, and DBH.

Information concerning the horizontal dimensions of plot and snag trees was compiled to create a graphical representation of each tree. The GPS location (latitude and longitude) of each tree and the crown radii measurements along each cardinal direction (north, south, east, and west) were used to generate polygons that visually represented the trees in the data. The vector-based polygons were converted to raster data in order to visually compare the field data with the LiDAR and MSS data. A total of 338 to 342 trees were identified for use in building the classification models and were distributed as evenly as possible among the six species (Engelmann spruce, ponderosa pine, lodgepole pine, Douglas-fir, grand fir, and subalpine fir), with lodgepole pine as the most numerous and subalpine fir as the least numerous.

ArcView 3.3 (ESRI, 1998) was used to graphically represent the field data. ArcView is a geographic information system software used to view spatial data, create maps, and perform basic spatial analysis. ERDAS IMAGINE (ERDAS, 1997) was used to export images containing site variables such as slope, aspect, concavity/convexity, and elevation from LiDAR ground return readings. ERDAS IMAGINE is a raster graphics editor and remote sensing application. It was also used for the orthorectification of MSS imagery, the process that corrects much of the distortion inherent in aerial imagery.

3.3 Classification Systems

Structural measurements and spectral reflectance variables formed the basis for species classification. Because of the underlying differences in the technologies of LiDAR and MSS data systems, the LiDAR data provided more stand level characteristics while the MSS data provided more individual tree characteristics. Individual species classification models were developed using a knowledge-based system and a neural network system for MSS data alone and MSS plus LiDAR data.

NeuralWorks Professional II/Plus (NeuralWare, 2000a) and eCognition 2.1 (Definiens Imaging, 2003) software packages were chosen for neural network and knowldege-based analysis, respectively, based on prior research and experience. Unlike software that uses the per-pixel approach to image classification, eCognition software is an object-oriented image analysis system that relies upon the segmentation of remotely sensed images derived from factors such as size, color, shape, and smoothness allowing the designation of trees crowns as objects. NeuralWorks Professional II/Plus provides several well-known neural network types (such as backpropagation) as built-in features and can create custom networks though script files.

CHAPTER V

KNOWLEDGE-BASED CLASSIFICATION

4.1 Methods

4.1.1 Approach

The knowledge-based classification approach was divided into two classification schemes. The first classification scheme utilized multi-spectral scanner (MSS) characteristics alone to differentiate trees. The second classification scheme combined LiDAR data with the MSS data, providing measurable site characteristics in addition to the spectral characteristics. Each classification scheme is summarized in a flowchart depicted in Figures 4.1 and 4.2.

The object-oriented image analysis software, eCognition 2.1, was used to create the classification schemes. Data stored in the remotely sensed images were broken up into image segments larger than individual pixels, with the image objects segmented by color and shape. The user determines whether the image segmentation relies primarily upon color or upon shape in creating the objects (Figure 4.3). The user also decides what scale to use, whether to break down the image into hundreds of objects or into hundreds of thousands of objects. Depending on the complexity of the data, multiple attempts at segmentation can be used simultaneously in creating the classification scheme (Figures 4.4 and 4.5).

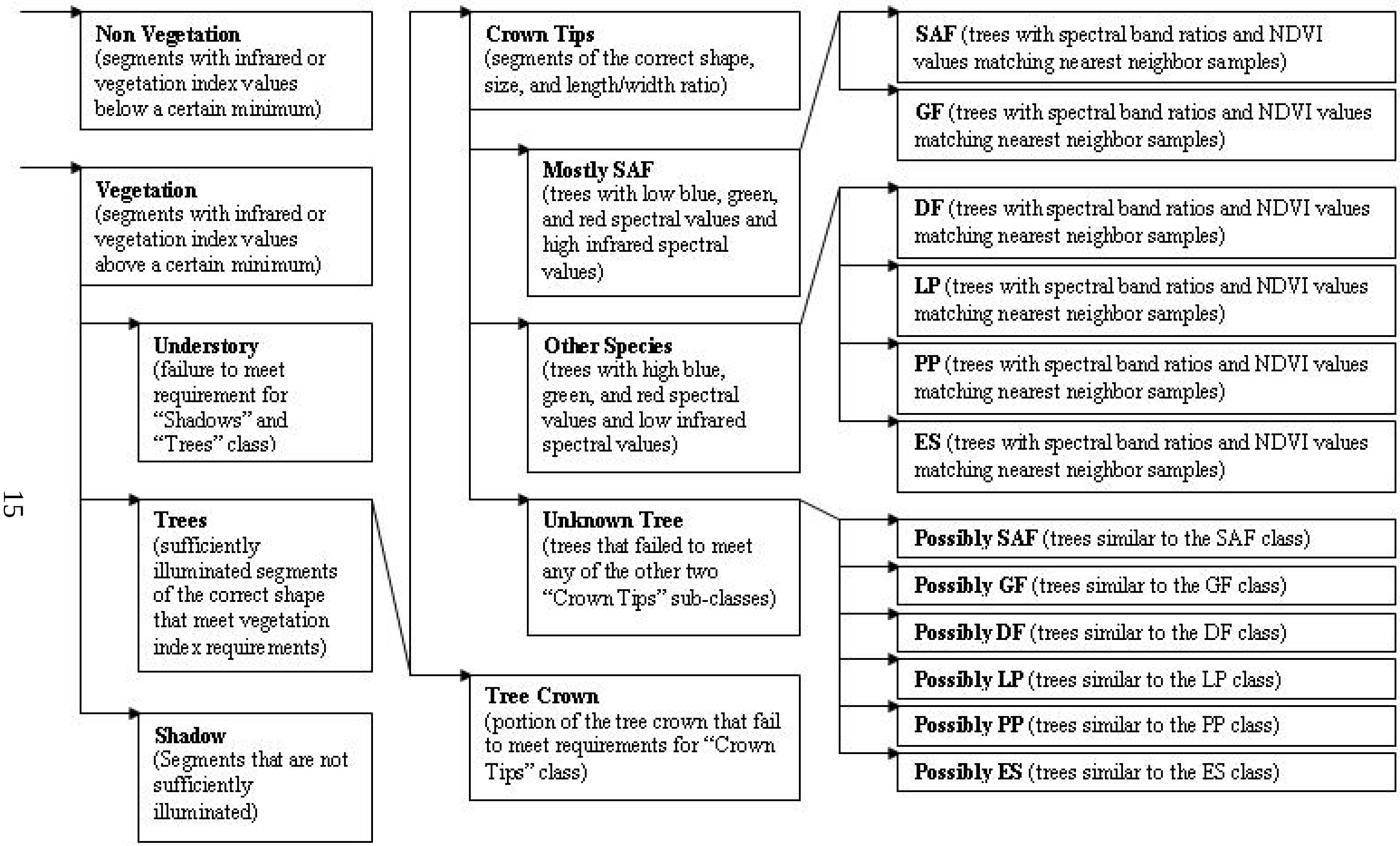

Figure 4.1 Diagram of the MSS data classification scheme used to separate six conifer species in the lower mountain regions near McCall, Idaho

Where, LP = Lodgepole pine, PP = Ponderosa pine, ES = Engelmann spruce, DF = Douglas-fir, GF = Grand fir, and SAF = Subalpine fir

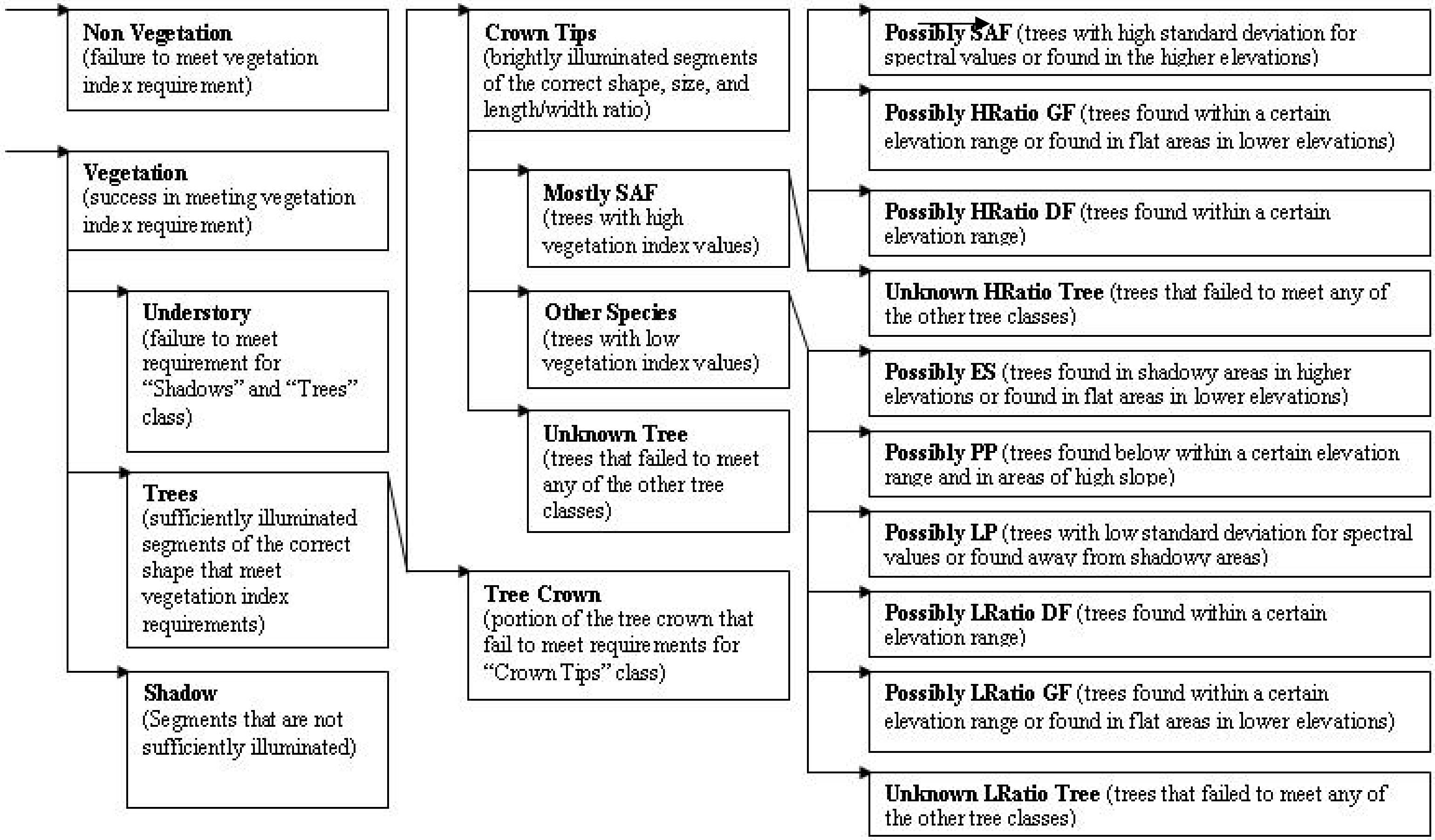

Figure 4.2 Diagram of the MSS plus LiDAR classification scheme used to separate six conifer species in the lower mountain regions near McCall, Idaho

Where, LP = Lodgepole pine, PP = Ponderosa pine, ES = Engelmannn spruce, DF = Douglas-fir, GF = Grand fir, SAF = Subalpine fir.

