Winter 2008

Evaluating scale to achieve optimal image classification accuracy in New Hampshire forests

Brianna L. Heath
University of New Hampshire, Durham

EVALUATING SCALE TO ACHIEVE OPTIMAL IMAGE CLASSIFICATION ACCURACY IN NEW HAMPSHIRE FORESTS

BY

BRIANNA L. HEATH

B.S., UNIVERSITY OF NEW HAMPSHIRE, 2006

THESIS

Submitted to the University of New Hampshire

In Partial Fulfillment of

The Requirements for the Degree of

Master of Science

in

Natural Resources: Forestry

December, 2008

This thesis has been examined and approved.

Thesis Director, Dr. Russell G. Congalton,
Professor of Remote Sensing and
Geographic Information Systems

Dr. Joel Hartter, Assistant Professor,
Department of Geography

Lucie Plourde, Research Scientist, Complex
Systems Research Center, University of
New Hampshire

November 25, 2008

Date

ACKNOWLEDGMENTS

I would like to sincerely thank my advisor, Russell Congalton, to whom I am indebted for not only the entirety of this project, but also for the guidance and opportunities he has provided me throughout my entire collegiate experience. Professor Congalton first introduced me to remote sensing and has provided me with the unparalleled support and advice indispensable to this project.

I would also like to extend my thanks to Lucie Plourde and Joel Hartter, of my committee. Lucie's knowledge of Pawtuckaway and the software, as well as her unfaltering optimism, were key to the development and momentum of this project. Professor Hartter's understanding of forestry and sampling theory were crucial to the successful planning and execution of field research.

Acknowledgements would not be complete without mentioning all of the current and past UNH students, faculty and staff that have contributed to my many fantastic experiences. A special thanks to Meghan Graham, who never minded the hours of troubleshooting. And a special thanks to the Natural Resources support staff that was always quick and willing to help.

A final, but no less important, thanks to my friends and family, especially my mother and father who have always believed in me. And to my best friend, Jeremy, who has kept me focused and down to earth.

A thanks to the New Hampshire Agricultural Experiment Stations for making this research possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

EVALUATING SCALE TO ACHIEVE OPTIMAL IMAGE CLASSIFICATION ACCURACY IN NEW HAMPSHIRE FORESTS

By

Brianna L. Heath

University of New Hampshire, December, 2008

New England forest complexity creates obstacles for land cover classification using satellite imagery. New methodologies such as object-oriented image analysis exhibit potential to improve classification. Although these methods have proven more accurate than traditional methods, it has been unclear what resolution yields the most accurate classification. As high resolution imagery increases classification difficulty and lower resolutions may not provide sufficiently detailed maps, this study explored the use of object-oriented classification to classify several resolutions of satellite imagery (Landsat TM, SPOT, IKONOS) at various spatial scales.

Although Landsat TM imagery yielded the highest accuracy, all classification results were unacceptable for practical use. While classification was inaccurate, segmentation successfully delineated forest stands. A comparison of 1-foot resolution aerial photography and 4-meter resolution

IKONOS imagery demonstrated little agreement between segmentation of individual tree canopies. This study indicates that finer resolution imagery is needed for segmentation and classification of individual trees.

INTRODUCTION

Meaningful scientific research is dependent on accurate data collection and analysis. Historically, forest classification and data collection have been accomplished through site visits and increasingly using remote sensing, in the forms of aerial photography and satellite imagery. Both are more efficient than sampling via site visits, which is costly and lacks total enumeration. In addition to an increased extent, remote sensing is capable of detecting more information than a human observer. Improved technology has led to an increase in both spectral and spatial resolution. Commercially available satellite imagery from Satellite Pour l'Observation de la Terre (SPOT) can achieve spatial resolutions as high as 5 meters (SPOT 2007) and GeoEye and Digital Globe data at 4 meters multi-spectral and 1-meter panchromatic (GeoEye 2007). This increase in spatial resolution brings forest classification via satellite imagery from a stand level (e.g. mixed forest) to a tree species level (e.g. hemlock). The ability to sense the environment at larger spatial scales begs the questions: what is the most appropriate scale for a particular analysis? At what spatial resolution is the highest degree of accuracy achieved?

New England forests are very complex and composed of a variety of species often within a single stand (Martin et al. 1998). These stands are often classified as "mixed forest" using a classification scheme for mapping from remotely sensed data. Mixed forest classification discounts any species differences, which makes it difficult, if not impossible, to determine species composition, which would be valuable information for a forester. Fine-resolution satellite imagery increases the amount of information available by allowing the identification of individual trees. While this increase improves the quantity of information available, it can create quality issues in terms of classification. Traditionally, land-cover classification has been approached with a single pixel method. Digital satellite imagery is essentially an equal area grid of squares (pixels), where each pixel is given a single value that symbolizes an averaging of everything on the ground "within" that square. Land-cover classification has treated these pixels as independent of neighboring pixels. The object-oriented approach to land-cover classification groups relates pixels based on predetermined parameters and characteristics. By considering neighboring pixels, this analysis can potentially identify objects at various scales (i.e., grasslands or individual trees) (Benz et al. 2004).

Although more accurate than single pixel classification (Lennartz 2004), large-scale object-oriented analysis presents several problems relating to accuracy. Resolution is currently such that a single tree and its shadow may fall within two distinct pixels, each with vastly different spectral properties. An object-oriented analysis would not consider these pixels as similar, even though they belong to the same tree. Therefore, a question arises not only as to what scale should be considered (e.g., stand or individual tree), but what spatial resolution should be used? The smallest measurement made on a map, the pixel size (spatial resolution) could be realistically small enough to detect an individual tree, or large enough so only landscape features are mapped. In the previous example, the use of a small pixel at a large scale (e.g., individual tree) would lead to the misclassification of the tree and its shadow. A study of optimal pixel size in land-cover classification using spectrometry data found that the smallest pixel did not necessarily yield the most accurate results (Rahman et al. 2003).

This study focused on identifying the most appropriate scales and spatial resolutions for land cover classification of Pawtuckaway State Park in southeastern New Hampshire, using high-resolution satellite imagery and object-oriented classification. The ultimate goal of this study was to aid satellite imagery users in achieving the most accurate land-cover mapping by allowing them to select the most effective combination of image scale and spatial resolution. The objectives for this research were to:

- create reference maps at three scales (i.e. large stand/2-acres, small stand/30-meters x 30-meters and individual tree).
- generate object-oriented maps from Landsat, SPOT and IKONOS imagery to test the effect of spatial resolution.
- compare the reference maps with the object-oriented maps to 1) assess the accuracy of each scale and 2) develop guidelines regarding the appropriate selection of imagery based on desired level of detail and accuracy.

More accurate data translates into more accurate planning and, often, better decision making. The results of this research could have wide reaching, interdisciplinary ramifications from natural resources (e.g., forestry or wildlife management) to engineering (e.g., development or land-use plans) and possibly beyond.

CHAPTER I

LITERATURE REVIEW

Background

From the first photograph taken by balloon to the multi-million dollar satellites that now orbit the planet, remote sensing has vastly improved our ability to record our environment. Remote sensing, or gathering information about something without touching it, has its roots in aerial photography. Chemically sensitive film types allow users to look into the infrared spectrum beyond what human eyes can remotely sense, allowing for data collection that would otherwise be impossible.

With the launch of the first commercial US satellite, Landsat 1, in July of 1972, remote sensing made a tremendous leap from privately flown, expensive photography to widely available satellite imagery. Although Landsat 1 was relatively short lived, it represented a new era, where multispectral sensors became the primary medium for the data collection frenzy that would follow in the next decades.

After a year in operation, Landsat 1 was replaced by Landsat 2, and then replaced by Landsat 4 and, subsequently Landsat 5, with the USGS and NASA sponsored program placing the most recent satellite, Landsat 7, into orbit in 1999. Although the Landsat program represents the longest continuously operated satellite remote sensing program in the world (USGS and NASA 2006), it is only one of many remote sensing platforms providing imagery to commercial and private entities. Built and launched by countries around the world, these satellites represent a range of resolutions and image scene sizes. For example, Landsat 1 was capable of acquiring imagery with an 80-meter spatial resolution in its multispectral bands. The IKONOS satellite, launched in 1999, is capable of capturing panchromatic imagery with 1-meter resolution (GeoEye 2007). The French SPOT 5 satellite imagery can be pan-sharpened to create 2.5-meter resolution images (SPOT Image 2007).

Satellite imagery is composed of bands, each able to sense in a different region of the electromagnetic spectrum. A more modern satellite, Landsat Thematic Mapper (Landsat TM) captures data in 7 bands: blue, green, red, near infrared, middle infrared, thermal infrared and another middle infrared, respectively. SPOT 5 imagery is characterized by 4 bands: green, red, near infrared and middle infrared. The IKONOS instrument also senses in 4 bands: blue, green, red and near infrared.

Despite the range in satellite imagery resolutions, demand continues to exist for improved technology and more data acquisition. High demand has led to a steady increase in both spatial and spectral resolution, including the development of sub-meter sensors and hyperspectral imagery. Digital processing methods, aimed at achieving the most accurate information extraction, have similarly developed to accommodate newer imaging techniques. However, many of these methodologies are still rooted in techniques of photointerpretation substituting the traditional term of minimum mapping unit for terms such as "spatial resolution."

Demand for improved resolution is intensified by the considerable scope of professions and applications that employ the use of satellite imagery. While many professions can benefit from remote sensing, much of its use has traditionally focused in military applications and the agriculture and natural resource disciplines (e.g., McCabe and Wood 2006). Other applications can include crop health assessment, timber management and water resource management, all of which include a land cover mapping component achieved through digital image processing.

Digital Image Processing

Advances in computer hardware, software and overall processing speed have facilitated the transition from analog images to digital ortho-imagery, a more flexible and useful data format widely used in land-cover classification. The optimal use of digital ortho-images for land cover classification purposes depends upon appropriate acquisition, correct processing and successful data exploration.

Acquisition of satellite image acquisition is dependent upon project goals and weather conditions. Ideally, images would be collected on a cloudless day to prevent shadows and/or missing data due to impenetrable cloud cover. In the hardwood forests of the northeast, season of image acquisition (e.g., leaf-on or leaf-off) can considerably affect the land cover classification results (Schriever

1992). An image acquired during leaf-off would render hardwood species classification challenging, if not impossible, but may still be useful for delineating coniferous species. An image acquired during senescence, on the other hand, may depict stark differences in hardwood species canopies.

Selection of spectral and spatial resolution for an image to be used in land- cover classification is largely dependent upon desired scale, and is the focus of this thesis.

Pre-processing – After acquiring the appropriate satellite images, each image must be geometrically rectified to account for satellite/sensor movement, curvature of the earth and terrain variations on the ground. As nearly all ground metrics (e.g., shape and distance) are sensitive to geometric distortions, the integrity of a map depends upon successful rectification to correct these distortions. Several rectification methods exist, although orthorectification has proven the most complete, as it accounts for terrain variation. Rocchini and DiRita (2005) found that although other rectification techniques performed well on unvarying terrain, only orthorectification performed well, regardless of terrain.

In addition to geometric corrections, rectification assigns a coordinate system (x,y) to the image, an important step for future image to image or image to map registration (Plourde 2000, Leica Geosystems 2005). Image registration is often accomplished through the use of ground control points (GCPs). GCPs are known points that can be identified in the imagery and on an existing map or via Global Positioning System (GPS) on the ground (Jensen 2005). Correct image registration, or the proper alignment of two images to a like coordinate system, is of paramount importance to the accuracy of land cover classification. Misalignment of the land cover classification map and the reference data could underestimate classification accuracy. Verbyla and Boles (2000) found that misregistration by introduced positional error caused up to a 33% change in classification of a Landsat TM image, when compared to the original classification. Studying the effect of misregistration on change detection, Dai and Khorram (1998) calculated that a registration accuracy of 1/5 of a pixel is needed to obtain less than a 10% error.

Data exploration –Data exploration, often referred to as the heart of Geographic Information Systems (GIS) (Plourde 2000), involves gaining an understanding of the imagery in terms of spectral pattern response to land cover. The purpose is to become familiar with the imagery, for a better understanding and interpretation of results.

Data exploration at its most rudimentary level involves visual inspection of the image (Plourde 2000). Generally, this includes color composite creations and variations to find band combinations that best distinguish between cover types or minimize atmospheric effects. Often the most useful composites contain some combination of visible light, middle infrared (MIR) or near infrared (NIR) bands.

Histogram analysis (plotting color response) and spectral profiles (plotting brightness response) can aid in the understanding of the spectral properties of the bands and the image as a whole (Jensen 2005). Unlike an ideal world where all objects would reflect large amounts of varying energy, the landscape objects often reflect relatively similar amounts of energy, resulting in a low-contrast image. Contrast enhancement allows for the entire range of brightness to be used (Jensen 2005) and can be simply accomplished in image processing software, such as ERDAS IMAGINE.

The derivation of additional bands from existing bands can provide useful indices, commonly based on the properties of vegetation, which draw from the knowledge of leaf physiology to provide "...dimensionless, radiometric measures that indicate relative abundance and activity of green vegetation..." (Jensen 2005). These derived bands (aka vegetation indices) can provide insights into biology (such as vegetation health) and aid in the automated, or software-based, classification of land cover. The Normalized Vegetation Index (NDVI) is a band ratio derived from the difference between the NIR band and the red band, divided by the sum of both (Jensen 2005). NDVI has been used as a seasonal gauge of vegetation activity and as is can reduce noise and spectral variation across an image (Jensen 2005).

With a multitude of potential combinations of bands and indices, a means of determining the most useful bands can be key to a successful classification. One such means is a separability, or divergence, analysis which plots each band's spectral response by class or cover type (Figure 1). This analysis can measure the class separability exhibited by each band. In the example below, the greatest separability exists within the NIR band. The separability indicates that the NIR band is an important band to use when distinguishing between species. Selecting the most suitable bands for classification reduces the processing time and dimensionality (Jensen 2005), preventing excessive complication for the researcher.

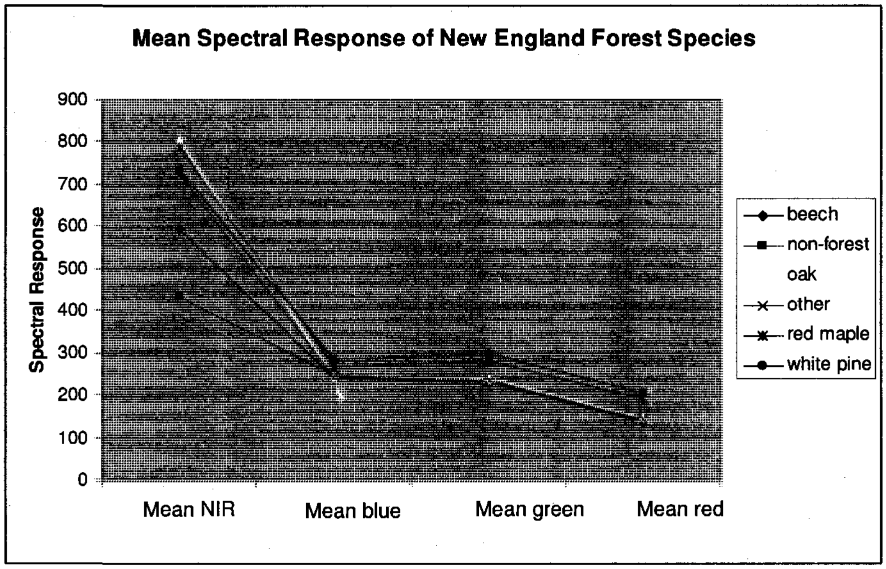


Figure 1. Sample mean spectral response (nm) based on an IKONOS image of six cover type classes.

Bi-spectral plots are also helpful to distinguish class spectral differences at a finer level. A bi-spectral plot (Figure 2) consists of two axes, each representing the spectral reflectance of a given band. The bi-spectral plot allows the researcher to determine which band is most useful for distinguishing differences between land cover types or species. Pixels are plotted based on their spectral reflectance properties, allowing a researcher to identify the "feature" space that a particular land cover class occupies (Jensen 2005). More advanced feature space analysis involves plotting spectral properties in the n^{th} dimension, representing n number of bands (Jensen 2005).

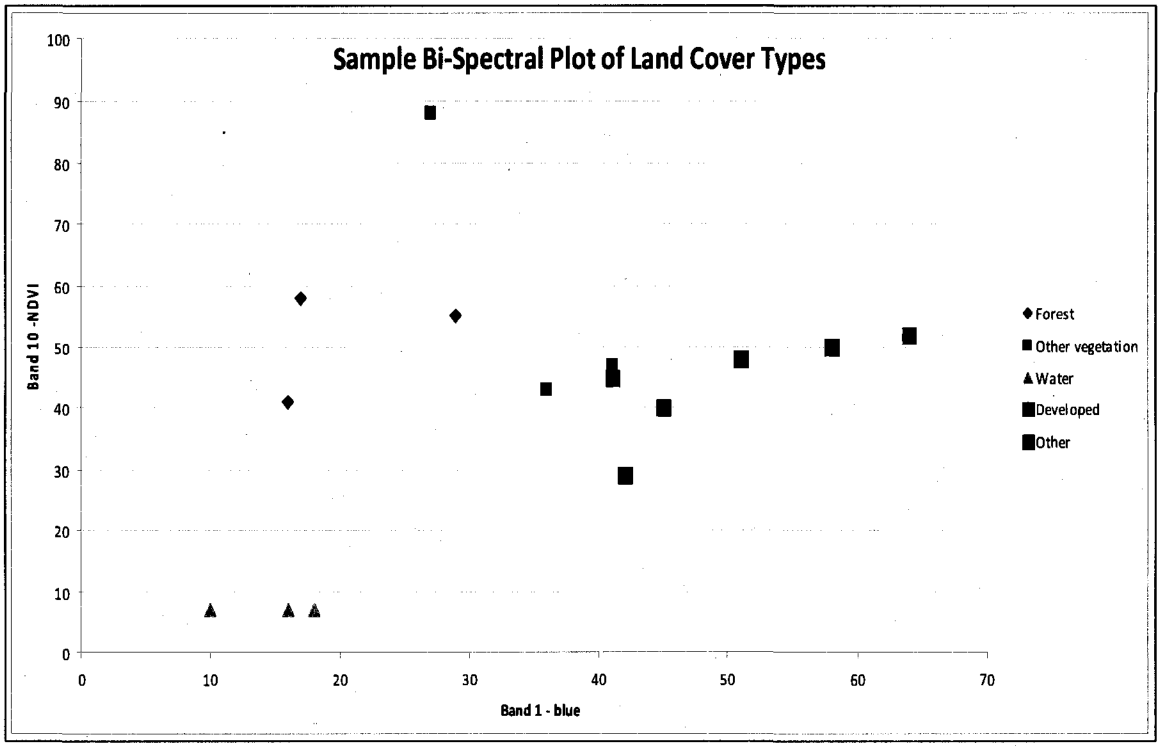


Figure 2. Sample bi-spectral plot of land cover types based on Landsat TM data.

Land Cover Mapping

Thematic land cover mapping is one of the most common uses of remotely sensed data (Foody 2002). Its attractiveness for land cover mapping manifests itself in the data's spatially continuous and map-like nature. Remotely sensed maps provide us with "bird's eye views," which are not only visually attractive, but allow for easier understanding of spatial relationships (Congalton and Green 1999). Land-cover mapping using remote sensing is possibly the only feasible way to track change at the global scale. In the past, it has been immensely valuable for tracking the land cover changes associated with the effects of global warming (Vitousek 1994).

In addition to global scales, remote sensing can also be useful for local-scale land cover mapping, such as forest cover estimation (Boyd et al. 2002) or land and water resource monitoring (Sawaya et al. 2002). These applications are of special interest to many natural resources managers, because they can provide information on available habitat for wildlife, water resources for hydro-management and forest composition. For example, Martin et al. (1998) used remotely sensed imagery to classify individual tree species in Harvard Forest, Massachusetts. Difficult to acquire through field mapping, these data would provide a forester with information on not only a species' presence or absence, but also their spatial distribution. Since land cover mapping via remote sensing can be more efficient than field sampling (depending on the research question), which lacks total enumeration, methodology is continually improving. The basic components of land cover classification by remote sensing can be loosely grouped into three stages: training stage, classification stage and the testing stage (Foody 1999).

Training Stage – The success of any remotely sensed mapping effort is largely dependent on the quality of data acquired during the training stage. Training data forms the basis for the land-cover classification. A software program will use training data, often acquired by ground visits or photo-interpretation, as standards for classification during the allocation stage. It is vitally important to select training data sites that are representative of the desired class, as the accuracy of the thematic map is dependent upon the quality of training data (Congalton and Green 1999, Foody 1999).

Classification Stage – The classification stage is guided by the training stage. Classification is the process of extracting information from remotely sensed data, comparing each pixel's spectral signatures to training data and classifying each pixel to a category with which it shares the greatest class membership (Jensen 2005). More simply stated, it is the statistical grouping of pixels into a class with the most closely related pixel properties, as determined by the training data. Classification consists of two parts: 1) labeling, which is guided by 2) a set of rules.

Testing Stage – This stage is most appropriately described by the term "accuracy assessment," as the value of any land cover map is a function of its accuracy, which is determined during the testing stage. Accuracy of a land cover map can more easily be thought of as "the degree of correctness" (Foody 2002:186). The testing stage is partially dependent upon the quality of the training data and the classification scheme (Congalton and Green 1999). However, this stage is also dependent on the quality and consistency of reference data, or what is believed to be on the ground. Land cover maps generated via remote sensing are tested against reference data (usually ground visited or photointerpreted) for classification correctness and consistency, often expressed as a percentage of agreement or using a Kappa coefficient of agreement (Story and Congalton 1986, Foody 1999, Lui et al. 2007).

Classification

The highest functionality of an image is achieved through information extraction. Although it is data, imagery must be translated into meaningful information (Jensen 2005), often thematic in nature. Land-cover classification, or using pattern recognition of spectral response to allocate portions of an image into pre-defined, discrete categories, is thematic information. Classification techniques can be grouped into three broad categories: unsupervised classification, supervised classification and hybrid classification.

Unsupervised Classification – Unsupervised classification initially requires less input from the researcher than supervised classification. Sometimes referred to as clustering, unsupervised classification is the grouping of homogenous areas of pixels into classes (Jensen 2005). Initially the researcher has only to define the number of classes (categories) into which an image will be divided. Division occurs based on specified parameters (usually spectral band properties of each pixel). Algorithms then merge pixels into like groups (clusters), which are to become classes. Once the clustering is complete, the researcher must become engaged, and assign each cluster a class, essentially labeling the clusters. *A priori* knowledge and appropriate interpretation is necessary for successful unsupervised classification (Leica Geosystems 2005).

Supervised Classification – Closely guided by the researcher, supervised classification is less computer-automated than unsupervised classification (Leica Geosystems 2005). Successful supervised classification begins with training data that must be collected based on a classification scheme with well defined categories that are mutually exclusive, totally exhaustive and hierarchical (Congalton and Green 1999). Training sites, best if located in homogenous areas, should be representative of the desired class. These sites are used for statistics extraction (Jensen 2005), which provides base information for each class (e.g., pixel spectral responses) and also to acquaint the researcher with the land cover. Guided by training statistics (the spectral and spatial properties derived from training data) and knowledge of ground conditions, the researcher identifies pixels representing recognized land cover classes (Leica Geosystems 2005). This identification "teaches" the computer the properties of land cover classes, which can then be used in algorithms that effectively compare unclassified pixels with the "known" pixels to allocate class labels.

Three commonly used supervised classification algorithms include the parallepiped, minimum distance and maximum-likelihood algorithms (Jensen 2005). The parallepiped algorithm incorporates the variation of training pixels when assigning values to unknown pixels, but may sometimes result in unclassified values. The minimum distance algorithm, which does not produce unclassified values, matches unknown pixels with the closest training data.

Perhaps the most commonly used, the maximum-likelihood algorithm accounts for both training pixel variation and similarity to training pixels.

Hybrid Classification – Hybrid classification is a combined unsupervised/supervised approach to pattern recognition. Various hybrid methodologies exist. Chuvieco and Congalton (1988) used a hybrid approach to develop training statistics through the clustering of unsupervised and supervised classification training fields (areas of known land cover), preserving the advantages of both techniques while minimizing disadvantages. Bauer et al. (1994) successfully applied the clustering technique to forest cover mapping in Minnesota. Hybrid classification provides the power to locate and label training areas using statistical clustering (unclassified) and then use those areas to classify the remaining unlabeled pixels (supervised). Hybrid classification's flexibility and higher accuracies than traditional classification (Chuvieco and Congalton 1988, Bauer et al. 1994, Lo and Choi 2004) have resulted in its widespread favor.

With all classification methodologies a common problem exists: mixed pixels. Mixed pixels are the result of landscape heterogeneity (e.g., structural, age, health and species differences). Regardless of pixel resolution, some heterogeneity will exist *within* a pixel. This creates a fundamental classification problem: the minority land cover within a pixel is not accounted for in the labeling. Often, mixed pixels are more prevalent in an image than "pure" pixels, making traditional “hard” classification approaches inappropriate (Foody 1999). Fuzzy set theory allows a pixel to have partial membership to more than one class (Jensen 2005). A user may set threshold values for class memberships, allowing fuzzy methodology to mimic environmental imprecision and human logic. ERDAS Imagine provides a fuzzy convolution tool that will “...assign the center pixel in the class with the largest total inverse distance summed over the entire set of fuzzy classification layers” (Leica Geosystems 2005). For example, a single pixel could be assigned partial membership to two or more classes (e.g., 85% hemlock and 15% wetland). This partial value assignment provides more insight into the make-up of mixed pixels. This fuzzy classification can then reduce the “salt and pepper” effect (Leica Geosystems 2005) found in more rigid classification schemes and also can allow for consideration of the natural variation within pixels by providing a more flexible interpretation of the classification results. However, fuzzy classification will not necessarily improve overall accuracy.

Error and Accuracy Assessment

The value of remote sensing data, particularly for land cover mapping, was quickly realized by the scientific community. However, the first few decades following the advent of satellite imaging were heavily focused on data collection, with less regard for data quality. In fact, accuracy was often ignored if maps looked or seemed accurate (Congalton and Green 1993). Although this mindset resulted in a plethora of data, it set a poor standard for data accuracy, with the testing stage only recently becoming a standard inclusion in land cover mapping procedure. Recently, much attention has focused on creating and meeting accuracy standards, understanding and measuring accuracy (Sader et al. 1995, Edwards et al. 1998, Congalton and Green 1999, Foody 2002) and developing optimal accuracy methods of land cover mapping (Foody 2002, Lennartz 2004). The importance of accuracy cannot be overstated; incorrect data can lead to misinformed management decisions, which in turn could have prolonged and widespread environmental ramifications.

There are many sources of error that have the potential to influence accuracy. These contributing sources can be reduced into two categories they affect: positional accuracy and thematic accuracy. Positional error refers to the spatial location or coordinates of any given object or pixel. In other words: does the map correctly identify the object's ground location? Thematic accuracy refers to the correctness of an object's or pixel's classification (Foody 2002) (e.g. is the tree the map classifies as a pine tree actually a pine tree). Positional error affects the ability to correctly locate an object, but if an object is misrepresented spatially, the likelihood of thematic error increases.

Sources of positional error vary widely, but are most commonly considered when GPS is part of a project effort (e.g., training data collection and ground referencing). When using a GPS, positional error affects the signal read by the GPS' antenna, thereby affecting the positional recording. Until 2000, the primary source of positional error was the U.S. military's Selective Availability (SA) system. Intentionally designed to corrupt satellite signals, the SA system's aim was improved national security. As a result, GPS users were left with much poorer accuracies.

Although the SA system is no longer active, satellite signal can still be degraded. Signal bounce, the result of environmental factors, contributes to a higher Position Dilution of Precision (PDOP). PDOP accounts for the constellation of satellites, and is effectively a measure of signal reliability. Higher PDOP numbers are indicative of decreased positional accuracy. Atmospheric interference or delay (which can alter signal travel time), inaccurate clocks and incorrect satellite orbit path can prove contributing factors in increased PDOP. Although PDOP is affected by satellite condition and position, environmental factors are a primary source of positional error. Research has shown that canopy cover and terrain both decrease positional accuracy (Deckert and Bolstad 1996, Rubens et al. 2002). Canopy and terrain can both obscure a user from satellite signal.

Data collection protocol also affects positional accuracy. Points can be collected instantaneously or the GPS can be set to average the point location over a specified period of time. Piedallu and Gégout (2005) found that the longer the averaging period (10 seconds compared to 1 second), the more accurate the acquired point's position, as it represents an average of multiple points. Therefore, a trade-off between speed and accuracy is established.

Dependent upon positional accuracy, thematic error relates to the quantity of cells in a map that are correctly classified when compared to the corresponding reference data (Pontius 2000). Sources of positional and thematic error can be found in all three component stages of land cover classification, with some sources of error common to all three stages. A vast body of literature has been generated in an attempt to recognize or eliminate these contributing sources of error (Lunetta et al. 1991, Congalton 1991, Congalton and Green 1993, 1999). Although each land cover mapping effort has unique challenges relative to landscape and data availability, many sources of inaccuracy are common regardless of budget, scale or landscape.

Training stage accuracy is adversely affected by a number of factors. First, data collection inconsistency can result in skewed or incorrect training data. An inappropriate sampling scheme can result in data gaps or missing trends, more commonly described as sample bias (Congalton and Green 1999). For example, choosing to only gather reference data in easily accessible area during a ground visit will likely miss trends only found on rougher or steeper terrain, which may be equally important for training purposes. Instead, the training data samples would be biased towards flat ground.

Second, training stage error can also be affected by observer error. Observer error is often as simple as collecting data in the wrong place, improper labeling (e.g. incorrect tree identification during a ground visit) or the incorrect photointerpretation of land cover features. It can also result from a misunderstanding of the classification scheme, especially when dealing with complex classification schemes (Congalton and Green 1999).

Since training stage error can negatively affect classification accuracy, it is important to minimize its impact with well designed sampling and classification schemes. Proper classification protocols can reduce some sources of error in data collection, which may lessen confusion during the classification stage. To minimize uncertainty and interpretation errors during data collection and processing, classification should begin with well-defined categories. These categories should be totally exhaustive, meaning no object is unable to be classified. Often, the inclusion of an "other" category resolves this problem. Categories should also be mutually exclusive: an object can only belong to one category. Finally, these categories should be hierarchical in nature (Congalton 1991). Thus, a hemlock tree and a pine tree should both be within the coniferous tree category.

Even using a well-defined classification scheme, some errors may be inherent. Further, classification success may be inversely related to data complexity. In mostly homogenous landscapes, large, simple land cover types will likely be more accurately classified than a highly complex forest with heterogeneous stands (Congalton and Green 1999). This classification pattern can be a function of resolution and averaging of features within a pixel. For example, a 3-meter pixel in an open field is easier to classify than a 3-meter pixel in a tropical forest, which is likely to contain the canopy branches of several different species. In the latter case, the pixel would be classified based on majority rule, resulting in underestimation of any other species within that pixel. Without corrective algorithms, shadows can also have a negative effect upon land cover classification.

Other sources of error in the classification stage are misregistration between the training data, image or reference data, inadequate or inappropriate resolution (Foody 2002) and changes in land cover. If land cover changes (e.g., fire, flood or timber harvest) occur between training data collection and image acquisition, the resulting classification will be incorrect (Congalton and Green 1993, 1999).

Inherently, training stage and classification stage error contributes to error during the accuracy assessment stage. Most importantly, however, reference data must be as representative of ground conditions as possible as it forms the basis for the accuracy assessment stage.

The Error Matrix

Although no one method of accuracy assessment is agreed upon (Foody 2002), Liu et al. (2007) maintain that the overall, user's and producer's accuracies should be reported as a minimum accuracy assessment requirement for any study. These accuracies are generated using an error matrix (Story and Congalton 1986), which compares reference data with the classified image data in a tabular form (Table 1). This coupling results in both a visual and a statistical measure (Plourde 2000). The error matrix is unique in that it provides not only the overall accuracy, but also the distribution of that accuracy amongst the land cover categories (Story and Congalton 1986).

Table 1. Sample error matrix. Rows represent classified data and columns represent reference data.

		Reference Data				
		Developed	Vegetation	Water	Other	Row Total
Map Data	Developed	10	18	13	19	60
	Vegetation	19	15	3	10	47
	Water	11	15	4	1	31
	Other	9	13	15	36	73
	Column Total	49	61	35	66	211

Overall Accuracy: 10+15+4+36/211 = .31 or 31%

	Producer's Accuracy	**User's Accuracy**
Developed	10/49 = .20 or 20%	10/60 = .17 or 17%
Vegetation	15/61 = .25 or 25%	15/47 = .32 or 32%
Water	4/35 = .11 or 11%	4/31 = .13 or 13%
Other	36/66 = .54 or 54%	36/73 = .49 or 49%

Relatively low in the sample error matrix (Table 1), the overall accuracy (31%) indicates that 31% of the map agreed with the reference data. The producer's and user's accuracies reveal how accuracies are distributed amongst the classes, from two perspectives: the producers and users. Both measurements are an indication of omission errors (i.e., an area is excluded from its correct category) and commission errors (i.e., an area is included in the incorrect category). The producer's and user's accuracies can identify in which categories these omissions and commissions most occur. For example, perhaps the user wishes to know how many times "Developed" was correctly classified as developed, and not, for example, "Vegetation." To calculate this error, the number of times developed was correctly classified, 10, (see Table 1), is divided by the number of times developed occurs in the reference data, 49. The resulting number, 0.20, indicates that developed was correctly identified as developed 20% of the time. Alternatively, the user's accuracy is calculated by dividing the number of times developed was correctly classified, 10, by the number of times it was classified on the map, 60. The resulting number of 0.17 indicates that there is only a 17% chance of visiting an area labeled as developed on the map and actually having it be developed.

Congalton and Green (1999) surmised that any incorrect classification within the error matrix was a result of four possible sources: error in the reference data, observer interpretation of classification scheme, inappropriate source of remote sensing technology or mapping error. However, data entry could also contribute to inaccuracies in any of these categories. Additionally, some reference data sampling schemes (e.g., systematic or random sampling) have been found to overestimate accuracies, as has sampling in homogenous areas (Plourde 2000) compared to other types of sampling in heterogeneous areas.