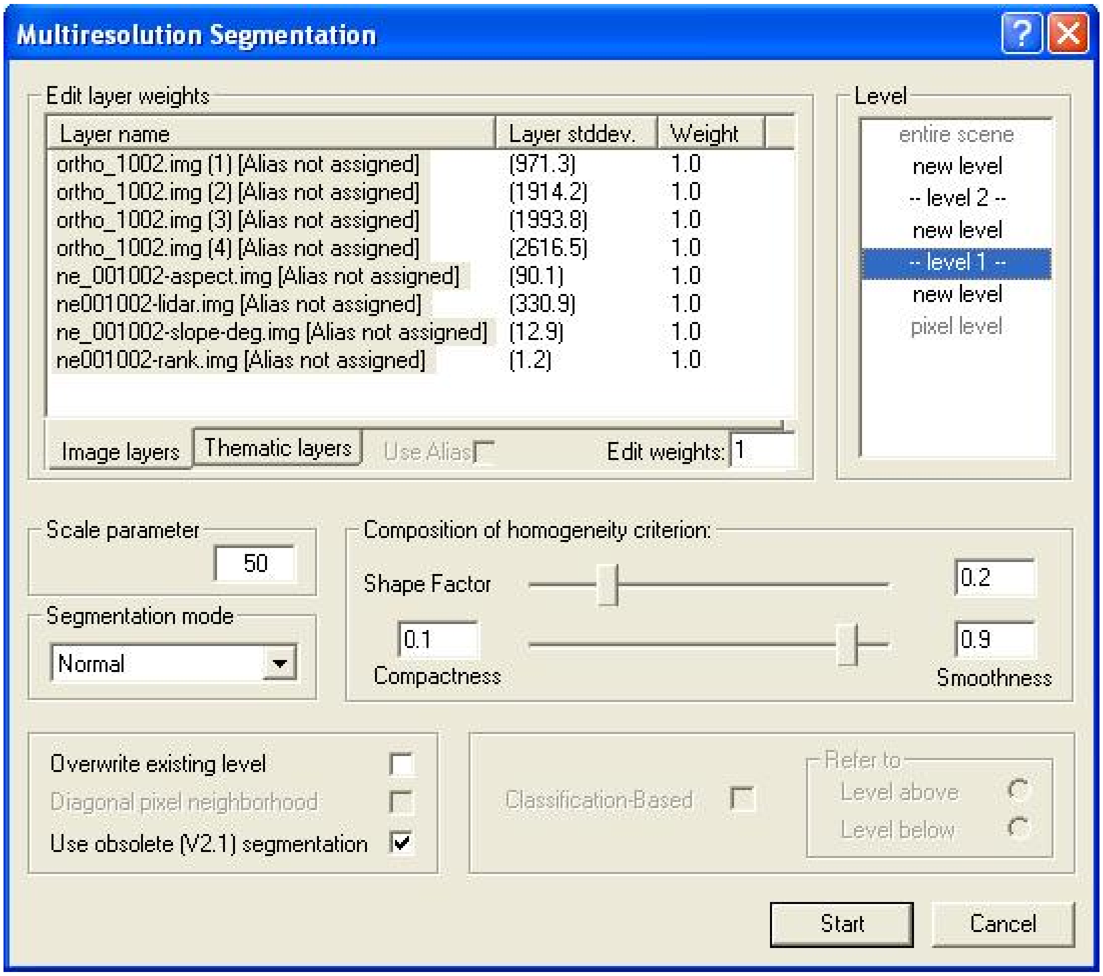

Figure 4.3 Multiresolution segmentation options available for use in eCognition Professional 4.0

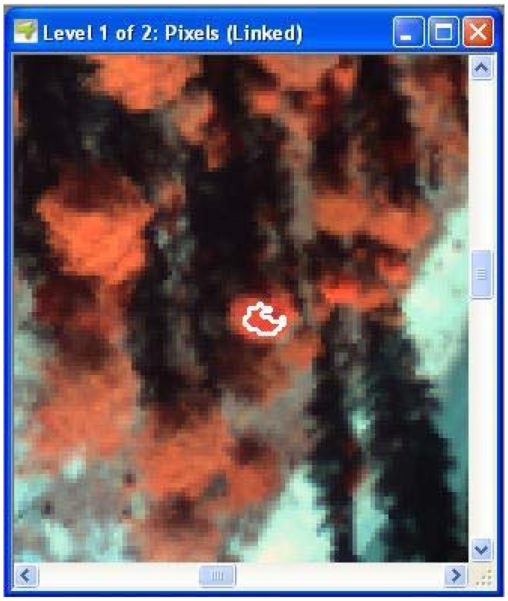

Figure 4.4 Image object in eCognition at a scale parameter of 50

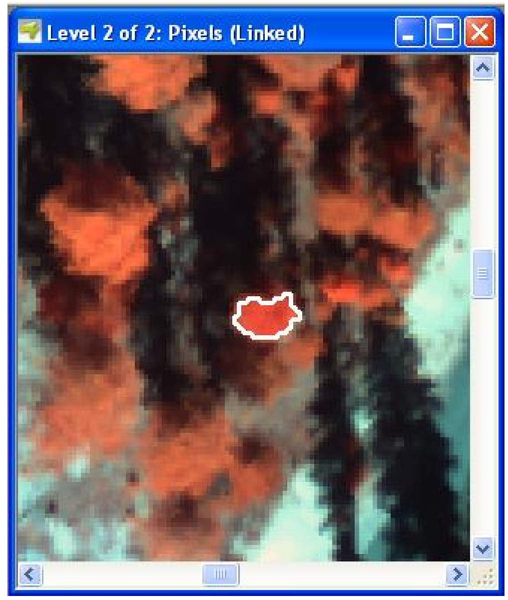

Figure 4.5 Image object in eCognition at a scale parameter of 150

4.1.2 Features and Membership Functions in eCognition

Features and membership functions define the predetermined classes into which eCognition objects are sorted. Features refer to the values used in determining the criteria for each class while membership functions characterize the probability that an object will match the criteria of a class.

The data used in determining classes are separated into object features and class-related features (Figure 4.6). They are either provided by eCognition 2.1 or are custom designed to work in eCognition. Vegetation indices are not readily available in eCognition and had to be derived and added as customized object features. Vegetation indices are transformations of the reflectance values of spectral bands that have been useful in identifying vegetative presence and activity (Jensen, 2005). The Normalized

Difference Vegetation Index (NDVI) was calculated as the quantity of the near-infrared band minus the red band divided by the quantity of the near-infrared band plus the red band (Figure 4.7).

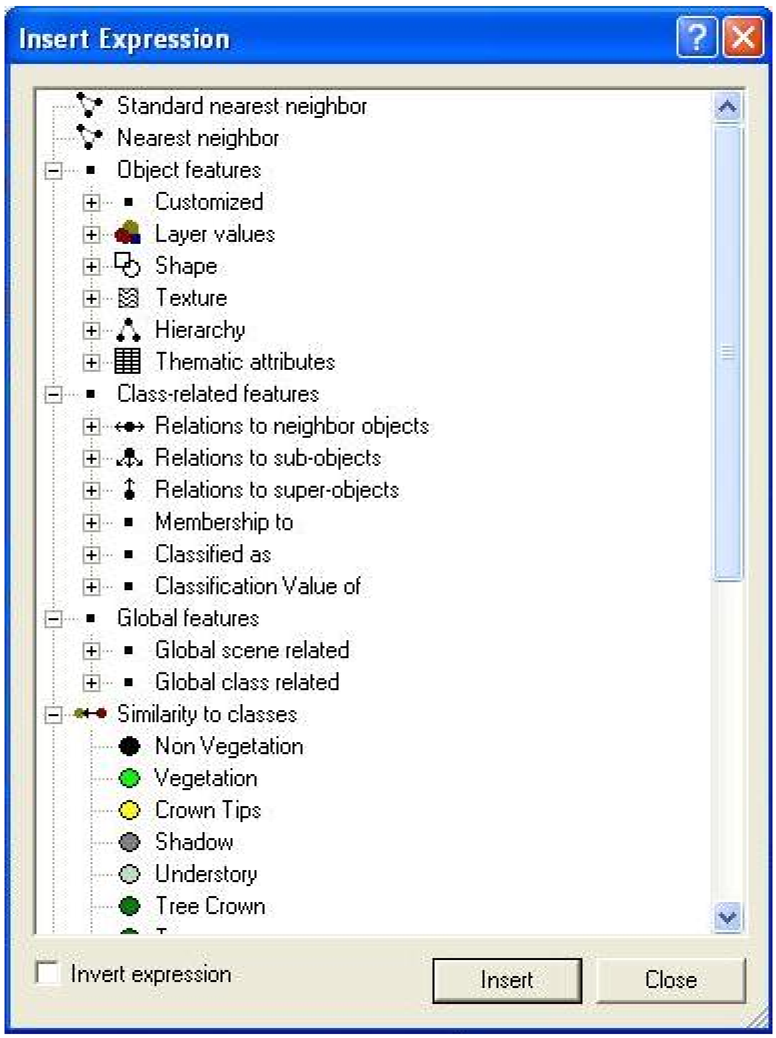

Figure 4.6 Object and class-related features available for use in eCognition Professional 4.0

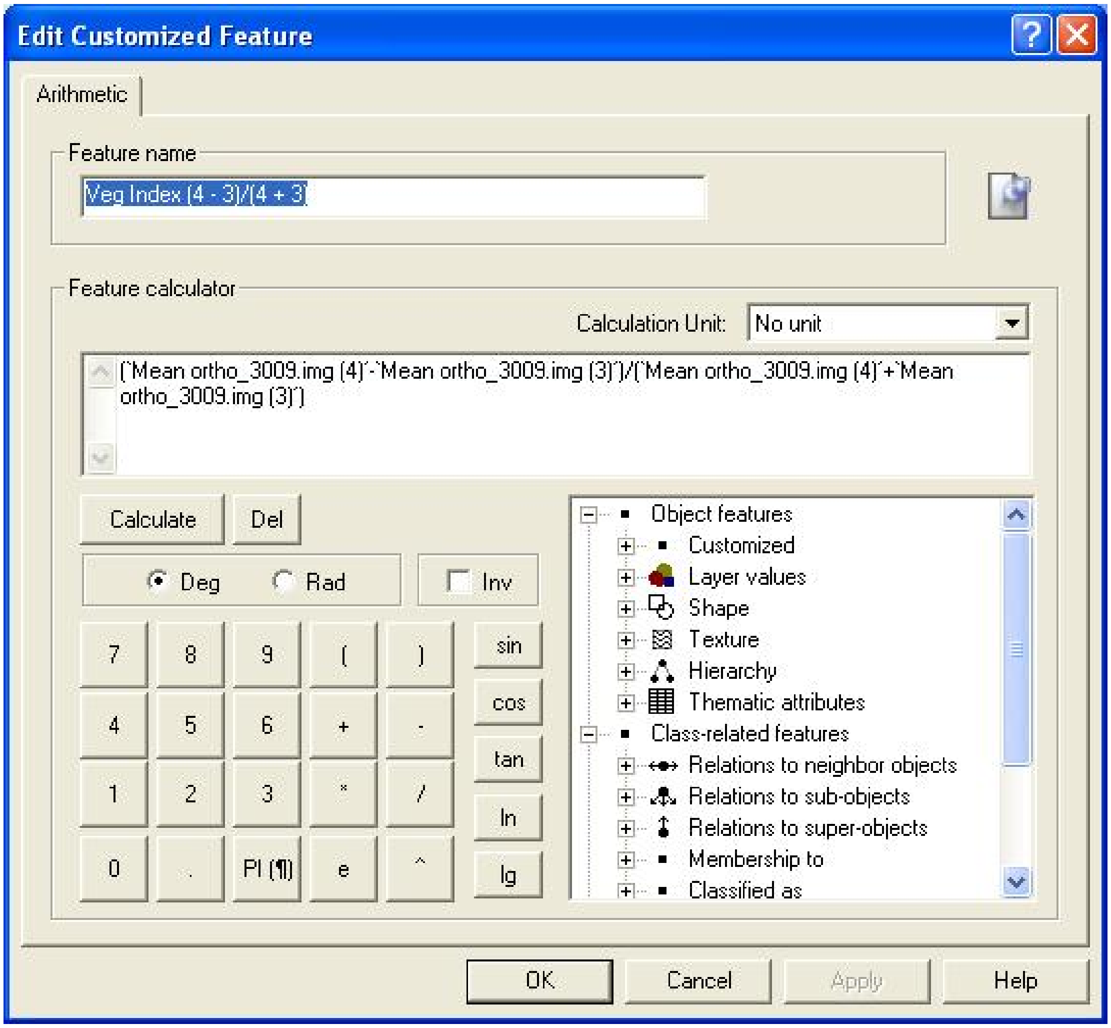

Figure 4.7 Calculation used to derive NDVI in eCognition

Object features are the arithmetic values of the segments and the data comprising the segments, along with associated statistics. eCognition provided many object features such as brightness, border length, shape index, and spectral band ratio that were used in the knowledge-based classification schemes (Appendices A and B).

Class-related features are relational. They compare and contrast object features of one segment with those of another segment. Together with any statistics associated with those comparisons, they inherit the results of past classifications and incorporate these features into the current classification.

The membership functions (Appendix Tables A.1 to A.7 and B.1 to B.7) of a class are primarily defined by the shape of the probability slope and the range of values (maximum, center point, and minimum) comprising the function slope. The following membership function definitions are examples used in the classification process.

In the form "Larger Than," the image object is classified as a member 100% of the time if its feature value is greater than the maximum range value of the membership function. It is classified as a member from 100% down to 50% of the time if its feature value is between the center point and the maximum range value. It is classified as a member from 50% down to 0% of the time if its feature value is between the minimum value and the center point. It is classified as a member 0% of the time if its feature value is less than the minimum range value. In the form "Smaller Than," the membership percentages are reversed (Figure 4.8).

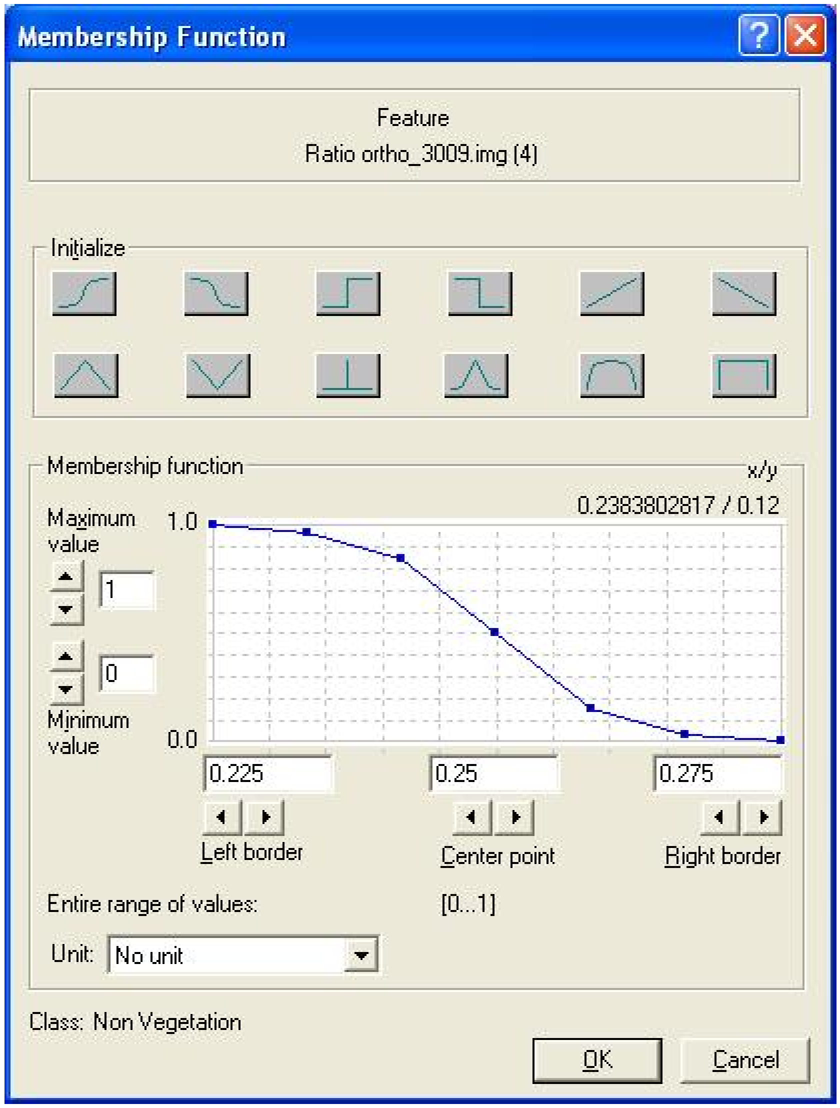

Figure 4.8 "Smaller Than" probability slope used in a membership function in eCognition

In the form “Larger Than (Boolean),” the image object is classified a member 100% of the time if its feature value is greater than the center point and is classified a member 0% of the time if its feature value is less than the center point. In the form “Smaller Than (Boolean),” the membership percentages are reversed. In the form “Approximate Gaussian,” the image object is classified using a Gaussian curve, with the center point signifying the 100% marker. In the form “Full Range,” all of the values between the minimum and the maximum range values are classified as a member 100% of the time.

4.1.3 Isolating Trees

The first step in the classification process involved isolating trees from the rest of the image. The segmented data from the MSS imagery (Appendix Table A.1) and from the combined MSS and LiDAR imagery (Appendix Table B.1) were divided into “Non Vegetation” and “Vegetation.” Objects classified as “Vegetation” were in turn split up into three groups: “Trees,” “Understory,” and “Shadow” (Appendix Tables A.2, B.2, and B.3). Each class had its own criteria with requirements exclusive to the other classes. The portions of the images identified as “Trees” were further divided into “Crown Tips” and “Tree Crowns” (Appendix Tables A.3 and B.4). This step facilitated the identification of tree species. If the classification were confined to one portion of the tree crown, the classification of different portions of the same tree as differing species would be less likely. Crown tips receive the most illumination; thus, they are less likely to have aberrant spectral characteristics. Segments classified as “Crown Tips” were separated from “Tree Crowns” according to criteria involving their size, shape, and illumination.

Ideally, the criteria created would apply universally across all the MSS imagery. Unfortunately, trees from the study sites differed in the size of their crowns even within the same species. Several of the larger trees contained multiple segments that were classified as “Crown Tips”.

4.2 Model Building

Once the image segments representing the test trees were identified through their GPS locations, a random sample of each tree species was taken to examine its spectral properties. The eCognition software displayed numerous statistical derivations of the spectral data graphically, which facilitated the examination of spectral properties for the sample tree. The sample values were then tabulated to determine the range for each species.

The ratio values for each spectral band were initially the most useful. A spectral band ratio is the mean band reflectance value for an image object divided by the sum of all the spectral band mean values for that object (Definiens Imaging, 2003). Each MSS image had different lighting conditions, thus adding a randomizing factor to the mean spectral values for each image object. Patterns that held true in one set of images would no longer hold true in another group. Band ratios normalized the spectral values and allowed observations to apply for all the MSS imagery. As the research progressed, the NDVI values (red band) and the other vegetation indices (blue and green bands) derived from it were also found to be useful in separating the species.

4.2.1 Multi-Spectral Scanner Imagery

Working with the MSS only dataset, at least one tree species, subalpine fir, could be partially isolated from all other species. Within the range of values that formed eCognition ratios of each spectral band (Figures 4.9 – 4.11), subalpine fir consistently appeared at one end of the range. Some species overlapping occurred, but in general, a dividing line could be found between subalpine fir (SF) and a group of three other species: Engelmann spruce (ES), ponderosa pine (PP), and lodgepole pine (LP). Douglas-fir (DF) and grand fir (GF) straddled the middle values that overlapped the two extremes.

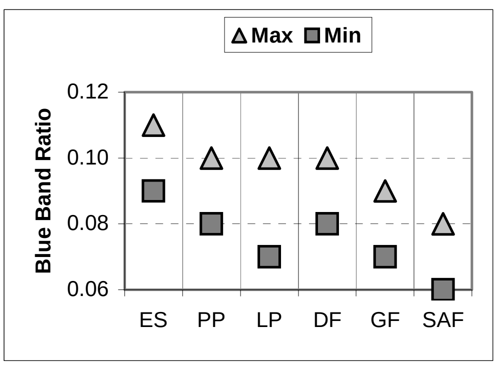

Figure 4.9 The range of eCognition ratio values within the blue spectral band in the multi-spectral sensor data associated with each of the six species found in the lower mountain regions study site near McCall, Idaho

Where, LP = Lodgepole pine, PP = Ponderosa pine, ES = Engelmannn spruce, DF = Douglas-fir, GF = Grand fir, SAF = Subalpine fir

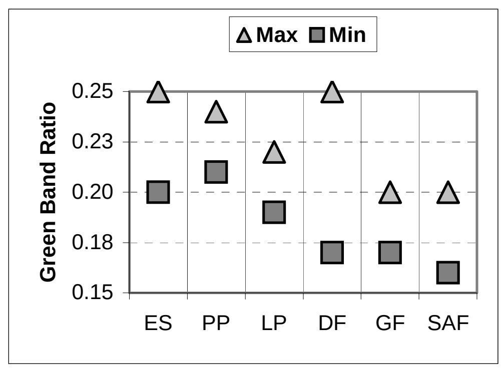

Figure 4.10 The range of eCognition ratio values within the green spectral band in the multi-spectral sensor data associated with each of the six species found in the lower mountain regions study site near McCall, Idaho

Where, LP = Lodgepole pine, PP = Ponderosa pine, ES = Engelmannn spruce, DF = Douglas-fir, GF = Grand fir, SAF = Subalpine fir

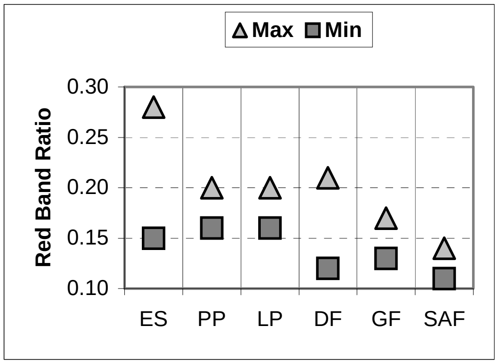

Figure 4.11 The range of eCognition ratio values within the red spectral band in the multi-spectral sensor data associated with each of the six species found in the lower mountain regions study site near McCall, Idaho

Where, LP = Lodgepole pine, PP = Ponderosa pine, ES = Engelmannn spruce, DF = Douglas-fir, GF = Grand fir, SAF = Subalpine fir.

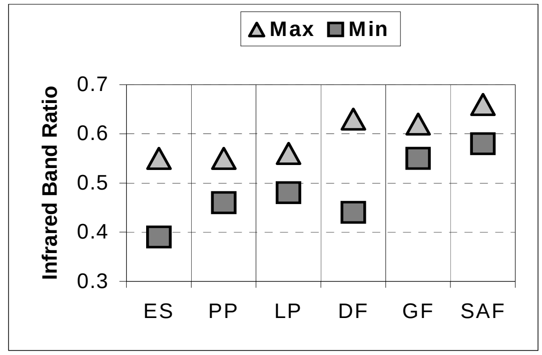

Figure 4.12 The range of eCognition ratio values within the infrared spectral band in the multi-spectral sensor data associated with each of the six species found in the lower mountain regions study site near McCall, Idaho

Where, LP = Lodgepole pine, PP = Ponderosa pine, ES = Engelmannn spruce, DF = Douglas-fir, GF = Grand fir, SAF = Subalpine fir.

Engelmann spruce and ponderosa pine were only found in the upper end of the blue spectral band (Figure 4.9) while subalpine fir was located in the lower end of the spectrum. An almost identical distinction among species groups was produced by the green (Figure 4.10) and red (Figure 4.11) spectral bands. In the infrared spectral band (Figure 4.12), the positioning of the species groups was reversed. Subalpine fir and grand fir occupied the upper extremes while Engelmann spruce, ponderosa pine, and lodgepole pine occupied the lower extremes.

Grand fir shared similar spectral values to subalpine fir in the blue (Figure 4.9), green (Figure 4.10), red (Figure 4.11), and infrared (Figure 4.12) spectral bands, though its range did not extend as far into the extremities. Lodgepole pine spectral values were similar to those of ponderosa pine (Figures 4.9, 4.11, and 4.12). Douglas-fir had a range straddling the middle of the spectrum (Figure 4.9), with values sometimes reaching the extremes (Figures 4.10 and 4.11).

Subalpine fir was the only species that could be identified in the model-building process from the combined data of all four MSS bands (Appendix Table A.4). The spectral band ratios of lodgepole pine, ponderosa pine, Engelmann spruce, and Douglas-fir could not be separated effectively. Trees that matched the spectral signature cutoff for subalpine fir but failed to meet further criteria for subalpine fir had a high possibility of being grand fir or Douglas-fir. The spectral band ratio values of the these two species occupied a broad range that was separable from the narrow range of spectral values for subalpine fir. As Douglas-fir and grand fir are the only species that have spectral values overlapping subalpine fir, they are the most likely to fall outside the range of the intertwined group and fail to meet the further requirements for subalpine fir.

Nearest neighbor classification was used to further the separability between subalpine fir and the other tree species. Nearest neighbor is the automatic generation of multidimensional membership functions based on sample objects (Definiens Imaging, 2003). It can be used to create criteria for classes the user has no experience in determining or to supplement classes the user has created. Samples were based on vegetation index values and spectral band ratios. Segments belonging to the "Crown Tips" class were isolated into "Mostly SAF," "Other Treetops," and "Unknown

Crown Tips" subclasses (Appendix Tables A.4 to A.7). Those with low values in the blue, green, and red spectral bands and high values in the infrared bands were sorted into the "Mostly SAF" class. Those with high values in the blue, green, and red spectral bands and low values in the infrared bands were sorted into the "Other Treetops". Those with values that matched neither criterion were sorted into "Unknown Crown Tips".

"Mostly SAF" (Appendix Table A.5) was subdivided into the "SAF" and "GF" classes. Segments in this grouping were thought to be subalpine fir primarily, with some grand fir that had similar spectral characteristics. "Other Treetops" (Appendix Table A.6) was subdivided into the "DF," "LP," "PP," and "ES" classes. Nearest neighbor classification was used to compare the segments of both groups against the profiles of known samples containing the spectral band ratios of all four MSS bands and the Normalized Difference Vegetation Index.

The spectral characteristics of a tree did not always match the extremes of their group. They overlapped with other species in one or more of the spectral bands. Segments that did not fulfill all of the spectral requirements for the "Mostly SAF" class and the "Other Treetops" were sorted into the "Unknown Crown Tips" class (Appendix Table A.7). Still, even if not all the spectral requirements were met, those same segments could come close enough to match those classified under the nearest neighbor classification. The segments similar to the "SAF" class were sorted into the "Possibly SAF" class. Those similar to the "GF" class were sorted into the "Possibly GF" class. Those similar to the "DF" class were sorted into the "Possibly DF" class.

Those similar to the "LP" class were sorted into the "Possibly LP" class. Those similar to the "PP" class were sorted into the "Possibly PP" class. Those similar to "ES" class were sorted into the "Possibly ES" class.

4.2.2 <u>Multi-Spectral Scanner plus LiDAR-Derived Imagery</u>

Several of the inseparable tree species have unique site requirements that should allow further differentiation to take place. Ponderosa pine, with spectral characteristics similar to Engelmann spruce, is found in the steep slopes of mountainous sites. Engelmann spruce, on the other hand, prefers flatter and wetter sites, such as valleys, bottoms, and northern slopes of mountain sides (Burns and Honkala, 1990). Lodgepole pine, with spectral characteristics that overlap ponderosa pine primarily, can be found in a wider range of site conditions. Lodgepole pine can reach higher elevations than ponderosa pine, is found in both wet and dry areas, and grows in both steep and flat terrain.