Accuracies reported through an error matrix are often accompanied by a Kappa statistic which calculates a K-hat value (Cohen 1960). Originally used in psychological statistics, Congalton and Mead (1983) found application for its use in reducing the effects of chance in representation of accuracies and allowing for the comparison of agreement to reference data between error matrices. The calculation for the K-hat value is as follows:

$$\text{K-hat} = (p_o - p_c)/(1 - p_c)$$

Where p_o is the actual agreement or number of correctly mapped samples (sum of the major diagonal) and p_c is the random agreement calculated by summing all of the proportions of samples in each map category multiplied by the proportions of samples in each reference category. The Kappa analysis normalizes the values of error matrices, by reducing the effects of chance. This normalization allows for accuracy comparisons between maps and error matrices derived from differing reference data (Foody 2002).

Sampling Design

A well-designed sampling scheme can lessen error in both the training and accuracy assessment stages. However, the many components of sampling design must meet statistical goals and project goals. Sample design must also be logistically possible and tailored to meet needs and challenges of an individual project.

Regardless of sample design, an important distinction between training samples and reference samples must be made. Although often similarly collected in the field, training samples and reference samples serve two purposes. The former serves to guide the classification stage of the land cover analysis, while the latter determines the correctness of the image produced. Both samples must be independent. That is, the sample units used to train the data can not be the same sample units used to assess the accuracy of the classification. Clearly, this would result in an inflated estimation of success.

Sample design begins with determining the appropriate sample unit, which may be a point or an area (i.e., pixel, polygon or fixed plots). Point sample units have no extent, while areas have some size associated with them. Pixel and polygon samples are closely linked to land cover, with pixels being uniform in size and shape and dependent upon an image resolution. Polygon samples may be based upon a specific land cover characteristic, such as a forest stand, but are bound, in some processing software, to the specific map on which they are made (Stehman and Czaplewski 1998); meaning that a polygon may not look correctly delineated when overlayed upon other imagery than the original. The ability of a researcher to locate a specific sample unit should be considered when determining sample unit size. For example, if a sample location is recorded using a GPS device, the sample unit should be large enough to account for any inaccuracy in the GPS position. Consider a handheld GPS, which may have an error of more than 15 feet and the effect that would have upon a 3-foot square pixel (polygon) sample. It would be possible for a GPS point to be 12 feet away from the pixel sample.

Sample size is often determined by project specific and statistical power requirements. For training stage data collection, several training sample units per class may be required to adequately represent the variability within a certain cover type (Joyce 1978). When collecting reference data samples, an adequate number must be selected to represent landscape variability across the study site (Stehman and Czaplewski 1998) and also to achieve sufficient statistical power. Although practicality and expense affects sample size, a "general rule of thumb" for accuracy assessment sample size when using an error matrix is a minimum of 50 samples per land cover categories (Congalton 1991, Congalton and Green 1999).

Sample placement should be considered in acknowledgement of spatial autocorrelation, which is the effect that one unit may have upon another sample in the same neighborhood. For example, if a point is collected in a hemlock stand, it is more likely that a nearby point would be hemlock than another species. This likelihood violates the assumption of sample independence (Congalton and Green 1999). Two sampling schemes that avoid placement bias, thereby increasing the likelihood of adequate sampling, are simple random sampling and systematic sampling (Plourde and Congalton 2003). However, some combination between the two may be necessary in light of field obstacles (e.g., gated roads, steep terrain, and access) and funding for field work (Congalton and Green 1999). Land cover heterogeneity should also be considered. While past accepted methods have favored placing samples in contiguous, homogenous cover types to reduce error, recent research has found such practices may overestimate overall map accuracy (Plourde 2000).

Sampling protocol (e.g., what attributes will be measured and how they will be measured) is of paramount importance to data consistency. For example, in a land cover sampling system, consider what the basis for measurement is. Will the observer use transects within a unit? Will the observer base classification on a majority rule? Will basal area, canopy enclosure or another "hard" measurement determine species dominance or will visual estimation suffice in determining cover type?

Realistic sample design often varies from the ideal. Stehman and Czaplewski (1998:342) best stated: "A practical accuracy assessment sampling strategy often represents a compromise, with the overall design goal being adequacy for all critical objectives, not optimally for any single objective."

Reference Data Collection

Reference data collection can be similar to training data collection, but should never be one in the same. In some cases, new reference data collection may not be necessary if suitable data already exists. However, if pre-existing data does not follow an appropriate classification scheme, is outdated, incorrect or otherwise inappropriate for use (Congalton and Green 1999), the user must collect reference data either by photointerpretation or field visits.

Congalton and Green (1999) found photointerpretation to be an effective method of reference data classification in situations with a few, simple categories. However, at some scales, photointerpretation was found to be an inappropriate method. Photointerpretation ideally should include field visits to ensure interpretation accuracy. Brogaard and Ólafsdóttir (1997) found photointerpretation costly and time consuming as it required camera calibrations and field work.

Sampling design and protocol must be considered when collecting reference data. Ideally, reference data collection should follow the same design and protocol as training data collection.

Per-pixel Image Processing v.

Object-Oriented (Segmented) Image Processing

Traditional satellite digital image processing techniques have focused on a single pixel approach, in which each pixel is classified independent of neighboring pixels. The advent of high spatial resolution (for the purposes of this study; less than or equal to 10-meter) satellite imagery such as IKONOS has created a demand for new processing techniques, capable of extracting new levels of information (Jensen 2005). This demand resulted in the development of the segmented or object-oriented image processing approach: a hybrid approach.

The object-oriented image classification approach more closely mimics the human process of object delineation and classification. Humans naturally delineate common objects on the basis of not only color, but texture and context, not on a per-pixel basis (Warner et al. 1998, Definiens AG 2006). Object-oriented software classifies by segmenting pixels into "zoned partial areas of differing characteristics" called image objects (Definiens AG 2006:3). Image objects are created based on the properties of spectral response, texture (smoothness and compactness) and context (relation to neighboring pixels), all of which are subjective. Some object-oriented image software, such as Definiens Professional, is capable of creating nested objects at various scales, allowing for classification at landscape and individual tree scales.

In addition to providing increased information, the inclusion of texture in a segmentation analysis can increase overall classification accuracy (Franklin et al. 2001, Lennartz 2004, Addink et al. 2007). Franklin et al. (2001) found that combining spectral and texture data increased accuracy to 75%, compared to 54% for isolated spectral data and 70% for isolated texture data when classifying forest structure and species. The use of contextual information (or spatial autocorrelation in segmentation), formerly achieved by a moving window filtering approach, may reconcile the physical differences detected by a computer and the human eye (Stuckens et al. 2000), resulting in a map that appears more visually correct (Figure 3).

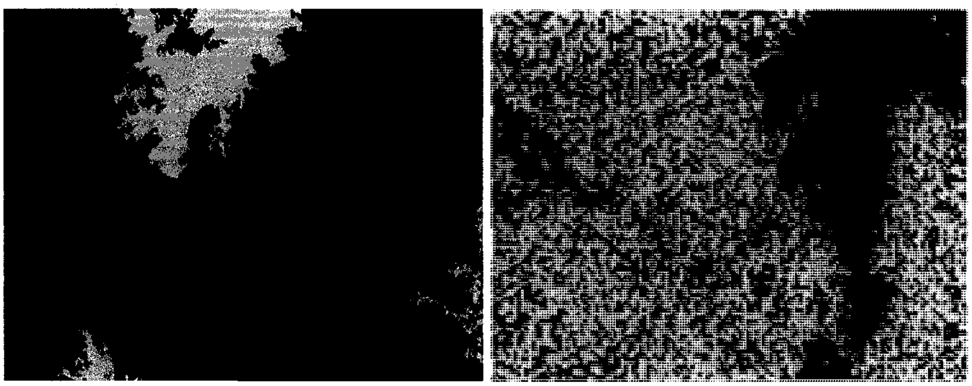


Figure 3. A sample segmented or object-oriented (left) and unsegmented or per-pixel (right) image classification.

Role of Scale and Pixel Size in Object-Oriented Segmentation

Object-oriented image analysis promises increased accuracies. However, it also provides increased complexities due to increasing spatial resolution of imagery. New England forest classification requires use of various scales, as species composition, stand density, stand size, individual crown size and shape varies (Warner et al. 1998). A common problem with thematic classification, regardless of resolution, is the averaging process that occurs within an individual pixel. For example, a 10-meter pixel classified as oak, may contain other species in addition to oak which are disregarded because oak comprises the majority of the pixel. At finer resolutions, a pixel may contain the branches of multiple species rather than an individual tree. Essentially, the pixel "covers" the space between two trees, creating a question as to how this should be properly classified. This resulting "mixel" problem is particularly common in continuous landscapes (e.g., a forest canopy) and can result in under- or over-represented land cover categories.

Selecting the appropriate remote sensing technology source is of paramount importance to achieving desirable accuracies and is responsible for minimizing the mixel affect. Besides logistical limitations, the type of remote sensing is largely dependent upon project goals, landscape and mapping scale. Spatial resolution should be selected to match desired scale, while type of data (e.g. image data or Light Detection and Ranging (LIDAR) data) should not only support project goals, but also be appropriate to landscape. For example, a project with an objective to classify a heterogeneous forested area using spectral response could incorporate an image acquired during leaf-on with a fine enough spatial resolution to detect the mixed nature of the forest.

While lower spatial resolution images (30-meter +) provide a decent representation of forest stands, higher spatial resolution imagery may be needed to identify individual tree species. However, the accuracy of high spatial resolution image classifications is not always superior to lower resolution classifications (Irons et al. 1985, Migeul-Ayanz and Biging 1997). Although a tendency exists to obtain the highest spatial and spectral resolution imagery that technology and funding allows, this may not always be the most appropriate solution (Jacquin et al. 2007) for achieving higher accuracies. Rahman et al. (2003) found that a pixel size of 6m was most suitable, compared to 4m to 20m pixels, to study the ecosystem function of plants in the grasslands and chaparral of southern California. Despite a decrease of mixed pixels with increased spatial resolution, there is increased spectral, within-class variations (Hay et al. 1996), potentially making classification difficult.

The increase in spectral variation inherent within classes must be considered when attempting to use imagery with increased spectral resolution to distinguish between spectrally similar species. Although the vast majority of satellites currently employ multispectral remote sensing systems, there is a wide range of spectral variation available. Spectral resolution is a measure of the number and size of wavelengths collected by sensors (Jensen 2005), often referred to as bands. Both the number and the width of available bands vary with imagery. For instance, hyperspectral imagery often features hundreds of narrow-width bands, while broad-width band images (such as IKONOS, SPOT or Landsat TM) have fewer than 10 bands.

At higher spatial resolutions, varying forest stem densities and crown sizes may create different texture patterns, even within the same species (Franklin et al. 2001). Consider two oak stands, one regenerating and one mature. Basal area and stem density will be vastly different between the two stands, despite the fact that they are composed of the same species. Therefore, texture-based segmentation, even with the inclusion of spectral data, may mistake these stands as two separate forest classes, rather than both as oak, which might be good if the goal is to detect two different age structures. The segmentation at the individual tree scale also has particular problems. Generally, the tops of trees are the brightest because they are sunlit (Warner et al. 1998),

Selection of spatial resolution and spectral resolution can have a profound impact upon the resulting classification accuracy (McCloy and Bocher 2007). It is also important to consider scale of segmentation, or what spatial scale (e.g., a stand vs. an individual tree), when selecting imagery as it can affect overall classification accuracy (Addink et al. 2007). This selection should be appropriate to the scale of the classification (McCloy and Bocher 2007, Jacquin et al. 2007). For example, a 30-meter resolution image would not be suitable for the identification and classification of individual trees, as it is likely that multiple tree canopies would be averaged within the pixel, which would defeat the purpose of attempting to classify an individual tree. However, the 30-meter resolution imagery may be suitable for a stand scale classification. It would then follow that higher spatial resolution is needed for finer classification scales.

In selecting imagery to use in an automated classification, it is also important to remember that automated land cover classification is heavily dependent up on spectral response of the land. Toll (1985) found that spectral and radiometric resolution (a measure of the satellite sensor digital capability) was more important than spatial resolution. Therefore, it may be more beneficial to sacrifice spatial resolution for improved spectral resolution. Understanding the trade-offs between spatial and spectral resolution in terms of achieving the most accurate classification at the desired scale is the essence of my research.

CHAPTER II

METHODS

Study Area

This research focused on the classification accuracies achieved at various scales and resolutions. Since classification accuracies tend to be higher, often artificially, in homogenous landscapes (Plourde and Congalton 2003), the study area was chosen for its diversified, heterogeneous nature. Representative of the complex structure of New England forest, the study area (Figure 4) is completely contained in Rockingham County in southeastern New Hampshire within the towns of Deerfield and Nottingham. Approximately half of the study area (4,146 acres) is publicly held land within the Pawtuckaway State Park. The remaining northern half of the site (4,621 acres) is privately held land, the majority of which can be characterized as a wooded upland (>25% of the landscape is forested) (Sperduto and Nichols 1994).

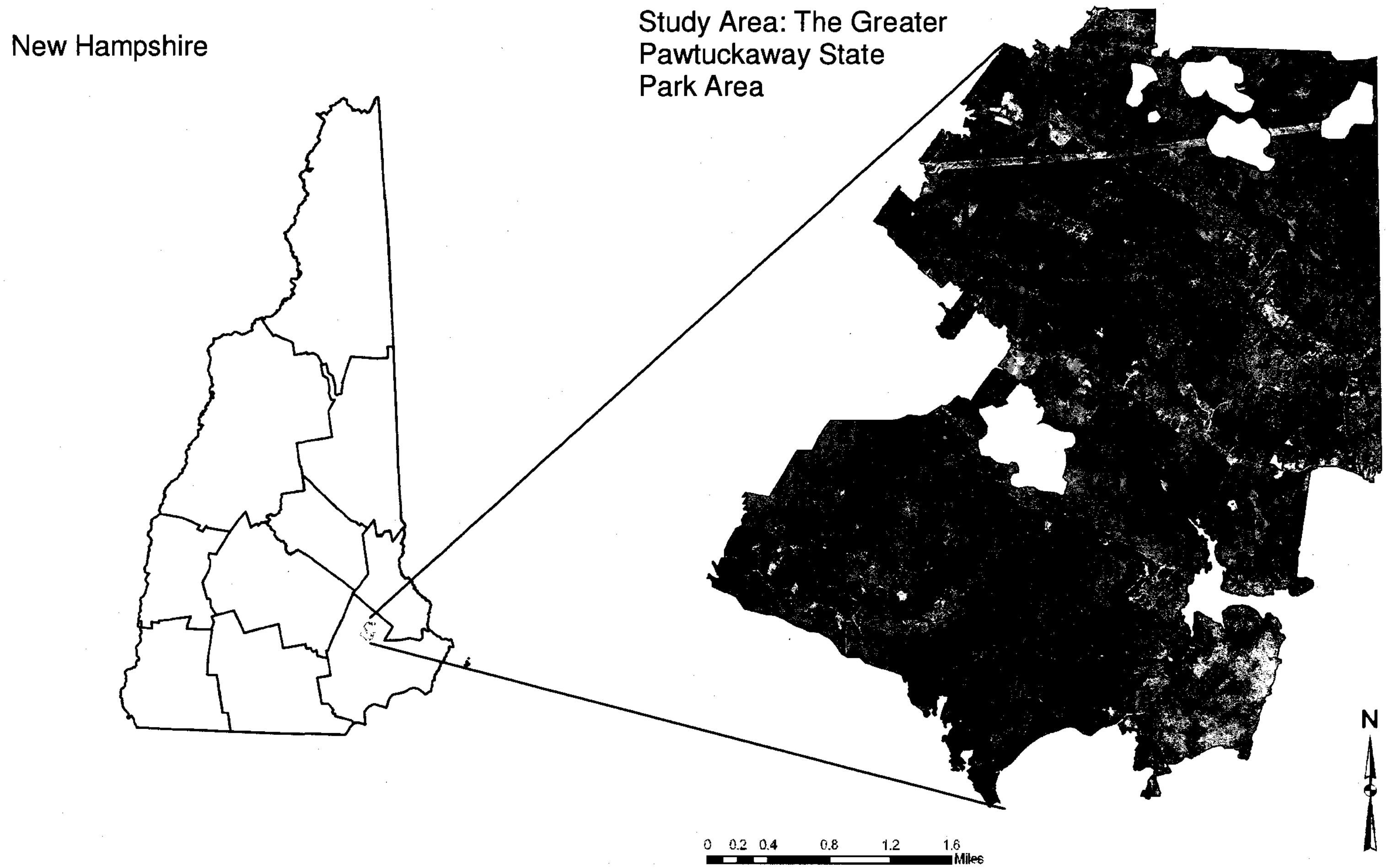


Figure 4. Research study site comprised of Pawtuckaway State Park and adjacent northern property, enlarged from the State of New Hampshire with county boundaries, depicting April 2005 1-foot resolution full-color aerial photographs clipped to the study area boundary.

Created by volcanic interaction in the late Devonian period, the landscape of Pawtuckaway features three apparent peaks: North Mt. (1,011 feet.), Middle Mt. (800 feet) and South Mt. (908 feet). Each peak's summit is characterized by exposed rock, while the majority of the area contains sedimentary rocks with plutonic rocks imposed. Shale, sandstone, dolomic limestone, phyllite, quartz-mica schist, quartzite, lime-silicate and shaly sandstone are the dominant rocks present throughout the study area (Freedman 1949). The summits and some lower elevation areas also have circumneutral cliffs, which occur when parent bedrock and fractured groundwater transport cations to the rock face (Sperduto and Nichols 1994). Study area base elevation begins at 250 feet above mean sea level, with the highest elevations located in the southern park portion of the study area.