LiDAR data were used to help characterize site variables and further differentiate among species. Digital elevation models and other information, such as terrain slope and flatness, and slope aspect, can be accurately derived from LiDAR data. These data were combined with the MSS data to evaluate any additional improvement in tree species separation. In adding LiDAR-derived image layers to MSS image layers, the spectral band ratios provided by eCognition were no longer useful in separating individual species or species groups (Figures 4.13 to 4.16). Clear species distinctions found previously by working with MSS imagery alone were no longer valid. This may have been due to differences in the segmentation of the two datasets.

The spectral value separation of the six tree species was recreated by the development of customized vegetation indices (Figures 4.17 – 4.19).

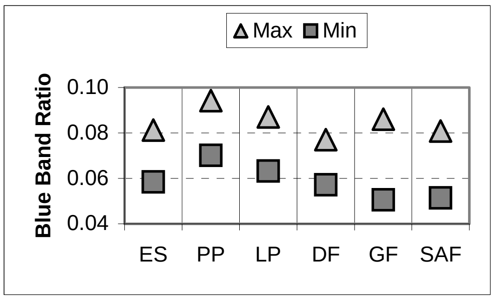

Figure 4.13 The range of eCognition ratio values within the blue spectral band in the combined multi-spectral sensor and LiDAR data associated with each of the six species found in the lower mountain regions study site near McCall, Idaho

Where, LP = Lodgepole pine, PP = Ponderosa pine, ES = Engelmannn spruce, DF = Douglas-fir, GF = Grand fir, SAF = Subalpine fir.

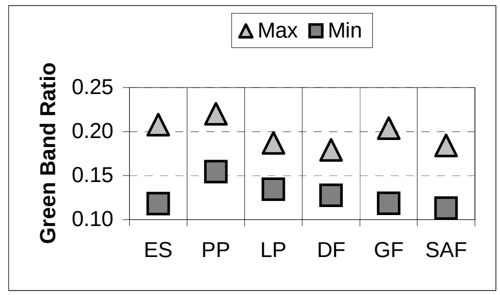

Figure 4.14 The range of eCognition ratio values within the green spectral band in the combined multi-spectral sensor and LiDAR data associated with each of the six species found in the lower mountain regions study site near McCall, Idaho

Where, LP = Lodgepole pine, PP = Ponderosa pine, ES = Engelmannn spruce, DF = Douglas-fir, GF = Grand fir, SAF = Subalpine fir.

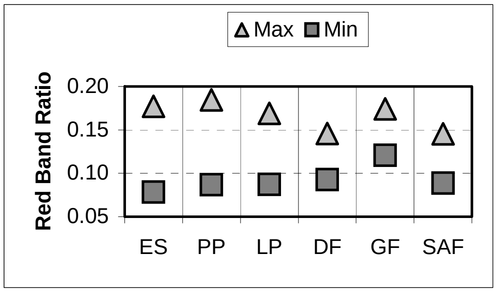

Figure 4.15 The range of eCognition ratio values within the red spectral band in the combined multi-spectral sensor and LiDAR data associated with each of the six species found in the lower mountain regions study site near McCall, Idaho

Where, LP = Lodgepole pine, PP = Ponderosa pine, ES = Engelmannn spruce, DF= Douglas-fir, GF = Grand fir, SAF = Subalpine fir.

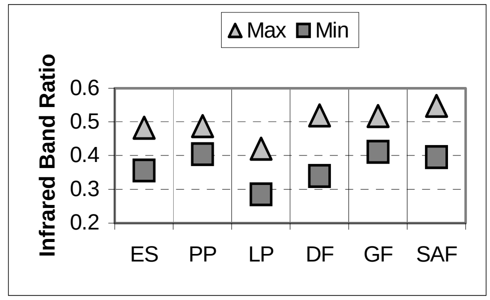

Figure 4.16 The range of eCognition ratio values within the infrared spectral band in the combined multi-spectral sensor and LiDAR data associated with each of the six species found in the lower mountain regions study site near McCall, Idaho

Where LP = Lodgepole pine, PP = Ponderosa pine, ES = Engelmannn spruce, DF = Douglas-fir, GF = Grand fir, SAF = Subalpine fir

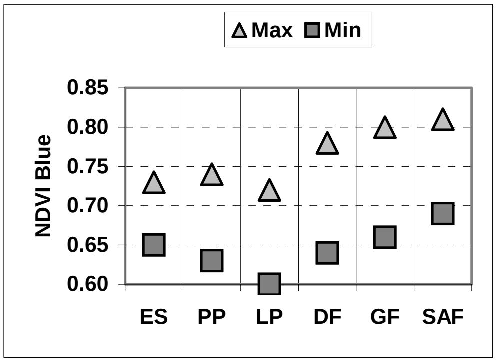

Figure 4.17 The range of NDVI values within the blue spectral band in the combined multi-spectral sensor and LiDAR data associated with each of the six species found in the lower mountain regions study site near McCall, Idaho

Where, LP = Lodgepole pine, PP = Ponderosa pine, ES = Engelmannn spruce, DF = Douglas-fir, GF = Grand fir, SAF = Subalpine fir

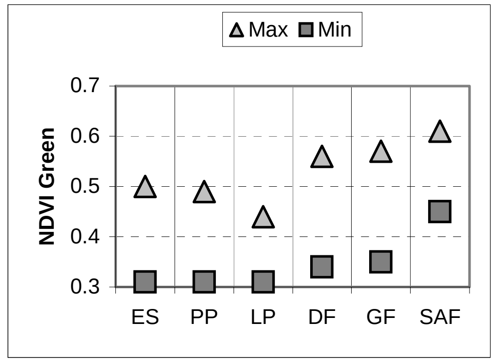

Figure 4.18 The range of NDVI values within the green spectral band in the combined multi-spectral sensor and LiDAR data associated with each of the six species found in the lower mountain regions study site near McCall, Idaho

Where, LP = Lodgepole pine, PP = Ponderosa pine, ES = Engelmannn spruce, DF = Douglas-fir, GF= Grand fir, SAF = Subalpine fir

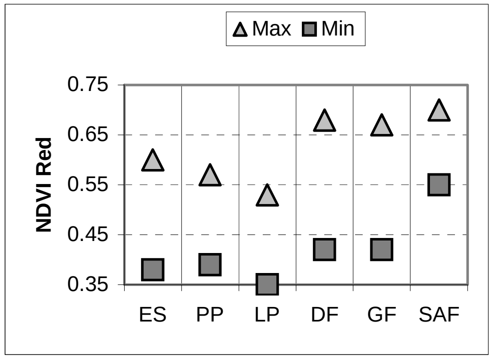

Figure 4.19 The range of NDVI within the red spectral band in the combined multi-spectral sensor and LiDAR data associated with each of the six species found in the lower mountain regions study site near McCall, Idaho

Where, LP = Lodgepole pine, PP = Ponderosa pine, ES = Engelmannn spruce, DF = Douglas-fir, GF = Grand fir, SAF = Subalpine fir

Subalpine fir, along with the upper range of values for grand fir and Douglas-fir, formed the upper extreme of the vegetative indices. Lodgepole pine, the lower range of values for grand fir and Douglas-fir , and the remaining species formed the lower extreme. Due to the distribution of these spectral values, tree segments isolated into the “Crown Tips” class could now be divided into two groups to aid identification: “Mostly SAF” and “Other Species” (Appendix Table B.5). Segments with calculated vegetation index values that fell below a cutoff point were grouped as “Other Species” while those with spectral values above the cutoff point were grouped as “Mostly SAF.” Those segments that failed to match the requirements for either sub-class were sorted into the “Unknown Tree” class.

The group “Mostly SAF” is composed of subalpine fir, grand fir, and Douglas-fir. It was subdivided into the classes of “Possibly SAF”, “Possibly HRatio GF”, “Possibly HRatio DF”, and “Unknown HRatio Tree” (Appendix Table B.6). The term “HRatio” refers to higher spectral values associated with these segments. Though subalpine fir and grand fir share similar spectral characteristics, the ecological range of grand fir does not reach elevations as high as those of subalpine fir (Burns and Honkala, 1990). It also does not tolerate the wide range of site conditions at higher elevations. Douglas-fir has a wider range in elevation and site conditions than grand fir, but it does not grow as high in elevation as subalpine fir. At higher elevations, subalpine fir is far more likely to occur than Douglas-fir or grand fir. Taking into account that subalpine fir can reach higher standard deviation values for its spectral bands (Tables 4.20 to 4.23), it is still separable from Douglas-fir and grand fir whose spectral values overlap its own.

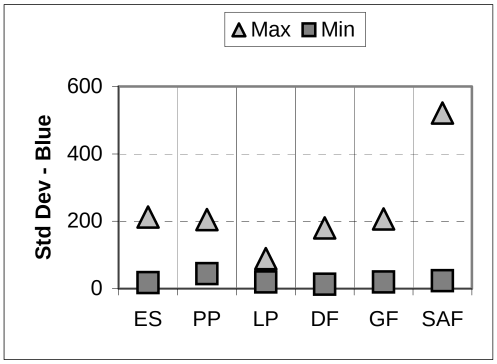

Figure 4.20 The range of standard deviation for eCognition ratio values within the blue spectral band in the combined multi-spectral sensor and LiDAR data associated with each of the six species found in the lower mountain regions study site near McCall, Idaho

Where, LP = Lodgepole pine, PP = Ponderosa pine, ES = Engelmannn spruce, DF = Douglas-fir, GF = Grand fir, SAF = Subalpine fir

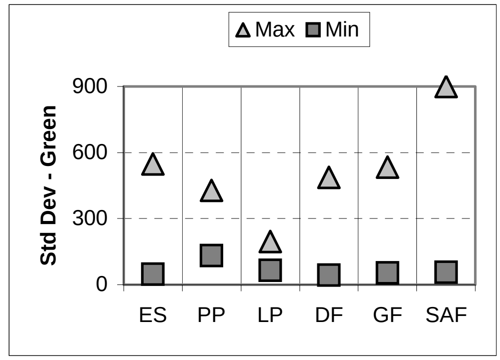

Figure 4.21 The range of standard Deviation for eCognition ratio values within the green spectral band in the combined multi-spectral sensor and LiDAR data associated with each of the six species found in the lower mountain regions study site near McCall, Idaho

Where, LP – Lodgepole, pine, PP = Ponderosa pine, ES = Engelmannn spruce, DF = Douglas-fir, GF = Grand fir, SAF = Subalpine fir

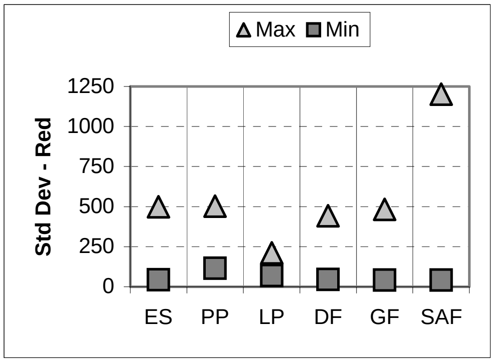

Figure 4.22 The range of standard deviation for eCognition ratio values within the red spectral band in the combined multi-spectral sensor and LiDAR data associated with each of the six species found in the lower mountain regions study site near McCall, Idaho

Where, LP = Lodgepole pine, PP= Ponderosa pine, ES = Engelmannn spruce, DF = Douglas-fir, GF = Grand fir, SAF = Subalpine fir

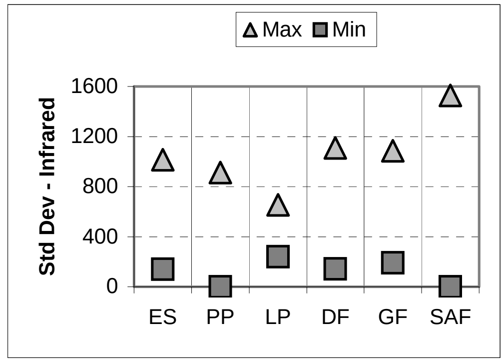

Figure 4.23 The range of standard deviation for eCognition ratio values within the infrared spectral band in combined multi-spectral sensor and LiDAR data associated with each of the six species found in the lower mountain regions study site near McCall, Idaho

Where, LP = Lodgepole pine, PP = Ponderosa pine, ES = Engelmannn spruce, DF = Douglas-fir, GF = Grand fir, SAF = Subalpine fir

Engelmann spruce, ponderosa pine, and lodgepole pine make up the “Other Species” group (Appendix Table B.5), along with Douglas-fir and grand fir whose spectral values also overlap subalpine fir. Engelmann spruce and ponderosa pine should be separable from each other due to their contrasting site requirements. However, lodgepole pine and Douglas-fir are found on sites for both Engelmann spruce and ponderosa pine (Burns and Honkala, 1990). No method of separating Engelmann spruce and ponderosa pine from lodgepole pine, Douglas-fir, and grand fir could be determined.

Despite the inclusion of LiDAR data, the intertwined spectral values of the species that made up the “Other Species” group were still not unique enough to be separable. Thus, nearest neighbor classification was added to the various site characteristics associated with each species. Samples were based upon vegetation index values and spectral band ratios. “Other Species” was subdivided into the “Possably LRatio DF”, the “Possably LRatio GF”, the “Possably LP”, the “Possably PP”, the “Possably ES”, and the “Unknown LRatio Tree” classes (Appendix Table B.7). The term “LRatio” refers to lower spectral values associated with these segments.

Ponderosa pine was more likely to be found in steeply sloped areas of lower elevations (Burns and Honkala, 1990). So trees found in higher elevations or in flatter regions were not classified as “Possibly PP”. In higher elevations, Engelmann spruce required the cooling shade of nearby trees to grow. It was confined to the cool bottoms and valleys in lower elevations. So, trees in cooler areas of the sites were classified as “Possibly ES”. Lodgepole pine had some of the most stable spectral values (Figures 4.20 to 4.23) as it is not shade tolerant. So trees that made up “Other Species” that had low standard deviations on their spectral band ratio values and were not too near areas of “Shadow” were grouped as “Possibly LP” (Appendix Table B.7).

Douglas-fir and grand fir both have spectral values that belong in the “Mostly SAF” and “Other Species” groups, and both have physical site characteristics that are shared by other species. Douglas-fir shares the same broad range of site conditions as lodgepole pine but tolerates light shade (Burns and Honkala, 1990). It was found on slopes of all aspects. Potential Douglas-fir segments that met the criteria for nearest neighbor and elevation were grouped as “Possibly HRatio DF” if its spectral characteristics were in the upper end of the spectrum or were grouped as “Possibly LRatio DF” if its spectral characteristics were in the lower end of the spectrum. Grand fir shares the same site preference as Engelmann spruce. It grows in moist areas of valleys and bottoms and along slopes and peaks of high elevations (Burns and Honkala, 1990). It is also highly shade tolerant. Potential grand fir segments with spectral characteristics in the upper end of the spectrum had to fulfill criteria for nearest neighbor, elevation, and focal rank to be grouped as “Possibly HRatio GF”. Potential grand fir segments with spectral characteristics in the lower end of the spectrum had to fulfill criteria for nearest neighbor, elevation, and focal rank to be grouped as “Possably LRatio GF”.

CHAPTER V

NEURAL NETWORK CLASSIFICATION

5.1 Methods

5.1.1 Approach

NeuralWare's NeuralWorks Professional II/Plus software was selected to construct a feed-forward back-propagation network. The back-propagation network was chosen because it is commonly used with biological data and classification problems (NeuralWare, 2000b). A multilayered neural network consists of an input layer of independent variable values, an output layer of predicted values, and one or more middle layers called hidden layers (Figure 5.1). Each layer is composed of a number of nodes that store and transfer calculated values to the nodes in the next layer. The input layer contains one node for each independent variable and the output layer contains one node for each class or species, as is the case in this model. Since there were six species to separate using MSS and LiDAR data, there were six nodes in the output layer.

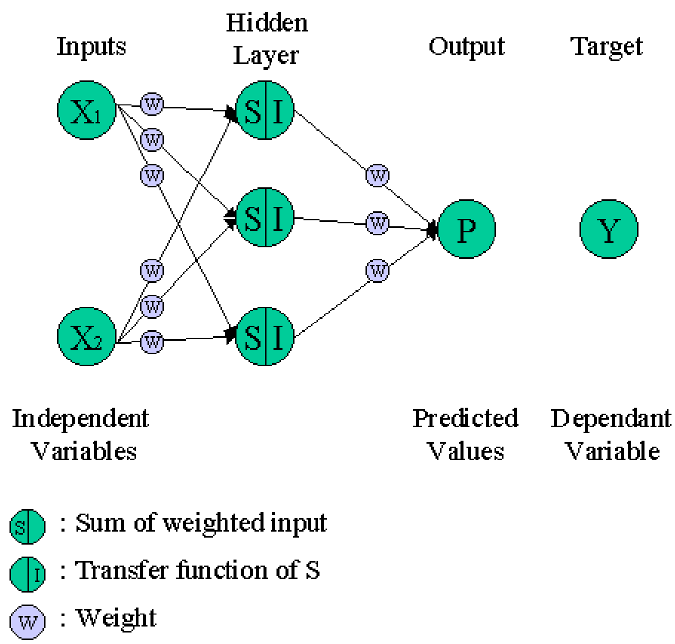

Figure 5.1 A simple multilayer fully connected, feed forward neural network (Meador, 2002)

The number of hidden layers and the number of nodes per hidden layer were determined by trial and error testing based on overall classification accuracies across all six species. In a fully-connected neural network, every node in a layer is connected or linked to every node in the adjacent layer. Input layer values for the independent variables travel in one direction, along the node-to-node connecting links, through the hidden layers to the output layer. Each link in the neural network is associated with an iteratively calculated weight, similar to a regression coefficient.

The construction of the network model takes place in a training stage followed by its evaluation in a testing phase. During the training phase, the input layer nodes receive scaled values of the independent variables. Each input node value is multiplied by its corresponding link weight and the product is transmitted to each connecting node in the next layer. Values for each node in the first hidden layer are determined by summing up all the values transmitted from the connecting nodes in the input layer and transforming them with a smoothing or transfer function. The weighting, summing, and transforming process continues from one hidden layer to the next until the output layer is reached. Output node values (predicted values) are descaled and compared to their corresponding observed dependent values. The differences between the observed and predicted values are used to adjust network weights by propagating errors back through the network according to the product of a learning rule, an error derivative, and the output of the previous layer. Network weights are adjusted for each observation or set of observations (epoch) presented to the network. The final product of the network is a model that is non-linear in its parameters (final network weights).

NeuralWorks' testing phase operates similarly to the training phase except there is no updating of weights. Randomly selected data subsets or independent datasets are used in the testing phase of the neural network. Input values are presented to the model that was built in the training stage, but no adjustments are made to weights. The differences between the predicted and observed values are used to calculate a collective error, usually RMSE, for comparing networks with different architectures and parameters and then selecting the best network (NeuralWare 2000b; Schultz et al. 1999, Weiss and Kulikowski, 1991).

5.1.2 Data

Object data output from eCognition was used as input to the neural network classification model. NeuralWorks requires input data in ASCII format, so information stored in the MSS and LiDAR imagery was exported from eCognition to an Excel spreadsheet. Each row of ASCII data represented a single segment created by the eCognition software from MSS and LiDAR imagery. Observations presented to the network were the data for each segment containing a plot tree or "Tree Crown Tips".

Training and testing datasets are required for building a neural network. Since there were only 338 "Tree Crown Tips" or observations in the MSS and MSS plus LiDAR datasets, testing was done on randomly presented sets (epochs) of observations from the training dataset. Therefore, the network's ability to generalize to an independent dataset of similar species and site type could not be evaluated.

5.2 Model Building

Separate neural network models were constructed for the MSS dataset and for the MSS plus LiDAR dataset. Network parameter values and the number of hidden layers and nodes per hidden layer were determined from recommendations by NeuralWorks, experienced modelers, and trial and error testing. Networks with one and two hidden layers were tested with a range of two to six nodes per hidden layer. There were a total of 338 "Tree Crown Tips" objects identified by eCognition whose feature data were available as input (training cases) to the neural network. There were 449 class, segment, and statistically derived variables derived from the MSS and

LiDAR data. To avoid overtraining the neural network and limiting its ability to generalize, the number of input variables and hidden layers was limited in proportion to the total number of training cases (NeuralWare, 2000c). A NeuralWorks input contribution instrument (bar chart) was used to determine which four to six variables of the 449 available inputs contributed most to the classification process. The dataset of all 449 variables was broken up into smaller datasets and used to train a neural network of set architecture and parameters. The variables which contributed most to the creation of the neural network were placed in a separate dataset and advanced to the next round of testing. Those variables that met a minimum percentage of contribution were selected. Four input variables were selected for the MSS network and two additional variables, one of which was LiDAR derived, were chosen for the MSS plus LiDAR network (Table 5.1).

Table 5.1 Variables selected for the MSS only and MSS plus LiDAR neural network models (Definiens Imaging, 2003)

MSS Only Data	MSS plus LiDAR Data
1) Brightness – sum of mean values of the layers containing brightness information divided by the quantity computed for an image object.	1) Brightness– sum of mean values of the layers containing brightness information divided by the quantity computed for an image object.
2) RatioBand2 (green) – mean layer value of image object divided by sum of all spectral layer mean values.	2) RatioBand2 (green) – mean layer value of image object divided by sum of all spectral layer mean values.
3) MeanDiffToNeighborsBand2 (green) – for each neighboring object the layer mean is computed and weighted to the length of the border between objects or the area covered by the neighbor objects.	3) MeanDiffToNeighborsBand2(green) – for each neighboring object the layer mean is computed and weighted to the length of the border between objects or the area covered by the neighbor objects.
4) StDevBand4 (infrared) – standard deviation calculated from the layer values of all pixels forming objects.	4) StDevBand4 (infrared) – standard deviation calculated from the layer values of all pixels forming objects.
	5) MeanDiffToSceneLiDARRelevObject – difference between the LiDAR elevation layer value of an image object and the LiDAR elevation layer mean value of the whole scene.
	6) Vegetative Index 1 – band 4 minus band 1 quantity divided by band 4 plus band 1 quantity.

Random observations from the MSS and MSS plus LiDAR datasets were presented to each neural network tested until a predetermined stopping criteria was met. After a number of networks were constructed with varying parameters and architectures, the MSS only and MSS plus LiDAR models with the best overall classification accuracies were selected. Network parameters for the two selected models are given in Tables 5.2 and 5.3. NeuralWorks outputs its models in C programming language code. Additional C code for the two selected models was written to calculate an error matrix and user, producer, and overall accuracies.

Table 5.2 NeuralWorks Professional II/Plus network parameters selected for the MSS only data model

Parameter	**Value**	**Description**
Network Type	Back-propagation	Back-propagation of errors
	Hetero-associative	Different input and output variables
	Min-max table	Scaled inputs
	Fully connected	All nodes connected in adjacent layers
Learning rule	Delta rule	Controls weights adjustments
Transfer function	Sigmoid	Smoothing function
Epoch size	1	No. inputs per weight update
No. hidden layers	1	
No. nodes/hidden layer	6	
Momentum	0.6	Modifies weights to prevent convergent behavior
Learning coefficients	HL1: 0.3	Multipliers in the calculation of weights, values
	Output: 0.15	Change after a set number of inputs (learn counts)

Table 5.3 NeuralWorks Professional II/Plus network parameters selected for the MSS plus LiDAR data model

Parameter	**Value**	**Description**
Network Type	Back-propagation	Back-propagation of errors
	Hetero-associative	Different input and output variables
	Min-max table	Scaled inputs
	Fully connected	All nodes connected in adjacent layers
Learning rule	Extended delta-bar-delta	Controls weights adjustments
Transfer function	Sigmoid	Smoothing function
Epoch size	1	No. inputs per weight update
No. hidden layers	1	
No. nodes/hidden layer	6	
Momentum	0.6	Modifies weights to prevent convergent behavior
Learning coefficients	HL1: 0.3	Multipliers in the calculation of weights, values
	Output: 0.15	Change after a set number of inputs (learn counts)

CHAPTER VI

RESULTS AND DISCUSSION

6.1 Analyses

Four separate classification analyses were conducted in an effort to separate the six tree species using remotely sensed data. Both MSS alone and MSS plus LiDAR datasets were used as inputs to eCognition's knowledge-base system and NeuralWorks' neural network modeler. The resulting four sets of classifications were compared to field data to determine: 1) if the addition of LiDAR data to the MSS data improved the classification of forest species in this study area and 2) which artificially intelligent modeling system gave the better results.

Classification results were reported for each of the four analyses by constructing error matrices (Congalton and Green, 1999) in which producer, user, and overall accuracies were recorded. The columns of the error matrices represent the samples assumed to be correctly identified and are called the reference data. The rows represent the classified observations generated from the remotely sensed data. Errors of both inclusion (commission) and exclusion (omission) can be observed in the matrices. A species classified within the incorrect category is a commission error. A species not classified within the correct category is an omission error. Producer's accuracy was calculated from the ratio of trees correctly classified out of all the reference trees for a species. User's accuracy was calculated from the ratio of trees correctly classified out of all the trees classified in a species category. Overall accuracy was calculated from the ratio between the total number of correctly identified trees for all species and the total number of trees used for the error matrix.

6.2 Knowledge-Based Classification

Subalpine fir was successfully identified as subalpine fir in both MSS (61%) and MSS plus LiDAR (71%) datasets; however, other species were often misidentified as subalpine fir (Tables 6.1 and 6.2). Both grand fir and Douglas-fir, species that share similar spectral characteristics with subalpine fir, were misclassified as subalpine fir as often, or more often, as they were correctly classified. Subalpine fir occupies the extremes of the spectral ratios and vegetation indices (Tables 4.9 to 4.12 and 4.17 to 4.19) and, not unexpectedly, had the second highest class accuracy in the MSS alone analysis and the highest in the MSS plus LiDAR analysis (Table 6.1).