Soil type varies throughout the study area with the Chatfield-Hollis-Canton complex accounting for approximately 45% of study area soil type. Canton gravelly fine sandy loam, greenwood, water and Montauk fine sandy loam comprise 23.22%, 7.11%, 6.22% and 4.70% of the study area soil coverage, respectively. The remaining study area is covered by thirteen soil types, each comprising less than 4% of the study area (USDA 2006).

The climate of the Pawtuckaway State Park is characterized by the typical seasonal changes of the region, including leaf senescence in autumn and persistent snow cover throughout the winter months. Average temperatures range from a mean of 70.2 degrees Fahrenheit in July to a mean of 23.5 degrees Fahrenheit in January. Annual precipitation averages 50.41 inches, with snowfall accounting for approximately 22 inches (New Hampshire State Climate Office 2008).

As evidenced by remaining stone walls, the greater Pawtuckaway area experienced an agricultural history similar to that of the rest of New Hampshire. Affected by European settlers and their descendents, the majority of the state was cleared for farm land by 1850 leaving only 45% of forests remaining statewide. However, demographic and lifestyle changes resulted in a resurgence of forested land to 87% statewide coverage in 1983 (NH DRED 1996).

A small portion of the study area contains residential housing, abutting the interior border of some study area boundaries. Pawtuckaway State Park is primarily used for recreation with various multiple-use trails throughout the park. The study area was clipped to exclude the campsites and camping facilities that are associated with the eastern edge of the park, as their impact is significant and detectable on satellite imagery. However, hikers and bikers do frequent the interior of the study area. Peak use occurs during summer weekends (Manning and Cornier 1980), with significantly less recreational impact occurring during the winter season. However, impact is confined to trails and ground-level plant growth as woody shrubs and trees are far more resistant to trampling (Cole 1995). As these larger tree species are the targets of the study, recreation is unlikely to affect results of this study.

As evidenced by visual field inspections of the Pawtuckaway State Park and discussions with personnel from the Department of Resources and Economic Development, the park's managing state agency, some small-scale forest harvesting occurs within the park. However, it is infrequent and covers little area (<100 total acres from 1998 to 2005). As was evidenced by recent paint markings and an accompanying sign, a small (<10 acres) portion of the park is slated for a future selective cut as part of the State of New Hampshire's park management plan. However, little literature is available regarding the frequency or extent of past or future forest management plans, with all of the state's lands subject to one statewide plan. Regardless, much of the forested land in the park is situated on steep inclines, with the lower terrain dotted by or completely comprised of wetland areas. Both the inclines and the wetland make forest harvesting for a vast area of the park unfeasible.

Unlike the southern portion of the study area, the adjacent private land contains some scattered houses. Although residential areas exist, the vast majority of the privately owned area is forested, of which a 25% portion in the northern area is actively harvested (Lennartz 2004) by its owner, a small scale lumber company. In total, over 4,400 acres of the study area are designated conservation lands (Society For the Protection of New Hampshire Forests 2007).

Current forest species combinations have been consistent for the past 2,000 years (NH DRED 1996), including a mix of coniferous, deciduous and integrated coniferous/deciduous. This type of species composition is characteristic of the Central Hardwoods-Hemlock-White Pine Forest Region of New England in which the study area is situated. Within this region, average date of last frost is May 1, with the average date of first frost falling on October 15, averaging 150 to 180 frost-free days (DeGraaf and Yamasaki 2001). Principally deciduous, the majority of trees in this region lose their leaves in the autumn (roughly September to November).

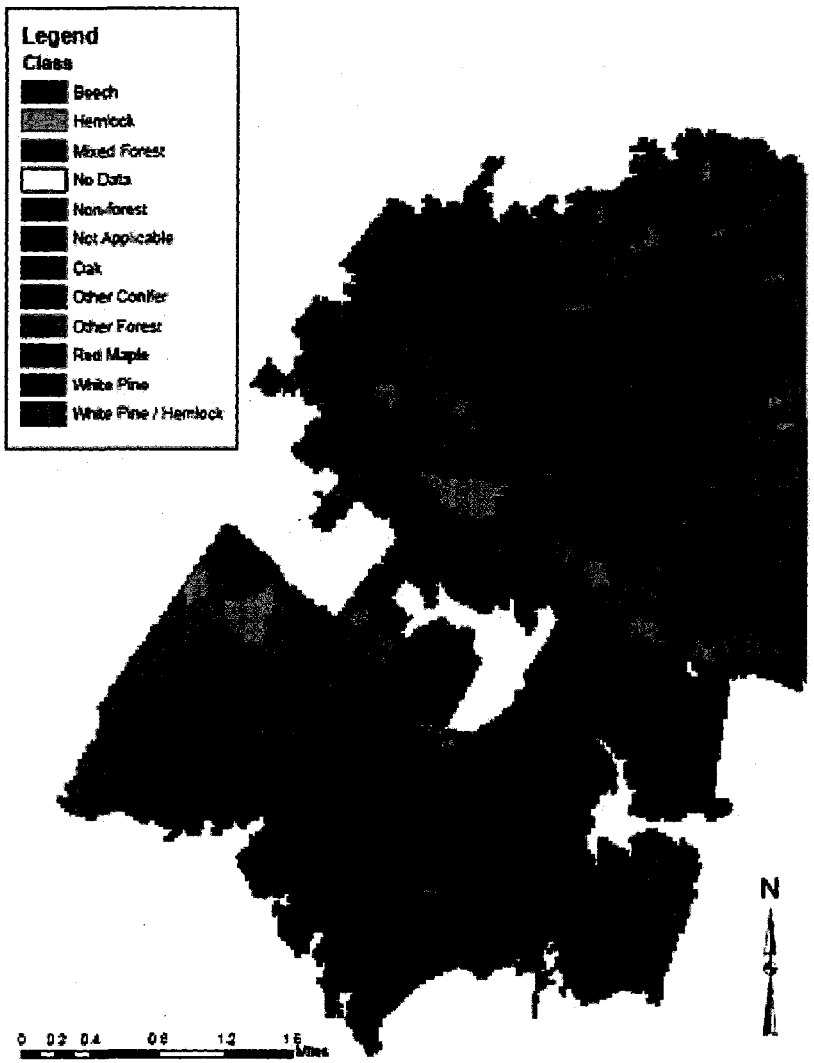
The Pawtuckaway area has several forested and non-forested natural community systems found in New Hampshire, including the aforementioned circumneutral cliffs as well as hemlock, hemlock-hardwood-pine and Appalachian oak (*Quercus spp.*) rocky woods forests, all of which are rich mesics (Sperduto 1995). Pawtuckaway's forests are indicative of well-drained, nutrient poor, acidic, glacio-fluvial soils. Pawtuckaway is also host to the rich red-oak (*Quercus rubra*), rocky woods system, which includes red maple (*Acer rubrum*) swamps. These forests are all defined as having greater than 25% tree cover and are best described as belonging to a group of mid-elevation community systems of New Hampshire, as opposed to the high-elevation spruce-fir systems (*Picea spp.*) (Sperduto and Nichols 1994).

Stand age varies from early successional species to mature forests. Wetlands, red maple swamps and small ponds are scattered throughout the landscape, although the majority of the area is forested. Dominant tree species in the greater Pawtuckaway State Park include eastern white pine, oak (*Quercus spp.*), eastern hemlock, maple (*Acer spp.*) and American beech (*Fagus grandifolia*).

These forests, particularly those with hemlock, provide excellent cover for wildlife, including white-tailed deer (*Odocoileus virginianus*). Their dense structure provides ample cover and decreased snow depth, allowing for easier wildlife movement. In addition to forested lands, the Pawtuckaway area has natural community systems such as vernal woodland pools and marsh habitats (Sperduto and Nichols 1994) which are often host to a variety of herptofauna, avian species and beaver in deeper water areas.

Pre-existing Data

Previous research within Pawtuckaway State Park (Pugh and Congalton 2001, Plourde and Congalton 2003, Lennartz 2004) has established a classification scheme meeting Congalton's (1991) criterion. This classification (APPENDIX A) is a quantitative interpretation of the Society of American Foresters (Eyre 1980) guide to forest stand type. As pre-existing reference data are crucial to this study, it was important to ensure that classification schemes were compatible. In addition to pre-existing, downloadable datasets from New Hampshire Geographically Referenced Analysis and Information Transfer System (NH GRANIT), two field-based datasets from prior research within the study area were used. Pugh (1997) created a 2-acre minimum mapping unit (MMU) vegetative reference map through field validation and photointerpretation. During her thesis research, C. Czarnecki (2006) collected 213 reference data points of stands in the summers of 2005 and 2006. Each point was taken in the center of a 30-meter x 30-meter area stand of uniform composition (not necessarily homogeneous) (Figure 5).



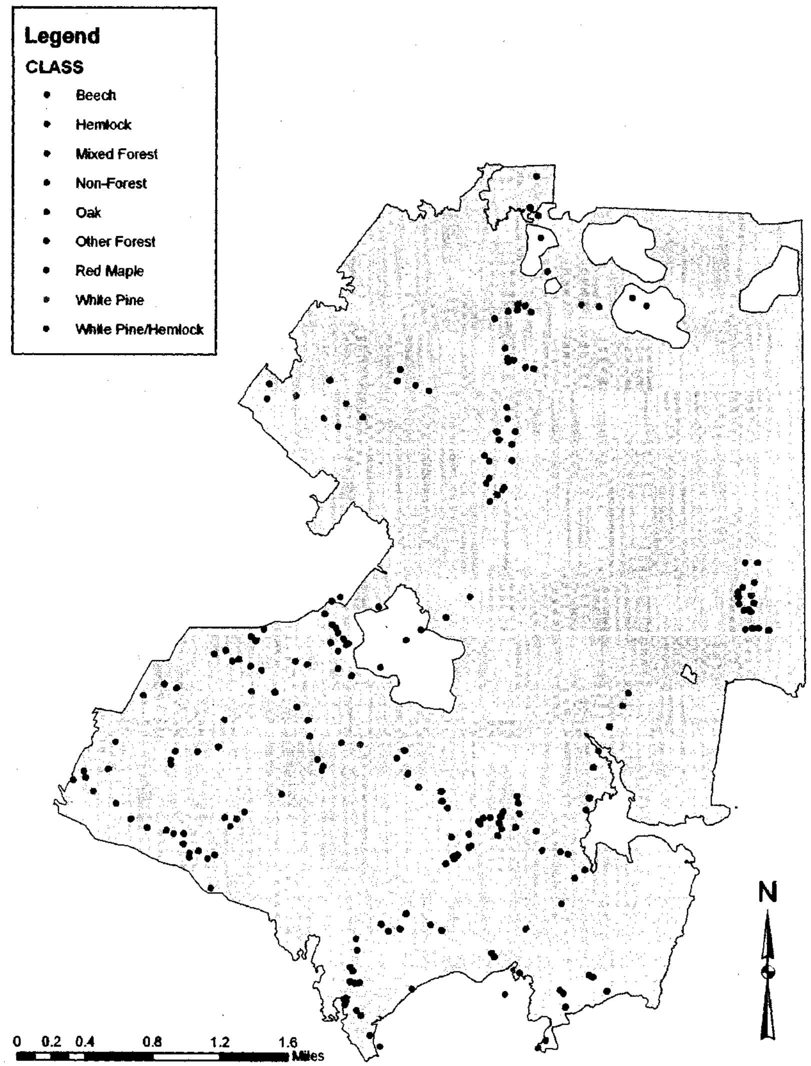


Figure 5. Previously existing reference data clipped to the study area boundaries. Figure at left is the 2-acre minimum mapping unit classified data. Figure at right depicts the 3- meter GPS data points.

Data Classification

As this study measured the accuracy of classifications, a well constructed classification scheme was crucial. Congalton (1991) stressed that classification schemes should have categories that are well defined, hierarchical, mutually exclusive and totally exhaustive. It is essential that the classification scheme be applicable to the overall species content, but also sufficiently simple for collecting field data. As pre-existing datasets were classified based on the Society of American Foresters classification scheme (Eyre 1980), field data collected and subsequent classification maps shared identical cover type categories and definitions to allow for comparison of classifications and their accuracies. Preliminary cover type investigations confirmed that class categories matched the predominant stand/land cover types (Figure 6). A total of nine cover classes were used in the large stand (2-acre) and small stand (30-meter x 30-meter scale) : White Pine, White Pine/Hemlock, Hemlock, Oak, Red Maple, Beech, Other Forest, Mixed Forest and Non-Forest. As they represent mixed species, Mixed Forest and White Pine/Hemlock were not used at the individual tree scale.

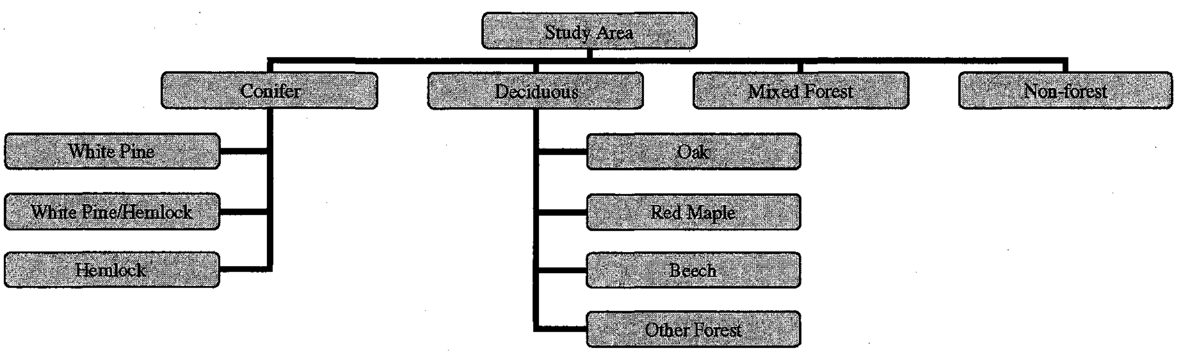


Figure 6. Classification hierarchy for labeling of study area landscape, based on the definitions of the Society of American Foresters and previous research in Pawtuckaway State Park. For definitions, see APPENDIX A.

Collection of Field Data

Field data was needed to test the accuracy of automated classification at three scales: 2-acre landscape areas, 30-meter x 30-meter stands and individual trees. Reference data for the 2-acre level were previously obtained (Pugh 1997), as was a partial dataset for 30-meter x 30-meter stands by Christina Czarnecki (2006). However, to achieve a minimum of 50 points per class (Congalton 1991, Congalton and Green 1999), more reference point samples of 30m X 30m uniformly comprised stands were needed.

Field work was conducted from September to early November 2007. Uniform 30m x 30m stands were located using 2005 1-foot resolution color aerial photography and the stand centers recorded using a handheld Garmin 12XL GPS unit. All points were averaged on-site for a minimum of two minutes (to minimize positional error) and then uploaded to a computer using a GPS to Geographic Information System (GIS) transfer software program called GPS DNRGarmin, developed by the University of Minnesota (2008) to transfer Garmin data into shapefiles. Once GPS positions were recorded, the stand composition was evaluated by visual estimation of canopy cover, as this represented the area most likely to be captured by satellite imagery. The visual estimate of the canopy cover was recorded, essentially capturing the observed percentage of each species (e.g., 50% oak, 30% maple, 20% pine). Class type was determined based on composition and the classification scheme and then recorded. Between 10 and 15 (dependent upon abundance) additional points were collected in the same manner to serve as training data.

The spatial distribution of the pre-existing 30-meter x 30-meter reference data points was considered when developing a sampling scheme for the supplemental collection of points. A visual analysis of the pre-existing point data overlayed upon a roads layer in a GIS indicated that all the data were collected on or immediately adjacent to the recreation roads and trails. As the majority of the roads and trails within the study area have a heavy forest canopy, it was concluded that there would be little to no bias as a result of the collection location. However, the collection of additional points was more carefully executed: Although no formal sampling scheme was implemented given the amount of data already existing, every effort was made to distribute the collection of additional points off-road/trail to capture as much landscape variation as possible. Care was taken to avoid collecting data in areas that had been obviously harvested within the past 10 years.

No pre-existing reference data were available at the individual tree scale. A total of 50 reference points per category were collected in November of 2007 using the same Garmin 12XL GPS unit. The positional accuracy of points, also collected using the automatic Garmin averaging function, was visually verified (again, to minimize positional error) by uploading and overlaying April 2005 full color aerial photography at a 1-foot resolution provided by NH GRANIT. The aerial photos were captured during leaf-off, allowing for the discrimination of individual trees and the verification of GPS reference points (reference trees). To insure integrity of reference data, reference trees adhered to several criteria as part of the sampling:

1. tree was tall enough for all of its canopy to be visible in satellite imagery
2. tree canopy diameter (at its widest) was >3 meters
3. reference tree canopy did not touch another reference tree's canopy
4. reference tree was not located immediately adjacent to a road or path

The canopy diameter for each reference tree was paced out at its widest point and recorded, as was tree species and classification. Additional points were collected in the same manner to serve as training data. Again, care was taken to avoid collecting data in areas that had been obviously harvested within the past 10 years.