Lodgepole pine's high producer accuracy (88%) in the MSS only dataset was probably due to the fact that much of its data had been gathered in pure stands. It was surprising, however, that when LiDAR data were added, the knowledge-based system was unable to accurately classify lodgepole pine. This species requires brightly-lit areas, so it could possibly be misclassified as shade-tolerant grand fir where it is located in closed-canopy stands. This explanation does not account for the misidentification of trees among the open, pure stands of lodgepole pine.

Table 6.1 Classification error matrix for the knowledge-based analysis on MSS only data from study sites near McCall, Idaho

User Species Class	**Reference Species Class**						
	ES	PP	LP	DF	GF	SAF	Sum
Engelmann spruce (ES)	17	10	2	7	3	1	40
Ponderosa Pine (PP)	24	24	3	11	4	2	68
Lodgepole Pine (LP)	2	1	66	5	1	1	76
Douglas-fir (DF)	8	11	2	9	9	5	44
Grand Fir (GF)	2	4	1	15	14	7	43
Subalpine Fir (SAF)	3	2	1	13	23	25	67
Sum	56	52	75	60	54	41	338

Accuracy

Producer	0.30	0.46	0.88	0.15	0.26	0.61
User	0.42	0.35	0.87	0.20	0.33	0.37
Overall	0.46					

Table 6.2 Classification error matrix for a knowledge-based analysis on MSS plus LiDAR data from study sites near McCall, Idaho

User Species Class	**Reference Species Class**						
	ES	PP	LP	DF	GF	SAF	Sum
Engelmann Spruce (ES)	2	1	3	1	2	1	10
Ponderosa Pine (PP)	30	33	4	11	11	3	92
Lodgepole Pine (LP)	1	0	1	1	1	1	5
Douglas-fir (DF)	3	5	5	8	4	3	28
Grand Fir (GF)	3	4	61	19	8	4	99
Subalpine Fir (SAF)	17	9	1	20	28	29	104
Sum	56	52	75	60	54	41	338
Accuracy							
Producer	0.04	0.63	0.01	0.13	0.15	0.71	
User	0.20	0.36	0.20	0.29	0.08	0.28	
Overall	0.24						

In a similar situation to lodgepole pine, Engelmann spruce should have benefited from the addition of LiDAR data but did not. Engelman spruce has similar spectral characteristics to ponderosa pine, and the addition of site variables was expected to differentiate between the two species, but accuracies were actually worse for the MSS plus LiDAR dataset. In areas of high slope, Engelmann spruce was often misclassified as ponderosa pine.

Ponderosa pine was the default classification for many of the species in the "Other Species" group (Figures 4.1.1 and 4.1.2). This misclassification could be explained from the fact that these species share similar spectral characteristics and are found in sloping areas. Ponderosa pine has an advantage over the other six species in remote classification, as it has one of the largest crowns, providing an easy to recognize characteristic. Ponderosa pine's producer accuracies recorded for the knowledge-based approach fell below those of subalpine pine in both analyses. In the MSS only analysis, they also fell below lodgepole pine but above the remaining species.

Both MSS only and MSS plus LiDAR analyses gave low producer accuracies (15% and 13%, respectively) for Douglas-fir. Douglas-fir was one of the most common species in the study, but shared spectral characteristics similar to other species. It was classified as ponderosa pine almost as often it was identified as grand fir and subalpine fir.

Grand fir and subalpine fir were expected to be difficult to separate because they are almost identical in physical appearance, differing only in total height and site elevation. Subalpine fir's spectral characteristics were slightly further on the extreme of the range and produced better producer and user classifications. Grand fir was identified as subalpine fir approximately 50% of the time in both analyses. It was classified as ponderosa pine in the combined MSS and LiDAR dataset as often as it was correctly identified.

Since separate identification of the six species was not readily obtained, error matrices were calculated for species group identification (Tables 6.3 and 6.4). It was assumed that MSS reflectance values would be influenced by needle morphology, and the six species were placed into two groups based on similar needle characteristics. Englemann spruce, ponderosa pine, and lodgepole pine formed one group while Douglas-fir, grand fir, and subalpine fir composed the another group. Classification accuracies were so greatly improved for both MSS alone and MSS plus LiDAR analyses that results suggested that needle characteristics play a major role in the ability to separate the species based on the imagery data. Accuracy rates between 77% and 81% for the MSS alone analysis (Table 6.3) could benefit applications where the quantity or distribution of these species was important.

Table 6.3 Classification error matrix for the knowledge-based species group analysis on MSS only data from study sites near McCall, Idaho

User Species Class	**Reference Species Class**		
	Spruce-Pine Group ES, PP, LP	**Fir Group** DF, GF, SAF	Sum
Spruce-Pine Group			
Engelmann spruce (ES), Ponderosa Pine (PP), Lodgepole Pine (LP)	149	35	184
Fir Group			
Douglas-fir (DF), Grand Fir (GF), Subalpine Fir (SAF)	34	120	154
Sum	183	155	338
Accuracy			
Producer	0.81	0.77	
User	0.81	0.78	
Overall	0.80		

Table 6.4 Classification error matrix for the knowledge-based species group analysis on MSS plus LiDAR data from study sites near McCall, Idaho

User Species Class	**Reference Species Class**		
	Spruce-Pine Group ES, PP, LP	**Fir Group** DF, GF, SAF	Sum
Spruce-Pine Group			
Engelmann spruce (ES), Ponderosa Pine (PP), Lodgepole Pine (LP)	75	32	107
Fir Group			
Douglas-fir (DF), Grand Fir (GF), Subalpine Fir (SAF)	108	123	231
Sum	183	155	338
Accuracy			
Producer	0.41	0.79	
User	0.70	0.53	
Overall	0.59		

6.3 Neural Network Classification

Neural networks with varying architectures and parameters values were generated for both the MSS alone and MSS plus LiDAR datasets until the overall accuracy values of the error matrices could no longer be improved. The neural network approach produced several classification schemes that were similar in producer, user, and overall accuracy values. The models were limited to one hidden layer and a maximum of six hidden layer nodes so as to avoid over training. The models with the highest overall accuracies for the MSS alone and MSS plus LiDAR datasets are presented in Tables 6.5 and 6.6, respectively. The architecture and parameter values for the selected models are given in Tables 5.2 and 5.3. The primary differences between the two models were the learning algorithm and the number of input variables.

Table 6.5 Classification error matrix for the neural network individual species analysis based on MSS only data from study sites near McCall, Idaho

User Species Class	**Reference Species Class**						
	ES	PP	LP	DF	GF	SAF	Sum
Engelmann spruce (ES)	33	8	9	1	3	2	56
Ponderosa Pine (PP)	8	41	2	4	5	2	62
Lodgepole Pine (LP)	1	2	64	6	0	0	73
Douglas-fir (DF)	7	11	13	26	0	1	58
Grand Fir (GF)	11	11	1	8	13	8	52
Subalpine Fir (SAF)	6	2	2	10	4	17	41
Sum	66	75	91	55	25	30	342
Accuracy							
Producer	0.59	0.66	0.88	0.45	0.25	0.41	
User	0.50	0.55	0.70	0.47	0.52	0.57	
Overall	0.57						

Table 6.6 Classification error matrix for the neural network individual species analysis based on MSS plus LiDAR data from study sites near McCall, Idaho

User Species Class	**Reference Species Class**						
	ES	PP	LP	DF	GF	SAF	Sum
Engelmann Spruce (ES)	35	9	1	6	1	3	55
Ponderosa Pine (PP)	4	43	1	7	5	1	61
Lodgepole Pine (LP)	2	2	64	6	0	0	74
Douglas-fir (DF)	6	5	7	38	2	0	58
Grand Fir (GF)	8	10	0	15	15	4	52
Subalpine Fir (SAF)	5	2	2	11	2	19	41
Sum	60	71	75	83	25	27	341
Accuracy							
Producer	0.63	0.71	0.86	0.66	0.29	0.46	
User	0.58	0.61	0.85	0.46	0.60	0.70	
Overall	0.63						

The MSS plus LiDAR model included all the MSS only model variables plus two additional variables and produced an overall classification accuracy of 63% compared to 57% for the MSS only model. Of the two additional MSS plus LiDAR variables only one was LiDAR derived: the difference between the LiDAR elevation layer value of a tree object ("Trees Crown Tips") and the LiDAR elevation layer mean value of the whole scene. The other additional variable was a derived vegetation index (Table 4.2.1).

Lodgepole pine was the most reliably identified species at 88% and 86% producer accuracy in the MSS only and MSS plus LiDAR analyses, respectively. Lodgepole pine's high occurrence in pure stands and low standard deviation of spectral characteristics may have distinguished it sufficiently to accurately classify.

Subalpine fir, whose spectral characteristics consistently occupied one extreme, was classified in the MSS only and MSS plus LiDAR analyses with 41% and 46% producer accuracies, respectively. Grand fir could not be reliably separated from the other five species and was misclassified as all the other species except lodgepole pine in approximately similar numbers.

MSS only and MSS plus LiDAR classifications for Douglas-fir (45% and 66%, respectively) and ponderosa pine (66% and 71%, respectively) were intermediate in accuracy to the other species. Douglas-fir was most often misclassified as lodgepole pine. As was the case in the knowledge-based approach, ponderosa pine and Engelmann spruce were often misclassified as each other, probably because their spectral characteristics are very similar.

Error matrices were also constructed by the same species groups as in the knowledge-base analysis, Englemann spruce, ponderosa pine, and lodgepole pine in one group and Douglas-fir, grand fir, and subalpine fir in the other group. Overall accuracies were improved over individual species identification as was the case for the knowledge-based analysis (Tables 6.3 and 6.4). The overall accuracy rates for the two species groups were very similar for the MSS alone (75%) and MSS plus LiDAR (78%) datasets (Tables 6.7 and 6.8). The addition of the LiDAR variable may not have played an important role in helping to distinguish between the two groups. As noted with the knowledge-based analysis, the multi-spectral variables may be more important in differentiating between the spectral characteristics associated with the species grouping by needle morphology.

Table 6.7 Classification error matrix for the neural network species group analysis on MSS only data from study sites near McCall, Idaho

User Species Class	**Reference Species Class**		
	Spruce-Pine Group ES, PP, LP	**Fir Group** DF, GF, SAF	Sum
Spruce-Pine Group			
Engelmann spruce (ES), Ponderosa Pine (PP), Lodgepole Pine (LP)	168	23	191
Fir Group			
Douglas-fir (DF), Grand Fir (GF), Subalpine Fir (SAF)	64	87	151
Sum	232	110	342
Accuracy			
Producer	0.72	0.79	
User	0.88	0.58	
Overall	0.75		

Table 6.8 Classification error matrix for the neural network species group analysis on MSS plus LiDAR data from study sites near McCall, Idaho

User Species Class	**Reference Species Class**		
	Spruce-Pine Group ES, PP, LP	**Fir Group** DF, GF, SAF	Sum
Spruce-Pine Group			
Engelmann spruce (ES), Ponderosa Pine (PP), Lodgepole Pine (LP)	161	29	190
Fir Group			
Douglas-fir (DF), Grand Fir (GF), Subalpine Fir (SAF)	45	106	151
Sum	206	135	341
Accuracy			
Producer	0.78	0.79	
User	0.85	0.70	
Overall	0.78		

6.4 Comparison of Knowledge-Base and Neural Network Approaches

The knowledge-based classification system succeeded in identifying two of the six species (lodgepole pine and subalpine pine) in the MSS only dataset with producer accuracies (88% and 61%, respectively) great enough to recommend the model to applications specific to these species. The addition of LiDAR derived site variables to the knowledge-based system drastically reduced the classification of lodgepole pine from 88% to 1% and greatly decreased overall accuracy. However, the producer accuracies for ponderosa pine and subalpine pine were increased to 63% and 71%, respectively. The neural network system showed less species sensitivity to differences between the MSS only and MSS plus LiDAR datasets than the knowledge-based system. Producer accuracy differences for any one species between the two neural network datasets ranged from 2% to 5%, with the exception of Douglas-fir at 21%. The neural network's overall classification accuracy of 63% would suggest its use in stratifying sampling schemes, estimating wildlife habitat, or allocating inventories or other characteristics to large forested areas where intensive sampling is not feasible.

There were several similarities between the two modeling systems. Lodgepole pine showed a high success rate in the MSS only knowledge-based classification and in both neural network dataset analyses. Grand fir was not predicted reliably with either modeling system. It was primarily misclassified as subalpine fir in the knowledge-based system and four of the other species in the neural network. Subalpine fir, whose spectral characteristics were at one extreme, was identified with greater accuracy in the knowledge-based system (61% to 80% producer accuracies).

In both classification schemes, all or most of the input variables were based on spectral characteristics. However, unlike the knowledge-based classification scheme, the addition of LiDAR data to the MSS data in the neural network increased overall accuracy slightly as opposed to decreasing it.