Image Acquisition

To facilitate a match between desired classification scale and image resolution, three images were acquired (Figure 7-9). An 8 bit Landsat 5 TM image (Figure 7) of the study area was obtained on September 7, 2007. Although a newer satellite, Landsat 7 imagery was not used due to sensor miscalibrations and resulting image striping. The image has seven bands: blue (0.45-0.52μm), green (0.52-0.60μm), red (0.63-0.69μm), near infrared (NIR) (0.76-0.90μm), middle infrared (MIR) (1.55-1.75μm), thermal (10.4-12.5μm) and middle infrared (MIR II)(2.08-2.35μm) portions of the electromagnetic spectrum with all but the thermal band (spatial resolution of 120 meters) having a spatial resolution of 30m. A SPOT 5 image (Figure 8) was acquired on August 16, 2007 with a 10 meter spatial resolution and NIR (0.78-0.89µm), red (0.61-0.68µm), green (0.50-0.59µm) and MIR (1.58-1.75µm) bands. Band order varies from traditional order as it was rearranged prior to purchase to display automatically as a Color Infrared (CIR) image. An IKONOS image (Figure 9) acquired on September 5, 2001 was also used in the study. The 16 bit radiometric resolution image had four bands covering the blue (0.45-0.52µm), green (0.51-0.60µm), red (0.63-0.70µm) and NIR (0.76-0.85µm) portions of the electromagnetic spectrum with a spatial resolution of 4m for each band. All images were nearly cloud free, except several small areas within the IKONOS imagery. The study area boundaries were adjusted to exclude cloud obscured land cover.

Although the radiometric resolutions and dates of acquisition vary for each of the images, there was no need to perform an atmospheric correction as training data would be derived from each image to classify each image individually (Jensen 2005). By calibrating training data to each image's spectral responses, the radiometric variations within the images are captured for the classification stage. Additionally, there was no aspect of change detection in this study making any spectral variation due to changed atmospheric conditions a non-issue.

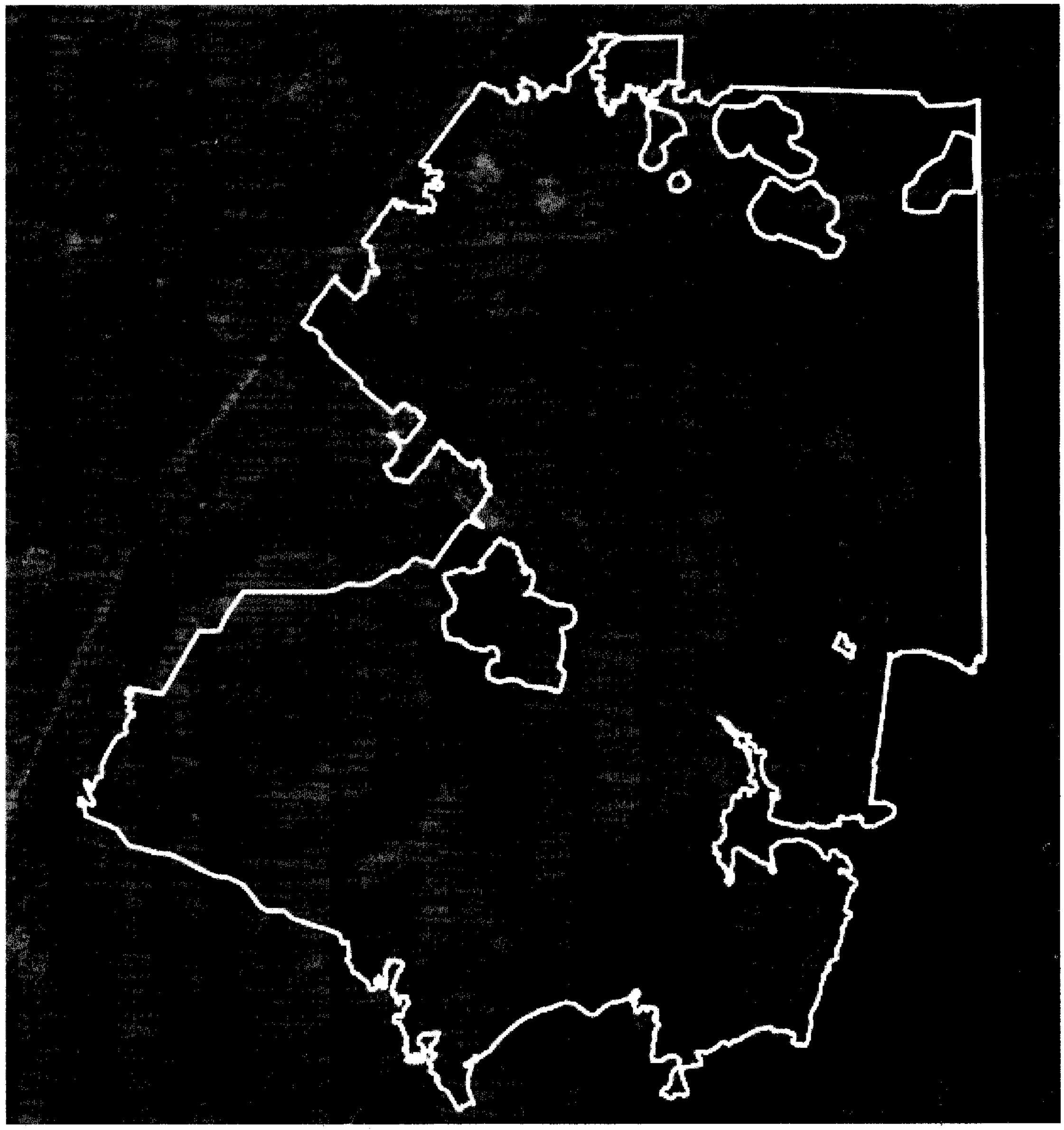


Figure 7. The Landsat TM imagery (displayed as NIR, Red and Green through R,G,B channels) acquired for the research project (Landsat Scene ID#: LT50120302007250EDC00), overlaid with the study area boundary of the greater Pawtuckaway State Park area.

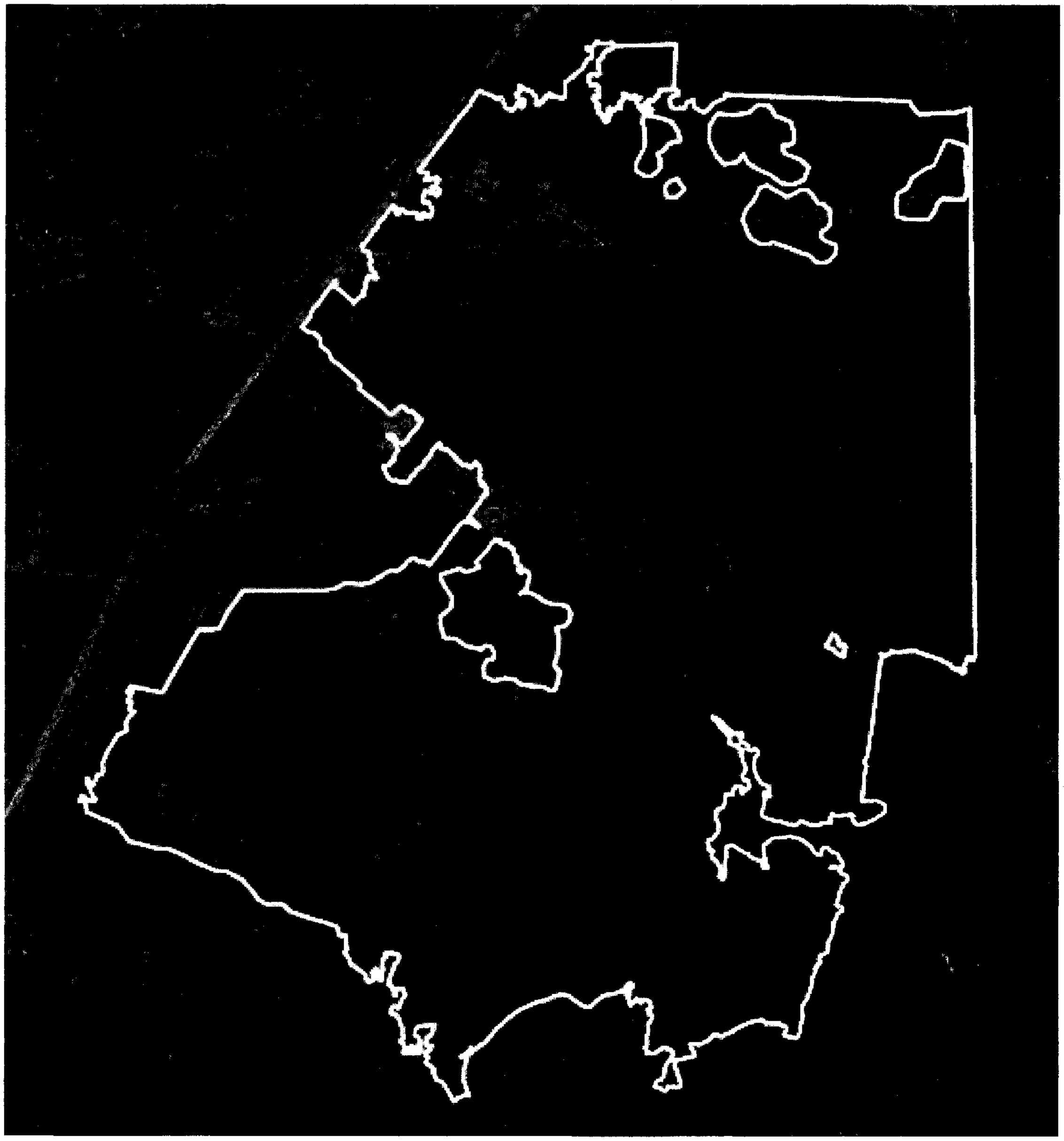


Figure 8. The SPOT imagery (displayed as NIR, Red and Green through R,G,B channels) acquired for the research project, overlaid with the study area boundary of the greater Pawtuckaway State Park area.

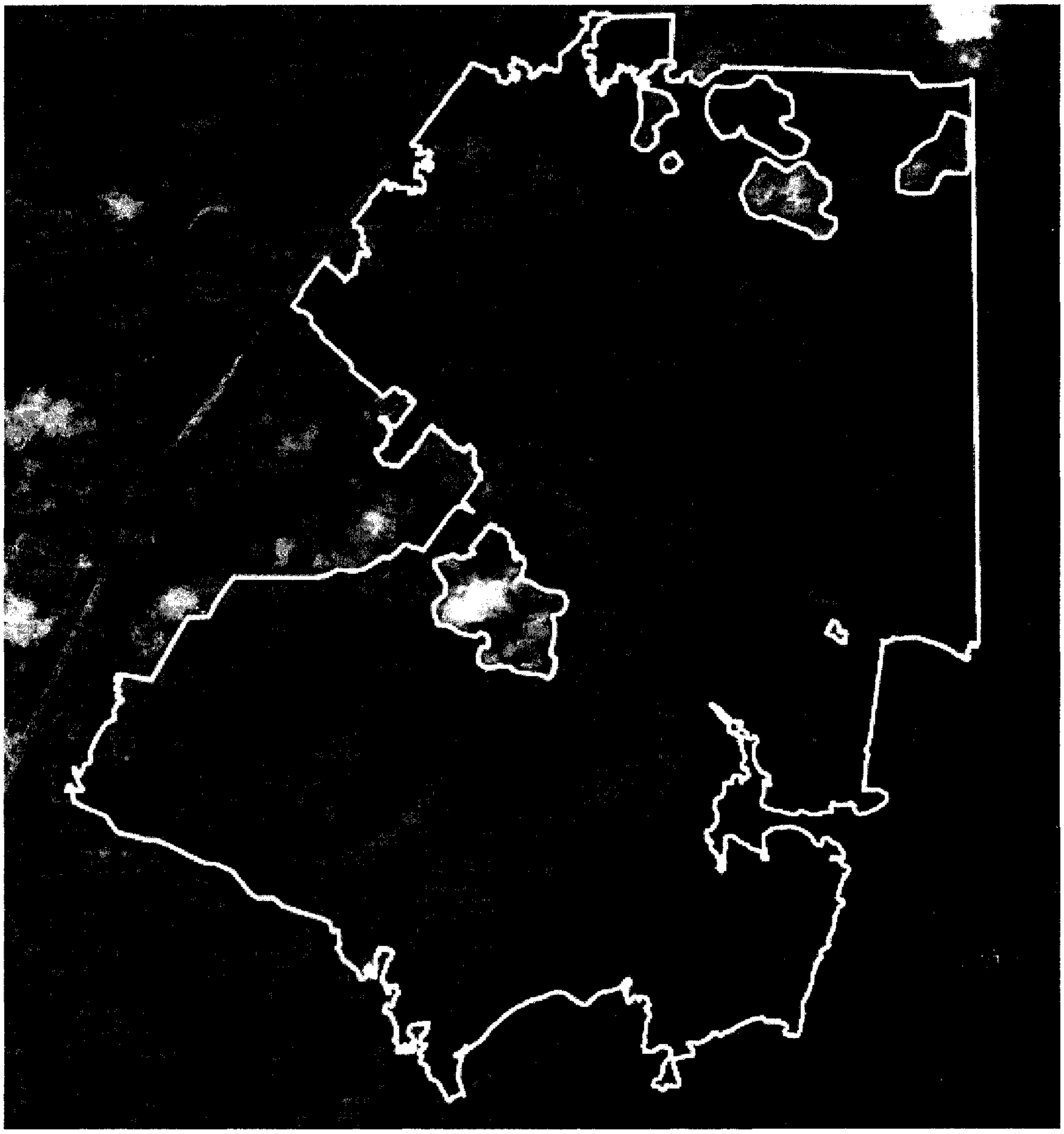


Figure 9. The IKONOS imagery (displayed as NIR, Red and Green through R,G,B channels) acquired for the research project, overlaid with the study area boundary of the greater Pawtuckaway State Park area.

Both the IKONOS and SPOT images were received already registered to NAD 1983 New Hampshire State Plane Feet (FIPS zone 2800). The Landsat TM image was reprojected from UTM meters (zone 19, WGS 84) into NAD 1983 New Hampshire State Plane Feet using ERDAS IMAGINE 9.1 software. Registration accuracy was high as it was visually verified using control points. The published registration accuracy for the IKONOS imagery was 11.8 meter Root Mean Square Error (RMSE). Locational accuracy for SPOT 5 data is published as better than 30 meters (SPOT 2007). Landsat TM accuracy is published at <20 meters 90% (USGS and NASA 2006).

Data Exploration

Data exploration includes any steps taken to better understand the variation of your data and how it relates to the variation on the ground. Initial data exploration requires an understanding of the dynamic ranges of all bands of data (APPENDIX B). To better understand this variation for this study, several additional bands were created for each image. Derivative bands created included NDVI, Tassel-Cap Transformation and Principal Components analysis, as well as simple ratio bands including infrared/red (IR/R), infrared - red (IR-R) and MIR/Red (MIR/R) (only available with SPOT and Landsat data). All derivative bands were re-scaled to the appropriate dynamic ranges of the original component bands, to facilitate an equal match between derived bands and original bands. Each image was restacked to include the newly-created bands.

Although the extra bands may provide insight, they may not be useful or may be redundant for classification purposes. Such extra bands may actually decrease classification accuracy. To avoid degrading the classification, the "best" bands were identified on a per image basis using a divergence analysis based on training data. Ten to fifteen field visited training points per class were digitally located on each of the three images using the "seed tool" (Leica Geosystems 2005). The seed tool grows areas of interest from a user-defined location based on the spectral similarity of neighboring pixels (Leica Geosystems 2005). By plotting the spectral properties of these training areas the user can visualize the separability or usefulness of each band. In addition to a visual examination of the bands, a statistical analysis was performed using the Jeffries-Matusita Divergence Analysis (Leica Geosystems 2005, Bruzzone et al. 1995). This analysis determines the best bands to use based on the user's input of desired bands (e.g., the user can parameterize the analysis to output the five most important bands) and which bands depict the most spectral variation. Based on both the statistical and visual inspection, the least useful bands were removed from the images.

Destriping

Launched in 1984, Landsat 5 TM is the longest running satellite imagery program currently in existence (USGS and NASA 2006). However, Landsat TM imagery appears increasingly striped due to satellite sensor miscalibrations. Readily apparent in bands 2 (green) and 3 (red) of the Landsat TM image acquired for this study, the periodic noise has the potential to affect classification accuracies. To minimize the striping, the Landsat TM image was processed within the Spectral Workstation in IMAGINE, traditionally used with hyperspectral data. Within the Workstation, a Maximum Noise Fraction Analysis (MNF) tool allows for the identification and rectification of striped layers, either automatically or manually. Filters or averaging substitutions can then be applied to the striped areas (Leica Geosystems 2005). Using the MNF tool, striped bands were identified and noise values replaced with the mean of all data. The destriped layers replaced the original bands of the image (Figure 10).