CHAPTER VII

CONCLUSIONS

Knowledge-based and learning system (neural network) approaches were employed to model the classification of six open-grown coniferous species near McCall, Idaho, using MSS data alone and MSS plus LiDAR data. The learning system produced better individual species identification accuracies, and the knowledge base produced better species group accuracies. The user-created classification schemes of the two individual species knowledge-based analyses could identify two species out of six with producer accuracies greater than 50%. The neural network classification schemes identified all but grand fir with accuracies ranging from 41% to 88%. The addition of LiDAR data to the MSS dataset decreased the overall success of the knowledge-based classification scheme by almost half (from 46% to 24%) while the learning system classification scheme was slightly improved (57% to 63%, respectively).

It was expected that knowledge-based classification would improved with the addition of site variables provided by the LiDAR data. However, the addition of more imagery in eCognition changed the values of the spectral band ratios associated with each species compared to the MSS alone analysis. Membership function definitions changed and the MSS plus LiDAR classification scheme did not perform as well as the MSS alone analysis.

Using the knowledge-based classification scheme, only lodgepole pine (88% producer accuracy) and subalpine fir (61% producer accuracy) could be reliably classified using MSS data alone. The spectrally similar species, Douglas-fir and grand fir, were also identified as subalpine fir (37% user accuracy). With the addition of LiDAR-derived data, accuracies improved for subalpine fir and ponderosa pine. Similar to the case with subalpine fir, species that shared spectral and site characteristics with ponderosa pine were misclassified as ponderosa pine (35% user accuracy). The tendency for the user-created classification scheme to identify trees as subalpine fir and ponderosa pine affected the accuracies of all six species. When the species were separated into two groups according to needle morphology (Douglar-fir, grand fir, and subalpine fir in one group and ponderosa pine, lodgepole pine, and Engelmann spruce in another group) overall group classification accuracy was 81%.

With the learning-system (neural network) classification scheme, ponderosa pine, lodgepole pine, and Engelmann spruce can be reliably identified with 66%, 88%, and 59% producer accuracies, respectively, using MSS data alone. Douglas-fir classification was less accurate at 45%. Grand fir could not be successfully separated from the other species. The addition of LiDAR data to the MSS neural network dataset improved overall accuracy making relatively small changes in individual species accuracies with the exception of Douglas-fir. Douglas-fir accuracy showed a large improvement going from 45% (MSS only) to 66% (MSS plus LiDAR).

The decrease in the accuracy of the knowledge-based system with the addition of LiDAR data, highlights the underlying differences between the knowledge-based and neural network approaches. A weakness of the knowledge-based approach is its dependence on the modeler to navigate, utilize, and distinguish among hundreds of derived variables. The nearest neighbor approach within the knowledge-base did provide some automation that produced improvements. It was logical to assume that the addition of LiDAR site variables would increase the distinction among species, but using hundreds of variable features with an almost infinite number of classification decisions may not be a task well-suited to manual cognitive procedures. The neural network, on the other hand, relied on the advantages of computer processing speed to automatically and exhaustively examine and test thousands of models with hundreds of object and class-related features and select the best one. The knowledge-based approach worked well (80%) for species group classification where needle morphology was identified as an important distinguishing feature in spectral reflectance. The neural network performed almost as well in group classification at 78%.

The results of this study indicate that four of the coniferous species in the mountainous regions of central Idaho can be classified with the assistance of neural network modeling and MSS plus LiDAR data with accuracies between 63% and 86%. Species group classifications can be reliably made using knowledge-based modeling and MSS data alone or neural network modeling. Applications in wildlife habitat management, inventory, forest damage assessment, and other problems that require estimates of individual species or species groups could be greatly facilitated by this type of modeling. Cost reductions compared to field data collection, especially in mountainous regions, and the ability to evaluate large areas are both advantages of classifying tree species with remotely sensed data.

BIBLIOGRAPHY

Burns, Russell M., and Barbara H. Honkala, tech. coords. 1990. Silvics of North America: 1. Conifers; 2. Hardwoods. Agriculture Handbook 654. U.S. Department of Agriculture, Forest Service, Washington, DC. vol.2, 877 p. (Online)(USDA Forest Service, St. Paul Field Office, assessed on December 3, 2003) Available from http://www.na.fs.fed.us/spfo/pubs/silvics_manual/table_of_contents.htm.

Casey, W. A. 1999. Forest species group classification using digital frame camera imagery. [Thesis] Mississippi State University, Mississippi State, MS. 54p.

Collins, C. A. 2003. Comparing integrated LiDAR and multispectral data with filed measurements in hardwood stands. [Thesis] Mississippi State University, Mississippi State, MS. 158p.

Congalton, R. G. and K. Green. 1999. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. Lewis Publishers. Boca Raton. 137p.

Conradsen, K. and J. Gunulf. 1986. A geological example of improving classification of remotely sensed data using additional variables and a hierarchical structure. Photogrammetric Engineering and Remote Sensing, 52 (8):1181-1187.

Definiens Imaging. 2003. User guide 3. Definiens Imaging, Munich, Germany. 475p.

ERDAS, Inc. 1997. ERDAS Field Guide. ERDAS, Inc., Atlanta, GA. 656p.

ESRI, Inc. 1998. Getting to Know ArcView GIS. ESRI Press, Redlands, CA. 654p.

Douglas. T. E. 2004. Classification of pine and hardwood by density and intensity of LIDAR canopy returns. [Thesis], Mississippi State University, Department of Forestry. 43p.

Franklin, S. E., A. J. Maudie, and M. B. Lavigne. 2001. Using spatial co-occurrence texture to increase forest structure and species composition classification accuracy. Photogrammetric Engineering and Remote Sensing, 67 (7):849-855.

Gimblett, R. H. and G. L. Ball. 1995. Neural network architectures of monitoring and simulating changes in forest resource management. AI Applications, 9(2). pp. 103-123.

Gopal, S. and C. E. Woodcock. 1996. Remote sensing of forest change using artificial neural networks. IEEE Transactions on Geoscience and Remote Sensing, 34(2): 398-404.

Hughes, J. S., D. L. Evans, P. Y. Burns, and J. M. Hill. 1986. Identification of two southern pine species in high-resolution aerial MSS data. Photogrammetric Engineering and Remote Sensing, 52 (8):1175-1180.

Knight, T. C. A. 2003. Analysis of competing hardwoods in mid-rotation loblolly pine plantations using remote sensing technology. [Thesis] Mississippi State University, Mississippi State, MS. 83p.

Jensen, J. R. 2005. Introductory Digital Image Processing: A Remote Sensing Perspective. Upper Saddle River, NJ: Pearson Prentice Hall, Inc. 526p.

Kraus, K. and N. Pfeifer. 1998. Determination of terrain models in wooded areas with airborne laser scanner data. Journal of Photogrammetry and Remote Sensing, 53: 193-203.

Lefsky, M. A., W.B. Cohen, S.A. Acker, C.G. Parker, T. A. Spies, and D. Harding. 1999. LiDAR remote sensing of the canopy structure and biophysical properties of Douglas-fir Western Hemlock forests. Remote Sensing of Environment. 70: 339-361.

Lillesand, T. M. and Kiefer, R. W. 2000. Remote sensing and image interpretation. New York: John Wiley & Sons, Inc. 724p.

Meador, A.J. 2002. Applications in growth and yield of longleaf pine (*Pinus palustris* Mill.): A comparison of artificial neural networks and other traditional modeling methodologies. [Thesis], Mississippi State University, Department of Forestry. 49p.

Means, J.E., S. A. Acker, D. J. Harding, J. B. Blair, M. A. Lefsky, W. B. Cohen, M. E. Harmon, and W. A. McKee. 1999. Use of large-footprint scanning airborne LiDAR to estimate forest stand characteristics in the Western Cascades of Oregon. Remote Sensing of Environment, 67:298-308.

Mehrotra, K., C. K. Mohan, and S. Ranka. 1997. Elements of artificial neural networks. Cambridge: MIT Press. p. 1-42.

Moshou, D., E. Vrindts, B. De Ketelaere, J. De Baerdemaeker, and H. Ramon. 2001. A neural network based plant classifier. Computers and Electronics in Agriculture, 31:5-16.

Nelson, R., W. Krabill, and J. Tonelli. 1988. Estimating forest biomass and volume using airborne laser data. Remote Sensing of Environment 24:247-267.

NeuralWare, Inc. 2000a. Reference guide. NeuralWare, Inc., Pittsburgh, PA. 286 pp.

NeuralWare, Inc. 2000b. Neural computing. NeuralWare, Inc., Pittsburgh, PA. 324 pp.

NeuralWare, Inc. 2000c. Using NeuralWorks. NeuralWare, Inc., Pittsburgh, PA. 157 pp.

Nilsson, M. 1996. Estimation of tree heights and stand volume using an airborne LiDAR system. Remote Sensing of Environment. 56:1-7.

Principe, J.C., N. R. Euliano, and W. C. Lefebvre. 2000. Neural and adaptive systems: fundamentals through simulations. New York: John Wiley & Sons, Inc. 656 p.

Rohde, W. G. and L.E. Olson, Jr. 1972. Multi-spectral sensing of forest tree species. Photogrammetric Engineering and Remote Sensing, 38:1209-1215.

Ritchie, J. C., D. L. Evans, D. Jacobs, J. H. Everitt, and M. A. Weltz. 1993. Measing canopy structure with an airborne laser altimeter. Transactions of the ASAE 36 (4):1235-1238.

Schultz, E. B., T. G. Matney, and J. L. Koger. 1999. A neural network model for wood chip thickness distributions. Wood and Fiber Science, 31(1):2-14.

Schultz, E. B., T. G. Matney, and W.F. Watson. 1997. Prediction of moisture contents for sprinkled and un-sprinkled stacked roundwood over time. 1997 Pulping Conference Book 1. October 19-23, Sna Francisco, CA. TAPPI Press. pp. 289-299.

Skirvin, S. M. and G. Dryden. 1992. Classification of Landsat Thematic Mapper image data, Chiricahua National Monument, Arizona. AI Applications, 11(2):99-106.

Stefik, M. 1995. Introduction to knowledge systems. San Francisco, CA: Morgan Kaufmann Publishers, Inc. 870 p.

Van Aardt, J. A. N. and R. H. Wynne. 2001. Spectral separability among six Southern tree species. *Photogrammetric Engineering and Remote Sensing*, 67(12):1367-1375.

Weiss, S.M and C.A. Kulikowski. 1991. Computer systems that learn: classification and prediction methods from statistics, neural nets, machine learning, and expert systems. San Francisco, CA: Morgan Kaufmann Publishers, Inc. 223 p.

Zhang, M., L.O. Hall, D.B. Goldgof, F.E. Müller-Karger. 2000. Knowledge-Guided Classification of Coastal Zone Color Images Off the West Florida Shelf. International Journal of Pattern Recognition & Artificial Intelligence, 14(8):987-1008.

Zimble, D. A. 2002. Characterizing vertical forest structure using small-footprint airborne LiDAR. [Thesis] Mississippi State University, Mississippi State, MS.

APPENDIX A

KNOWLEDGE-BASED CLASSIFICATION SCHEME FOR MSS DATASET

A.1 Membership function definitions for the “Non Vegetation” and “Vegetation” classes in the knowledge-based classification scheme utilizing MSS data

Class Name	**Classification Scheme**			
Non Vegetation	1 or 1 requirement			
or	Ratio Band 4 : Smaller Than			
	Range:	0.225 - 0.275	Center:	0.250
or	NDVI [(Infrared-Red)/(Infrared+Red)] : Smaller Than (Boolean) *			
	Range:	-0.1 - 0.5	Center:	0.2
Vegetation	1 requirement			
	not			
	"Non Vegetation" class			

* NDVI = Normalized Difference Vegetation Index

A.2 Membership function definitions for the “Shadow,” “Trees,” and “Understory” classes in the knowledge-based classification scheme utilizing MSS data

Class Name : Vegetation	**Classification Scheme**			
Shadow	1 requirement			
	Brightness : Smaller Than			
	Range:	600 - 1200	Center:	900
Trees	3 requirements			
and	Ratio Band 4 : Larger Than			
	Range:	0.35 - 0.45	Center:	0.4
and	Border length : Smaller Than (Boolean)			
	Range:	0 - 150	Center:	75
and	not			
	"Shadow" class			
Understory	2 requirements			
	not			
	"Shadow" class, "Trees" class			

A.3 Membership function definitions for the “Crown Tips” and “Tree Crown” classes in the knowledge-based classification scheme utilizing MSS data

Class Name : Trees	**Classification Scheme**			
Crown Tips	3 requirements			
and	Border length : Smaller Than (Boolean)			
	Range:	20 - 50	Center:	35
and	Length/width : Smaller Than			
	Range:	3 - 5	Center:	4
and	Shape Index : Smaller Than			
	Range:	1.9 - 2.1	Center:	2.0
Tree Crown	1 requirement			
	not			
	"Crown Tips" class			