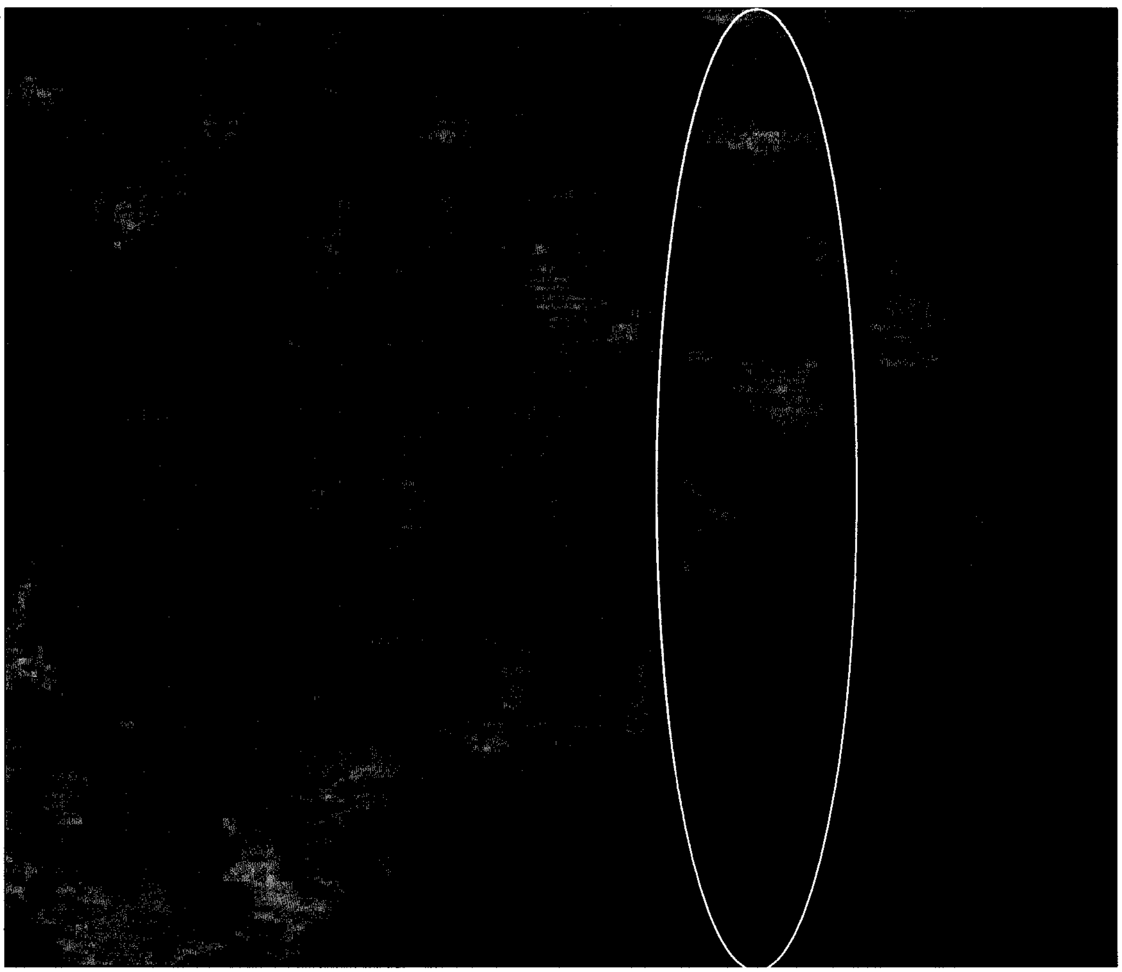


Figure 10. A "swipe" of a portion of the striped Landsat TM image (right) compared with the destriped Landsat TM image (left). The area of contrast between the images is indicated by the white circle. Both images are displayed as NIR, Red and Green through R,G,B channels.

Segmentation

Segmentation and classification analysis were performed using Definiens Professional (v.5) software. Each image was segmented separately and, due to the varying resolution of the images, segmentation parameters were unrelated between images. Definiens Professional uses color and shape parameters (Figure 11) to control the boundaries of segments, also known as image objects (Definiens AG 2005). The weighting of color and shape in the segmentation analysis is based on a sliding scale of 0 to 10 (e.g., 9/10 of the segmentation is based on pixel color and 1/10 is based on resulting segment shape). The shape parameter is further partitioned into smoothness and compactness, also on a sliding scale of 0 to 10. For example, segmentation may be 90% based on color, but 90% of the shape parameter is based on smoothness (Definiens AG 2005).

The Definiens scale number is arguably the most important segmentation parameter as it determines the mean size of the image objects. An arbitrary number, Definiens scale settings are dependent upon the imagery resolution. Thus, a Definiens scale of 9 in 4-meter data will result in different mean image object sizes than a Definiens scale of 9 in 30-meter data. Image object size is roughly equivalent to the desired level of classification (e.g., landscape scale versus individual tree scale).

Since this research focused heavily on determining which scales yielded the most accurate classification results, a variety of Definiens scale parameters were experimented with for each image. As part of the trial and error basis, all results were visually inspected for appropriateness before beginning the classification stage of the analysis. That is, various Definiens scales were examined to determine what resulted in the best segmentation size and placement for study scale (e.g., 30-meter stands). For instance, over 30 separate segmentations were run, each using a different scale parameter, on the IKONOS imagery. Those that resulted in image objects close to the size of 30-meter stands, 2-acre stands or individual tree scales were selected for further classification. The same process occurred for each image. Essentially, segmentation is an iterative process that requires a variety of trials to obtain satisfactory delineation of the image objects that would be similar to how a manual photointerpretation would delineate those objects.

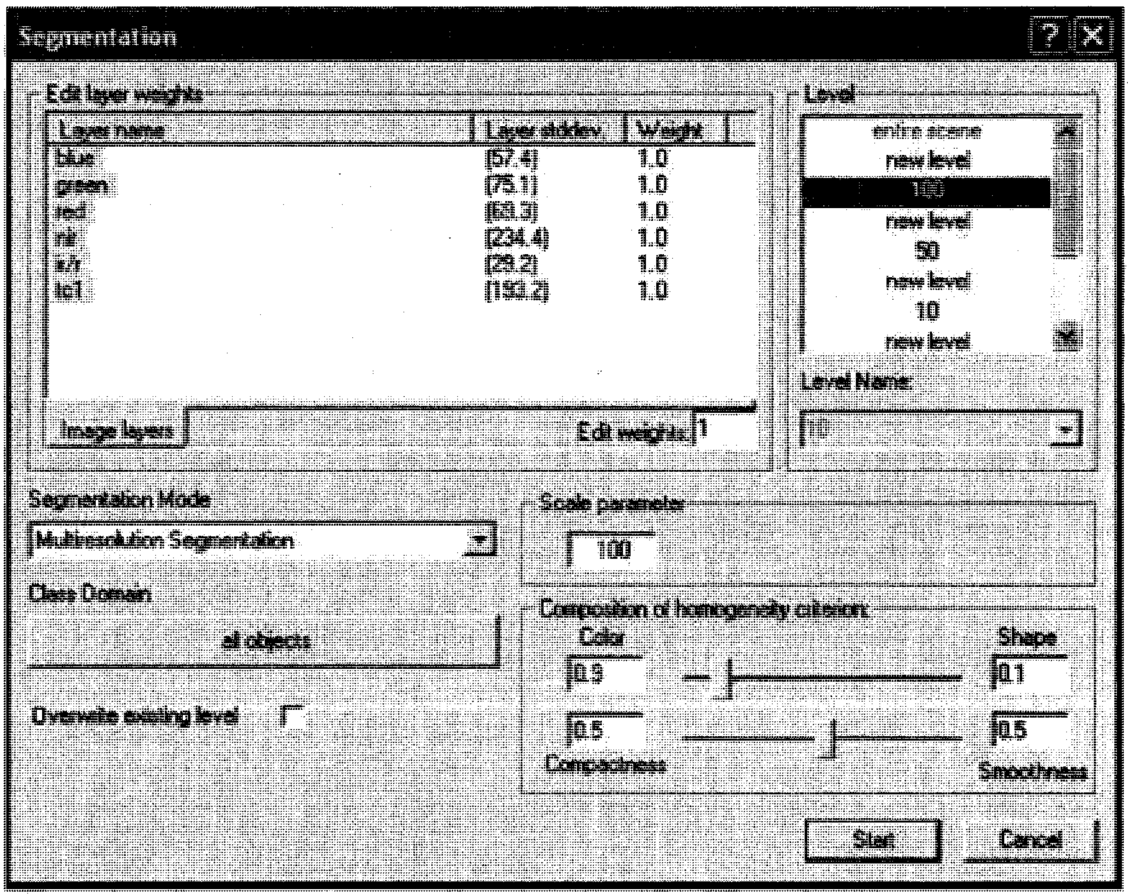


Figure 11. Dialogue box illustrating segmentation parameters in the Definiens software.

Classification

Classification was completed using the sample editor and nearest neighbor sampling application within the Definiens Professional software. Image object samples were selected based on segment size and spatial agreement with the Areas of Interest training data generated in the ERDAS IMAGINE software. Recall that these samples for each class are based on field verified training areas.

Both the mean spectral values and the standard deviation between the class image object samples were used in the classification of unknown image objects. Following the selection of image objects as samples, the classification was set to run with class related features, meaning hierarchical relationships of classes were accounted for during the classification. To better facilitate this feature, coniferous, deciduous, mixed and non-forest species classes training data were grouped and used to classify the image first. Then, a more specific classification was completed to filter the general classes into the study classes (Figure 12).

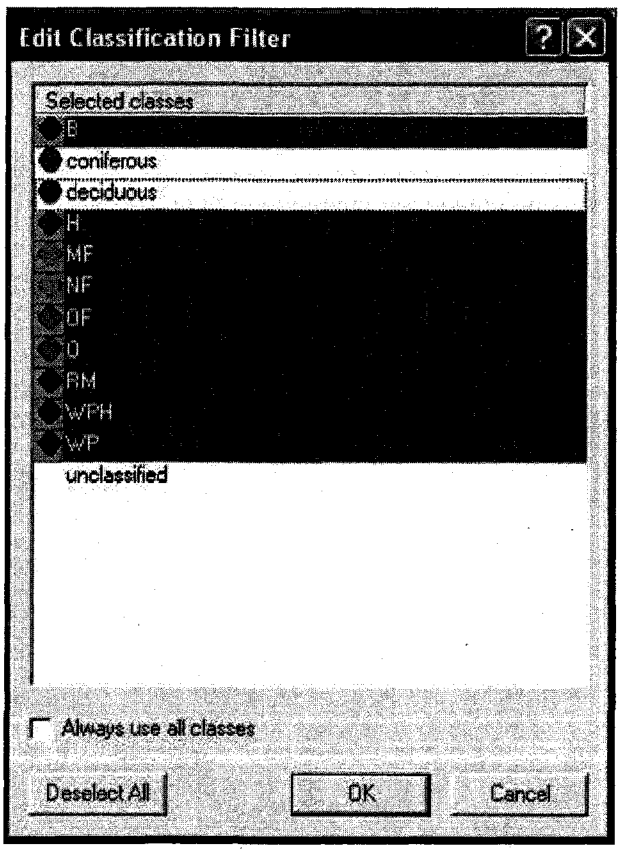


Figure 12. Dialogue box indicating classes used in filtering step of classification in Definiens software.

Number of cycles (iterations) was altered between image classifications to test the effect of iterations upon classification accuracy (Figure 13). Classification results were exported to shapefiles to better facilitate accuracy assessment.

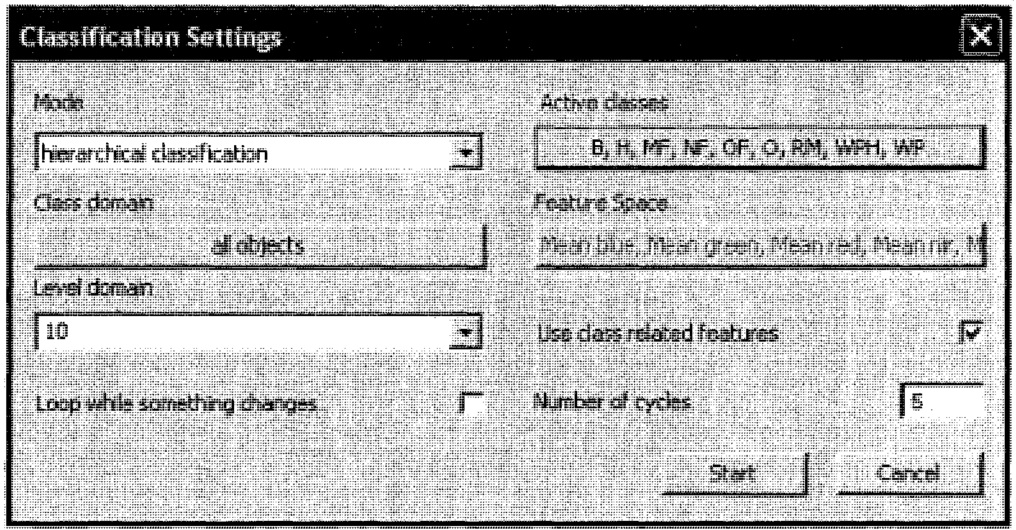


Figure 13. Dialogue box indicating classification parameters in Definiens software.

Accuracy Assessment

An error matrix accuracy assessment was completed for each map, allowing for identification of commissions and omissions. The field-sampled reference points for each class were used to generate the error matrix comparison for IKONOS and SPOT classifications. Randomly generated sample points (50 per class) were extracted from the pre-existing 2-acre scale coverage to test maps created with the Landsat TM imagery. The overall, user's and producer's accuracies were reported (Story and Congalton 1986) for each error matrix. In addition to an error matrix, the K-hat value was calculated and reported to account for chance agreement between the map and reference data (Congalton and Mead 1983). A K-hat value ranges from 0 to 1, with 0 indicating a random assignment of classes and 1 indicating total agreement of classes. A Z-score, calculated along with a K-hat value, allows for between-matrix comparisons. A Z-score of greater than 1.96 (at a 95% confidence interval) indicates significance between two matrices (Congalton 1991, Lennartz 2004). See Figure 14 for a flow chart detailing the classification and accuracy assessment process.

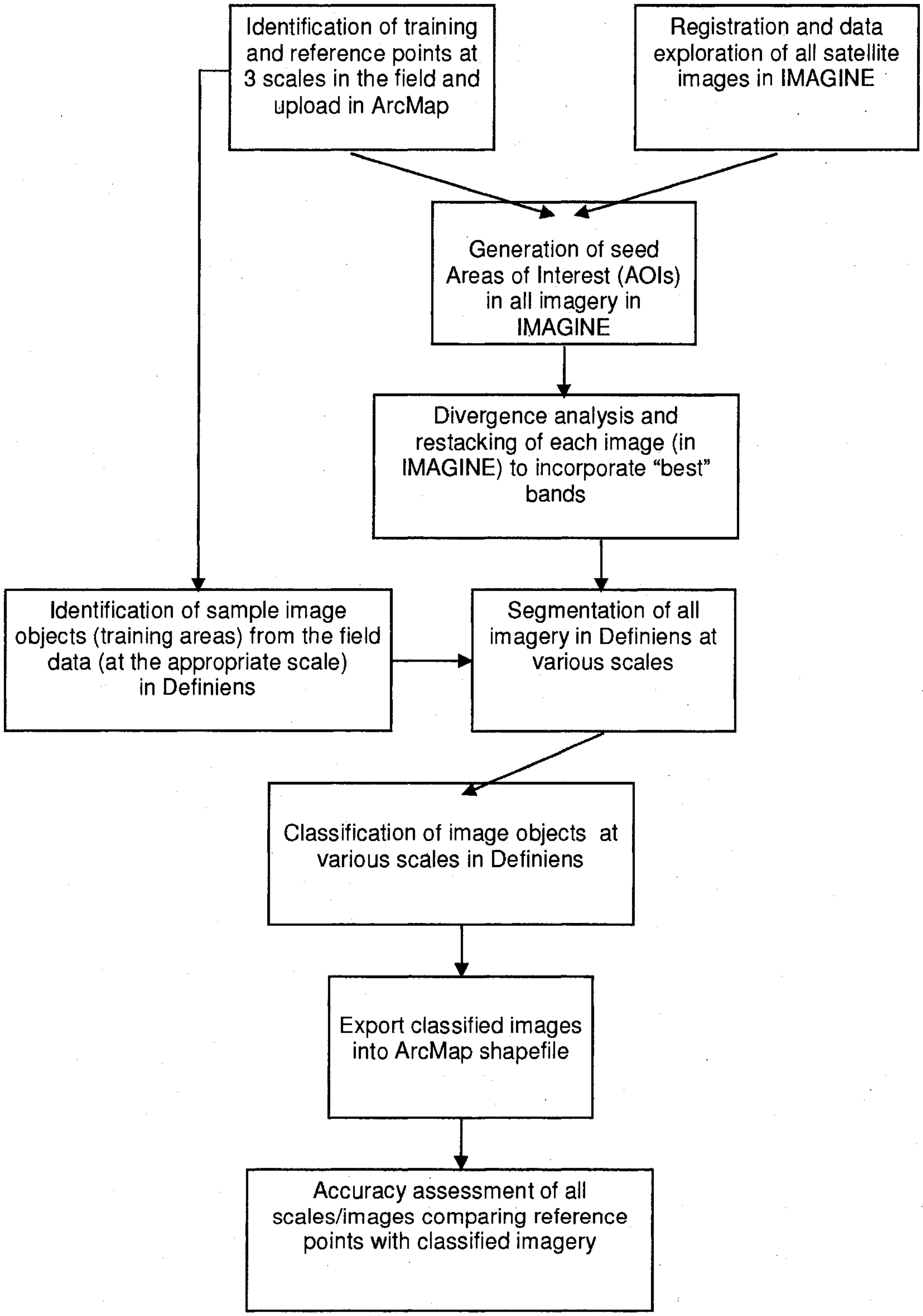


Figure 14. Flowchart illustrating the steps and software platforms incorporated into this classification study.