A.4 Membership function definitions for the “Mostly SAF,” “Other Species,” and “Unknown Crown Tips” classes in the knowledge-based classification scheme utilizing MSS data

Class Name : Crown Tips	**Classification Scheme**			
Mostly SAF	4 requirements			
and	Ratio Band 1 : Smaller Than (Boolean)			
	Range:	0.071 - 0.091	Center:	0.081
and	Ratio Band 2 : Smaller Than (Boolean)			
	Range:	0.200 - 0.210	Center:	0.205
and	Ratio Band 3 : Smaller Than (Boolean)			
	Range:	0.140 - 0.150	Center:	0.145
and	Ratio Band 4 : Larger Than (Boolean)			
	Range:	0.540 - 0.550	Center:	0.545
Other Treetops	4 requirements			
and	Ratio Band 1 : Larger Than (Boolean)			
	Range:	0.074 - 0.084	Center:	0.079
and	Ratio Band 2 : Larger Than (Boolean)			
	Range:	0.190 - 0.200	Center:	0.195
and	Ratio Band 3 : Larger Than (Boolean)			
	Range:	0.140 - 0.150	Center:	0.145
and	Ratio Band 4 : Smaller Than (Boolean)			
	Range:	0.56 - 0.58	Center:	0.57
Unknown Crown Tips	2 requirements			
	not			
	"Mostly SAF" class, "Other Treetops" class			

A.5 Membership function definitions for the “SAF” and “GF” classes in the knowledge-based classification scheme utilizing MSS data

Class Name : Mostly SAF	**Classification Scheme**
SAF	2 requirements
and	NN * : Ratio Bands (B, G, R, IR)
	TTA Mask
and	NN * : NDVI [(Infrared-Red)/(Infrared+Red)] **
	TTA Mask
GF	2 requirements
and	NN * : Ratio Bands (B, G, R, IR)
	TTA Mask
and	NN * : NDVI [(Infrared-Red)/(Infrared+Red)] **
	TTA Mask

* NN = Nearest Neighbor

** NDVI = Normalized Difference Vegetation Index

A.6 Membership function definitions for the "DF," "LP," "PP," and "ES" classes in the knowledge-based classification scheme utilizing MSS data

Class Name : Other Treetops	Classification Scheme
DF	2 requirements
and	NN * : Ratio Bands (B, G, R, IR) TTA Mask
and	NN * : NDVI [(Infrared-Red)/(Infrared+Red)] ** TTA Mask
LP	2 requirements
and	NN * : Ratio Bands (B, G, R, IR) TTA Mask
and	NN * : NDVI [(Infrared-Red)/(Infrared+Red)] ** TTA Mask
PP	2 requirements
and	NN * : Ratio Bands (B, G, R, IR) TTA Mask
and	NN * : NDVI [(Infrared-Red)/(Infrared+Red)] ** TTA Mask
ES	2 requirements
and	NN * : Ratio Bands (B, G, R, IR) TTA Mask
and	NN * : NDVI [(Infrared-Red)/(Infrared+Red)] ** TTA Mask

* NN = Nearest Neighbor
** NDVI = Normalized Difference Vegetation Index

A.7 Membership function definitions for the “Possibly SAF,” “Possibly GF,” “Possibly DF,” “Possibly LP,” “Possibly PP,” and “Possibly ES” classes in the knowledge-based classification scheme utilizing MSS data

Class Name : Unknown Crown Tip	**Classification Scheme**
Possibly SAF	1 requirement
	Similarity To "SAF" class
Possibly GF	1 requirement
	Similarity To "GF" class
Possibly DF	1 requirement
	Similarity To "DF" class
Possibly LP	1 requirement
	Similarity To "LP" class
Possibly PP	1 requirement
	Similarity To "PP" class
Possibly ES	1 requirement
	Similarity To "ES" class

APPENDIX B

KNOWLEDGE-BASED CLASSIFICATION SCHEME FOR MSS PLUS LIDAR DATASET

B.1 Membership function definitions for the “Non Vegetation” and “Vegetation” classes in the knowledge-based classification scheme utilizing MSS and LiDAR data

Class Name	**Classification Scheme**			
Non Vegetation	1 requirement			
	NDVI [(Infrared-Red)/(Infrared+Red)] : Smaller Than (Boolean) *			
	Range:	0.00 - 0.30	Center:	0.15
Vegetation	1 requirement			
	Not Similar to			
	"Non Vegetation" class			

* NDVI = Normalized Difference Vegetation Index

B.2 Membership function definitions for the “Shadow,” “Trees,” and “Understory” classes in the knowledge-based classification scheme utilizing MSS and LiDAR data in the southwest study site

Class Name : Vegetation	**Classification Scheme**			
Shadow	1 requirement			
	Brightness : Smaller Than			
	Range:	425 - 475	Center:	450
Trees	3 requirements			
and	NDVI [(Infrared-Red)/(Infrared+Red)] : Larger Than (Boolean) *			
	Range:	0.2 - 0.4	Center:	0.3
and	Border length : Smaller Than (Boolean)			
	Range:	0 - 150	Center:	75
and	Not Similar to			
	"Shadow" class			
Understory	2 requirements			
	Not Similar to			
	"Shadow" class, "Trees" class			

* NDVI = Normalized Difference Vegetation Index

B.3 Membership function definitions for the “Shadow,” “Trees,” and “Understory” classes in the knowledge-based classification scheme utilizing MSS and LiDAR data in the northeast study site

Class Name : Vegetation	**Classification Scheme**			
Shadow	1 requirement			
	Brightness : Smaller Than			
	Range:	600 - 650	Center:	625
Trees	3 requirements			
and	NDVI [(Infrared-Red)/(Infrared+Red)] : Larger Than (Boolean) *			
	Range:	0.2 - 0.4	Center:	0.3
and	Border length : Smaller Than (Boolean)			
	Range:	0 - 150	Center:	75
and	Not Similar to			
	"Shadow" class			
Understory	2 requirements			
	Not Similar to			
	"Shadow" class, "Trees" class			

* NDVI = Normalized Difference Vegetation Index

B.4 Membership function definitions for the “Crown Tips” and “Tree Crown” classes in the knowledge-based classification scheme utilizing MSS and LiDAR data

Class Name : Trees	**Classification Scheme**			
Crown Tips	4 requirements			
and	Border length : Smaller Than (Boolean)			
	Range:	20 - 50	Center:	35
and	Brightness : Larger Than (Boolean)			
	Range:	600 - 650	Center:	625
and	Length/width : Smaller Than			
	Range:	4.0 - 5.0	Center:	4.5
and	Shape Index : Smaller Than			
	Range:	1.9 - 2.1	Center:	2
Tree Crown	1 requirement			
	Not Similar to			
	"Crown Tips" class			

B.5 Membership function definitions for the “Mostly SAF,” “Other Species,” and “Unknown Tree” classes in the knowledge-based classification scheme utilizing MSS and LiDAR data

Class Name : Crown Tips	**Classification Scheme**			
Mostly SAF	3 requirements			
and	NDVI [(Infrared-Blue)/(Infrared+Blue)] : Larger Than *			
	Range:	0.65 - 0.70	Center:	0.675
and	NDVI [(Infrared-Green)/(Infrared+Green)] : Larger Than *			
	Range:	0.35 - 0.45	Center:	0.4
and	NDVI [(Infrared-Red)/(Infrared+Red)] : Larger Than *			
	Range:	0.45 - 0.55	Center:	0.5
Other Species	3 requirements			
and	NDVI [(Infrared-Blue)/(Infrared+Blue)] : Smaller Than *			
	Range:	0.70 - 0.80	Center:	0.75
and	NDVI [(Infrared-Green)/(Infrared+Green)] : Smaller Than *			
	Range:	0.45 - 0.55	Center:	0.5
and	NDVI [(Infrared-Red)/(Infrared+Red)] : Smaller Than *			
	Range:	0.55 - 0.65	Center:	0.6
Unknown Tree	2 requirements			
	Not Similar to			
	"Mostly SAF" class, "Other Treetops" class			

* NDVI = Normalized Difference Vegetation Index

B.6 Membership function definitions for the “Possibly SAF,” “Possibly HRatio GF,” “Possibly HRatio DF,” and “Unknown HRatio Tree” classes in the knowledge-based classification scheme utilizing MSS and LiDAR data

Class Name : Mostly SAF			**Classification Scheme**				
Possably SAF			2 or 1 requirement				
	or	and	LiDAR Elevation : Larger Than				
			Range:	1500 - 1525	Center:	1512.5	
		and	NN * : Ratio Bands (B, G, R, IR) and NDVI (B, G, R) **				
			TTA Mask				
	or	or	StDev Band 1 (Blue) : Larger Than (Boolean)				
			Range:	0 - 400	Center:	200	
		or	StDev Band 2 (Green) : Larger Than (Boolean)				
			Range:	0 - 1000	Center:	500	
		or	StDev Band 3 (Red) : Larger Than (Boolean)				
			Range:	0 - 1000	Center:	500	
		or	StDev Band 4 (Infrared) : Larger Than (Boolean)				
			Range:	0 - 2200	Center:	1100	
Possably HRatio GF			2 or 3 requirements				
and			NN * : Ratio Bands (B, G, R, IR) and NDVI (B, G, R) **				
			TTA Mask				
and	or		LiDAR Elevation - Full Range				
			Range:	1220 - 1675	Center:	1447.5	
	or	and	LiDAR Elevation : Smaller Than				
			Range:	1525 - 1550	Center:	1537.5	
		and	Focal Rank : Smaller Than				
			Range:	0 - 0.50	Center:	0.25	
Possably HRatio DF			2 requirements				
		and	NN * : Ratio Bands (B, G, R, IR) and NDVI (B, G, R) **				
			TTA Mask				
		and	LiDAR Elevation - Approximate Gaussian				
			Range:	550 - 2500	Center:	1525	
Unk. HRatio Tree			3 requirements				
			Not Similar To				
			"Poss. HRatio DF" class, "Poss. HRatio GF" class, "Poss. SAF" class				

* NN = Nearest Neighbor
** NDVI = Normalized Difference Vegetation Index

B.7 Membership function definitions for the “Possibly LRatio DF,” “Possibly LRatio GF,” “Possibly LP,” “Possibly PP,” “Possibly ES,” and “Unknown LRatio Tree” classes in the knowledge-based classification scheme utilizing MSS and LiDAR data

Class Name : Other Species			Classification Scheme			
Possably LRatio DF			2 requirements			
and			NN * : Ratio Bands (B, G, R, IR) and NDVI (B, G, R) **			
			TTA Mask			
and			LiDAR Elevation - Approximate Gaussian			
			Range:	550 - 2500	Center:	1525
Possably LRatio GF			2 or 3 requirements			
and			NN * : Ratio Bands (B, G, R, IR) and NDVI (B, G, R) **			
			TTA Mask			
and	or		LiDAR Elevation - Full Range			
			Range:	1220 - 1675	Center:	1447.5
	or	and	LiDAR Elevation : Smaller Than			
			Range:	1525 - 1550	Center:	1537.5
		and	Focal Rank : Smaller Than			
			Range:	0 - 0.50	Center:	0.25
Possably LP			3 or 3 or 3 requirements			
and			NN * : Ratio Bands (B, G, R, IR) and NDVI (B, G, R) **			
			TTA Mask			
and			Inverse Distance to Shadow Neighbor-objects : Smaller Than			
			Range:	0 - 15.0	Center:	7.5
and	or		StDev Band 1 (Blue) : Smaller Than (Boolean)			
			Range:	0 - 200	Center:	100
	or		StDev Band 2 (Green) : Smaller Than (Boolean)			
			Range:	0 - 400	Center:	200
	or		StDev Band 3 (Red) : Smaller Than (Boolean)			
			Range:	0 - 400	Center:	200
	or		StDev Band 4 (Infrared) : Smaller Than (Boolean)			
			Range:	0 - 1100	Center:	650
Possably PP						
and			NN * : Ratio Bands (B, G, R, IR) and NDVI (B, G, R) **			
			TTA Mask			
and			Inverse LiDAR Elevation : Larger Than			
			Range:	1900 - 1950	Center:	1925
and			Inverse Degree Slope : Smaller Than			
			Range:	0 - 30	Center:	15
Possably ES			3 or 3 requirements			
and			NN * : Ratio Bands (B, G, R, IR) and NDVI (B, G, R) **			
			TTA Mask			
and	or	and	LiDAR Elevation : Larger Than			
			Range:	1775 - 1825	Center:	1800
		and	Distance to Shadows Neighbor-objects : Smaller Than			
			Range:	0 - 15	Center:	7.5
and	or	and	LiDAR Elevation : Smaller Than			
			Range:	1550 - 1600	Center:	1575
		and	Focal Rank : Smaller Than			
			Range:	0 - 0.50	Center:	0.25

B.7 cont.

Class Name : Other Species	Classification Scheme
Unknown LRatio Tree	5 requirements
	not Similar To "Poss. LRatio DF" class, "Poss. LRatio GF, "Poss. LP" class, "Poss. PP" class, "Poss. ES" class

* NN = Nearest Neighbor
* DNVI = Normalized Difference Vegetation Index