CHAPTER III

RESULTS

This section was written to provide the reader with an understanding of the overall results and to evaluate the success of the method's components. Understanding what components worked well and what components need to be altered, replaced or omitted is key in advancing the improvement of methodologies.

Classification Scheme Complications

Preliminary field work revealed a need to include two other stand types in the classification definitions. Although not prevalent enough to merit unique categories, the few stands composed primarily of sugar maple and other conifer species were integrated with the red maple and white pine classes, respectively. The decision was made not to create an "other" category to incorporate these species, as there were grossly inadequate numbers of each stand type to achieve the recommended 50 samples per class.

Field Data

A total of 438 reference points at the 30-meter scale were collected using a handheld Garmin GPS unit (Figure 15). An overlay of the uploaded GPS points onto 2005 1-foot aerial photography combined with ground knowledge confirmed the positional accuracy of the GPS points to homogenously comprised 30-meter x 30-meter sample area.

Each class had between 25 and 68 reference points, with the majority of the classes (except beech and other forest) having between 47 and 68 reference points. The numbers of reference points in the two classes with the lowest amount of points were limited by the scarcity of class type as well as limited accessibility.

A total of 450 reference points at the 2-acre scale were generated from the field verified, pre-existing, 2-acre reference map (Figure 16).

A total of 350 individual tree reference points were collected for the individual tree scale, resulting in exactly 50 individual tree samples per category. As individual trees can only be single species, several categories (e.g. White Pine/Hemlock) utilized in other reference scale datasets were excluded.

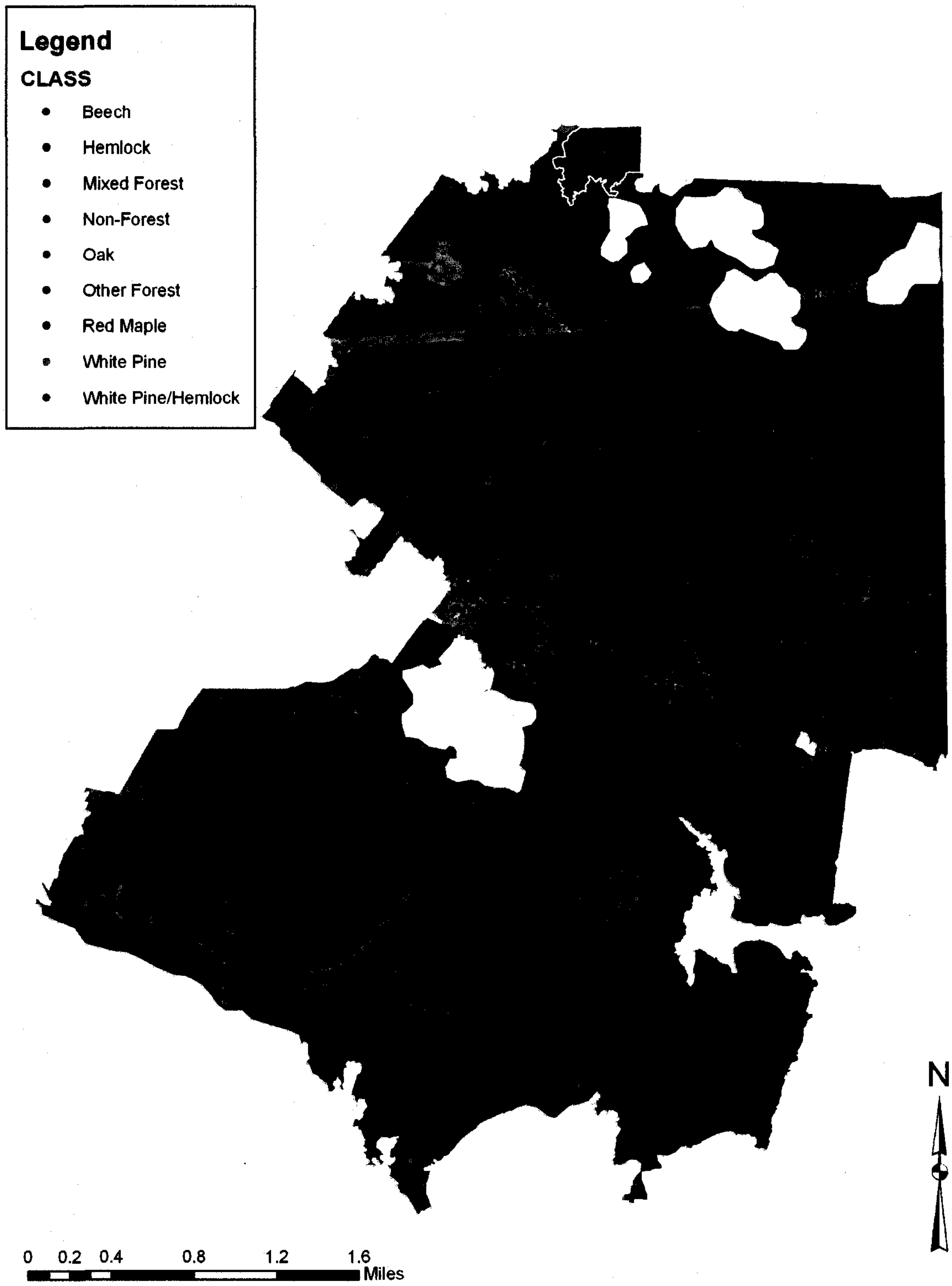


Figure 15. Overlay of 30-meter GPS-located field reference points within the study area overlaid on the SPOT image (displayed as NIR, Red and Green through R,G,B channels).

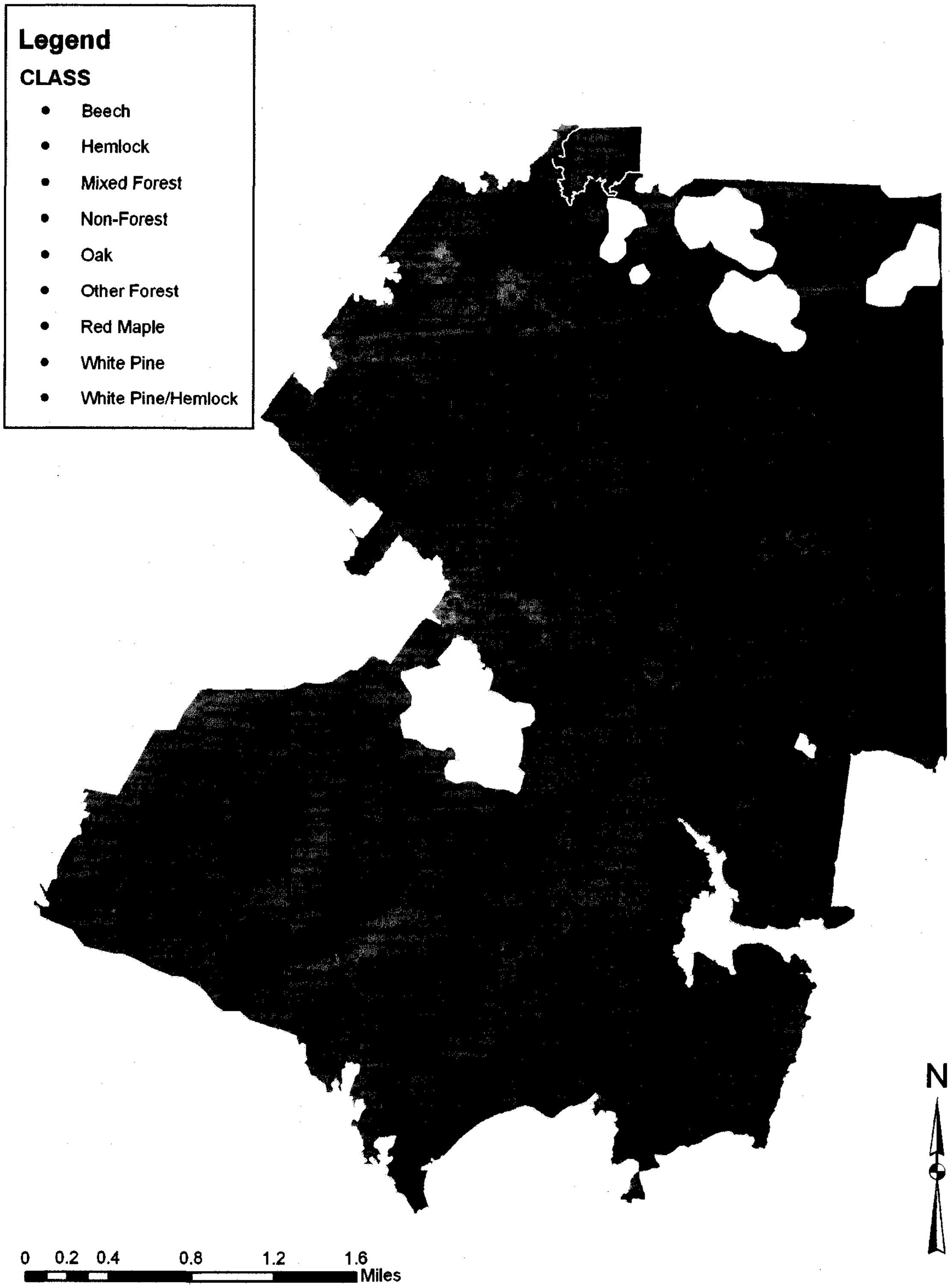


Figure 16. Overlay of 2-acre reference points selected with stratified random sampling of pre-existing reference data overlaid on the Landsat TM image (displayed as NIR, Red and Green through R,G,B channels).

Divergence Analysis

A visual inspection of each images' divergence analysis indicated which bands contained the most spectral variation (Figures 17-19). The divergence analyses were generated by Area of Interest (AOIs) training areas. Both a consideration of the Jefferies-Matusita analyses and the divergence analyses resulted in a reduction of each images' bands (including some derived bands). Based on where there was agreement between the two analyses, the best bands were retained, and the remainder discarded. The best bands (see Figures 17-19) to use in land cover classification with Landsat TM imagery were: blue, green, red, NIR, MIR, MIR II, IR/R and Tasselcap 1. The best bands for use with SPOT imagery were: green, red, NIR, MIR and IR/R.

The best bands for use with IKONOS imagery were: blue, green, red, NIR, IR/R and Tasselcap 1. In each case, redundancy existed between NDVI and IR/R bands. To reduce confusion, IR/R (not NDVI) was selected for use in all imagery. It is important to explain the variation between the numbers of bands selected. Further, in an effort to maintain consistency and the intrinsic value of the imagery, original bands were maintained and derivative bands common to all images were selected, with the exception of SPOT. SPOT data had the least number of bands, as it did not have the required wavelengths to generate a Tassel-Cap band.

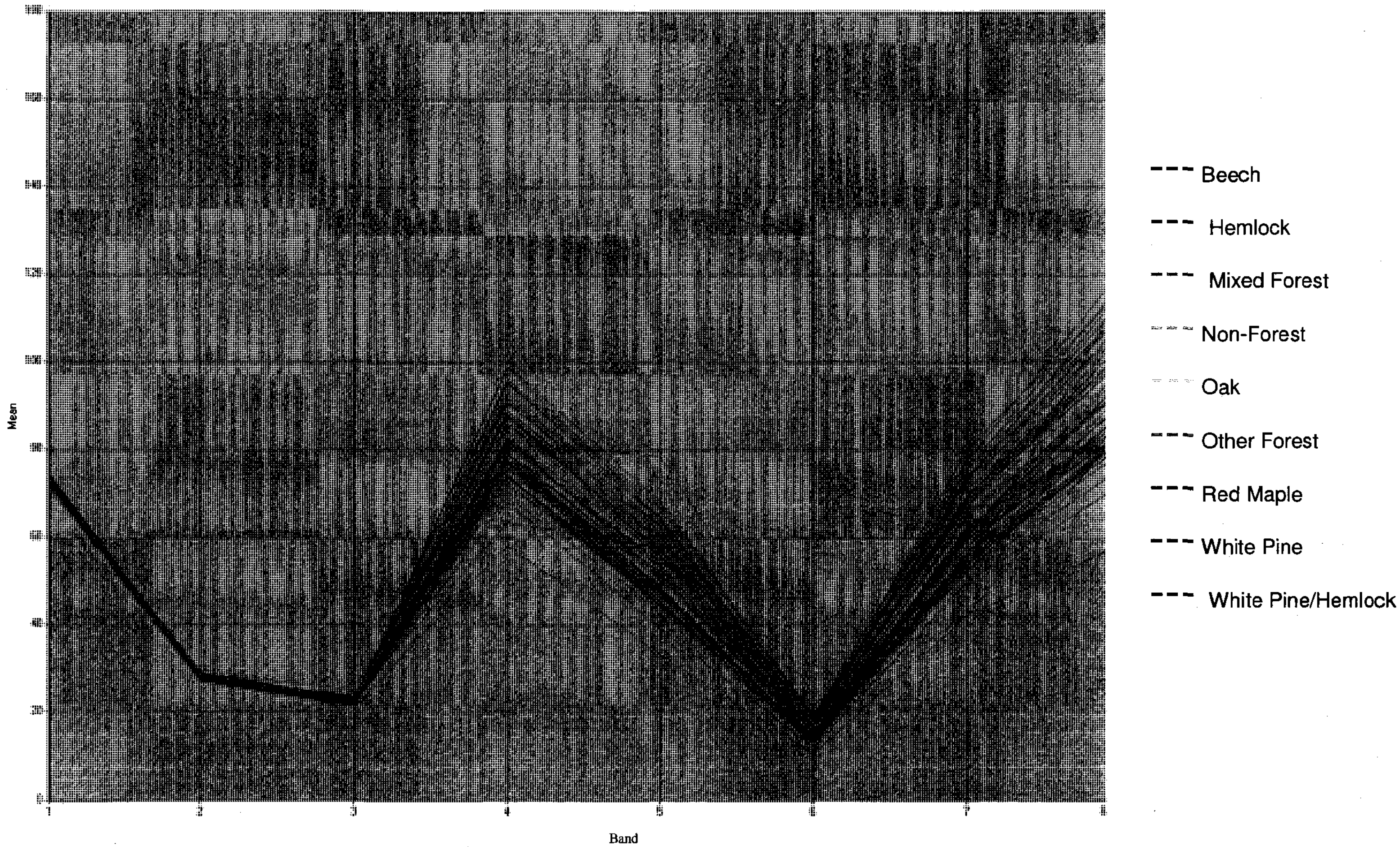


Figure 17. Spectral Pattern/Divergence Analysis of selected bands for Landsat TM imagery using field verified training points in ERDAS IMAGINE software. Bands 1-8 are: blue, green, red, NIR, MIR, MIR II, IR/R and Tasselcap 1.

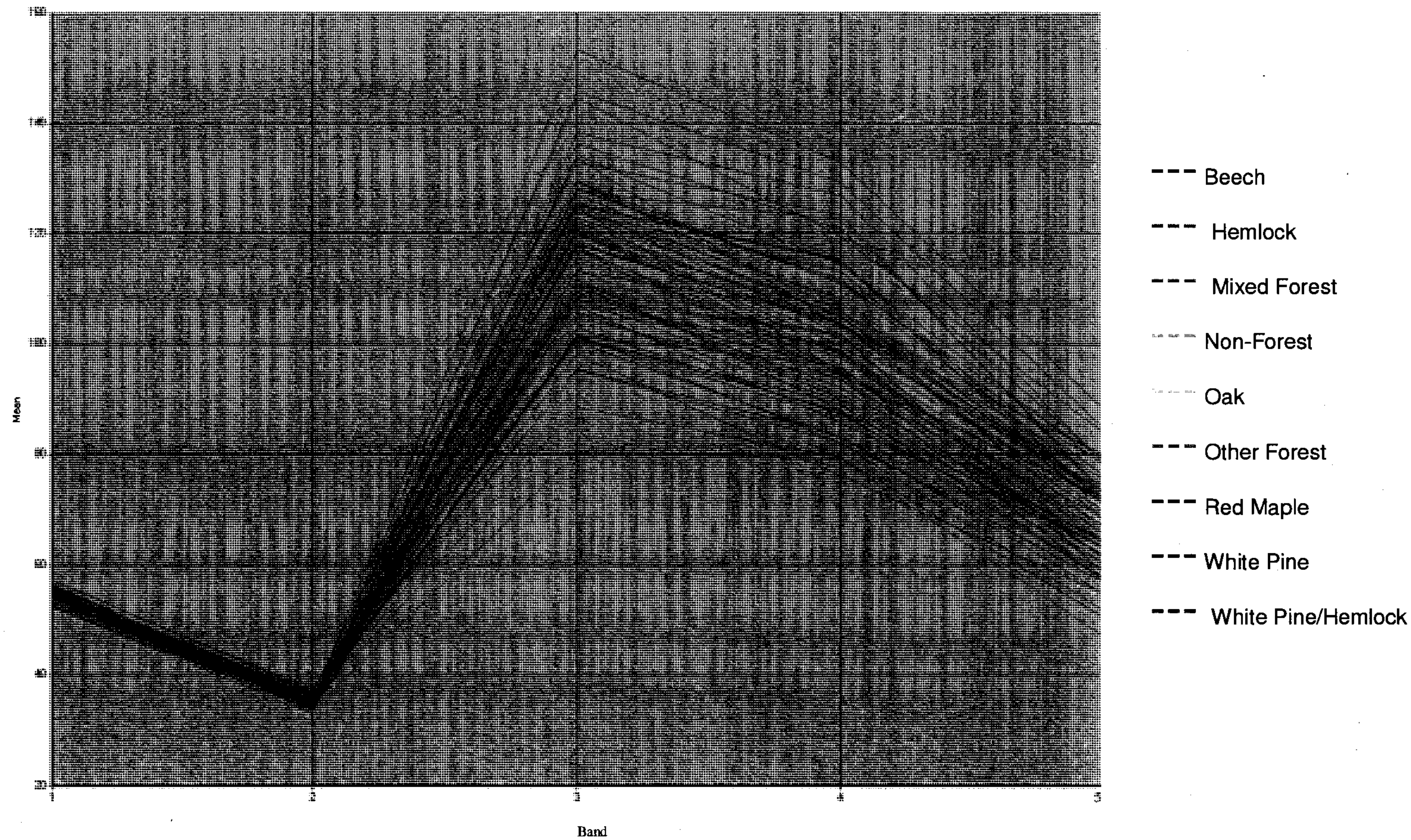


Figure 18. Spectral Pattern/Divergence Analysis of selected bands for SPOT imagery using field verified training points in ERDAS IMAGINE software. Bands 1-5 are green, red, NIR, MIR and IR/R.

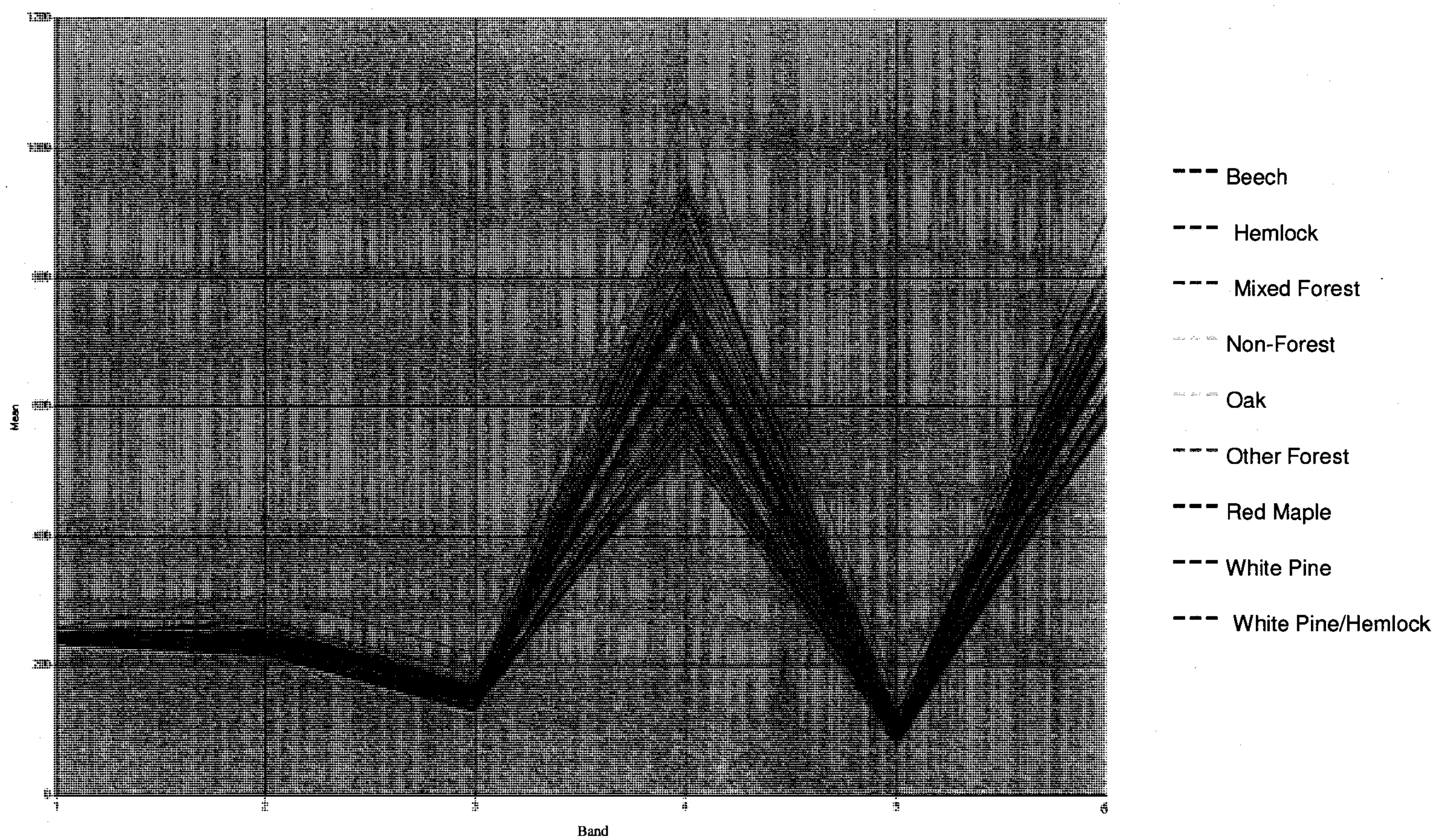


Figure 19. Spectral Pattern/Divergence Analysis of selected bands for IKONOS imagery using field verified training points in ERDAS IMAGINE software. Bands 1-6 are: blue, green, red, NIR, IR/R and Tasselcap 1.

In the Landsat TM image, the greatest spectral variation among species was apparent in the NIR and Tasselcap bands; Bands 4 and 8, respectively (Figure 17). The greatest spectral variation for the SPOT imagery was shown in the NIR and MIR bands; bands 3 and 4, respectively (Figure 18). The NIR and Tassel-Cap bands, bands 4 and 6, respectively, exhibited the greatest spectral variation for the IKONOS imagery (Figure 19). These bands showing the greatest spectral variation are most important in distinguishing between the majority of species.

Segmentation Parameters

A variety of Definiens segmentation parameter combinations were investigated with very little difference between object delineations, excluding the scale parameter. Based on observed iterations, since color and shape parameter change had little effect upon segmentation, a single set of color and shape parameters was selected for use between imagery. This standardization served to reduce the variation contributing to results (caused by testing multiple parameters), making it easier to identify the optimal imagery, and also streamlined the process. The segmentation parameters for each image were set at those that consistently produced the best results: color = 0.9 and smoothness = 0.5. Shape was set at 0.1 and compactness was set at 0.5.

The average segment area and actual object delineation, however, was dependent upon imagery type and scale input. At the best classification accuracy and therefore, the best scale parameter for use, the Landsat TM imagery yielded an overall average segment area of 6.03 acres at a Definiens parameter scale of 5. At the same Definiens parameter scale, the SPOT imagery had an average overall segment area of 0.74 acres. At Definiens parameter scales of both 10 and 15, the SPOT imagery yielded a segment area of 8.76 acres. At a Definiens parameter scale of 10, the IKONOS imagery segment areas averaged 0.04 acres. As the Definiens scale parameter is very dependent upon imagery resolution, the same scale parameter used to segment SPOT and Landsat TM imagery resulted in different segment areas.

While the above segmentation results appeared to be correct following a visual inspection, the segmentation parameters on the finest resolution imagery (IKONOS) were clearly unable to accurately segment individual trees (Figure 20). No scale parameter was able to accurately delineate the individual trees' canopies, as shadowed areas and overlapping tree branches created a large amount of spectral confusion. However, canopy delineation through segmentation was satisfactorily achieved using the 2005, full color 1-foot, digital aerial photography and can be seen as the white lines on Figure 20. A preliminary statistical analysis using Student's t-test demonstrated a significant statistical difference between the segmented IKONOS image and the segmented aerial photography areas ($p < 0.0001$). This p-value confirmed the visual inspection, in which the segmented IKONOS imagery neither matched canopy boundary nor individual tree location. Given there was no apparent correlation between individual trees and the segmented IKONOS imagery, an attempt at classifying the segments seemed imprudent and was therefore abandoned.

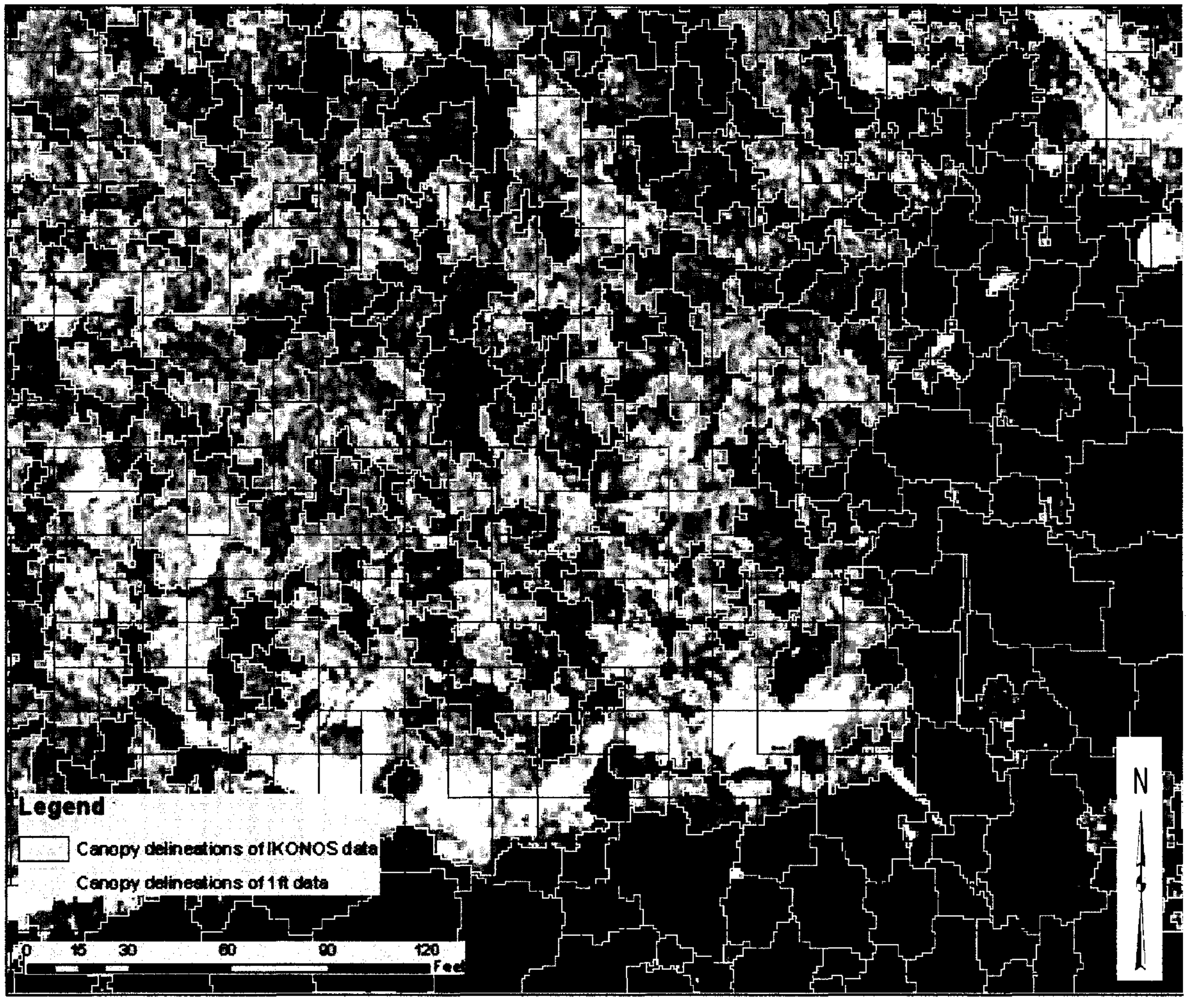


Figure 20. Comparison of segmentation capabilities between the IKONOS imagery data and the 2005 1-foot aerial photography (seen in background). Red outlines indicate the canopy delineation/segmentation in the IKONOS imagery, which clearly does not register correctly with the individual tree canopies. It is clear the IKONOS segmentation does not correspond to actual canopy location.

Classification Results

The best accuracy results of all classification trials are presented for the IKONOS, SPOT and Landsat TM data in both summary matrix (Table 2) and error matrix forms (Tables 3-8). Results of two reference scales (3- meter and 2-acre stands) and at least one classification per image are reported here. Overall five classification trials were chosen to represent the best classification accuracy results based on segmentation and classification parameters. The error matrix and Kappa analysis results (Table 2) indicate that the Landsat TM imagery, with a reference size of 2 acres, yielded the highest overall accuracy (34.2%) and highest K-hat value (0.26) and was also the only imagery with better than random results. Although the Landsat TM results are not considered high, they are higher than the "next best" results: SPOT imagery with a reference size of 30 m with an overall accuracy of 21.8% and a K-hat value of 0.12. The best classification trial of the IKONOS imagery had the lowest overall accuracy (21.1%) and the lowest K-hat value (.11).

The Z-score results indicate that the IKONOS and SPOT classifications are not significantly different (Table 2), nor are any of the trials of the SPOT classifications. In fact, two SPOT imagery trials with differing scale parameters, yielded identical results (Tables 6 and 7). However, the Landsat TM imagery classification is significantly different when compared to both the SPOT and IKONOS classifications. As is indicated by "unclass" in some error matrices (Tables 3-7), the software was often unable to assign a class to a segmented object, resulting in unclassified image objects. Non-forest stands consistently yielded higher producer's accuracies than most other stands. The corresponding classification maps for the error matrices are also presented (Figures 21-24). No pattern was observed in the distribution of the classification schemes that would indicate spatial autocorrelation.

Table 2. Summary table of error matrix and K-hat values for the "best" landscape classifications. K-hat value = 0 indicates no better than random class assignment. K-hat value = 1 indicates perfect classification. * indicates a significance at the 95% confidence interval level.

Comparison of Classification Accuracies between Images and Segmentation Scales							
Imagery	Definiens Segmentation Scale	Average Segment Size (Acres)	Reference Sample Size	Overall Accuracy (%)	K-hat Value	Variance	Z-Score
IKONOS	10	.04	30m	21.1	0.113	0.0005013	5.045
SPOT	5	.74	30m	21.8	0.118	0.0005038	5.252
SPOT	10	8.76	2 acres	21.4	0.116	0.0006839	4.418
SPOT	15	8.76	2 acres	21.4	0.116	0.0006839	4.418
Landsat TM	5	8.76	2 acres	34.2	0.260*	0.0006366	10.309

Table 3. Summary table of image classification comparisons using the Kappa Test Statistic approach. A Z score ≥1.96 indicates a significant difference between maps (at the 95% confidence interval level). Bold comparisons denote significantly different pairs.

Imagery and Scale (in parentheses) Comparison		Z-Score
IKONOS 10	**LANDSAT TM 5**	**-4.3630**
IKONOS 10	SPOT 5	-0.1555
IKONOS 10	SPOT 10	-0.0753
IKONOS 10	SPOT 15	-0.0753
LANDSAT TM 5	**SPOT 5**	**4.2121**
LANDSAT TM 5	**SPOT 10**	**3.9787**
LANDSAT TM 5	**SPOT 15**	**3.9787**
SPOT 5	SPOT 10	0.0679
SPOT 5	SPOT 15	0.0679
SPOT 10	SPOT 15	0

Table 4. Error matrix accuracy assessment results (overall, producer's and user's accuracies) for the classification of IKONOS imagery at a segmentation scale of 10, using 30-meter reference data.

Error Matrix

Map Classification: IKONOS Segmentation Scale of 10

		Reference Data									
		B	**H**	**MF**	**NF**	**O**	**OF**	**RM**	**WP**	**WPH**	**Totals**
Map Data	**B**	10	8	4	4	7	5	6	1	3	48
	H	6	11	7	6	5	15	8	10	3	71
	MF	7	3	8	1	2	3	6	7	3	40
	NF	0	2	0	18	0	2	1	2	0	25
	O	3	6	7	5	11	9	9	5	1	56
	OF	5	4	7	0	9	12	7	1	1	46
	RM	0	4	7	2	6	11	7	6	2	45
	WP	1	6	9	7	1	4	9	9	6	52
	WPH	2	7	4	4	10	7	5	8	6	53
	Totals	34	51	53	47	51	68	58	49	25	436
			1unclass		1unclass						

Overall Accuracy = **21.1%**

	Producer's Accuracy	**User's Accuracy**
B = beech	29.4%	20.8%
H = eastern hemlock	21.6%	15.5%
MF = mixed forest	15.1%	20.0%
NF = non-forest	38.3%	72.0%
O = oak	21.6%	19.6%
OF = other forest	17.6%	26.1%
RM = red maple	12.1%	15.6%
WP = eastern white pine	18.4%	17.3%
WPH = eastern white pine/eastern hemlock	24.0%	11.3%

Table 5. Error matrix accuracy assessment results (overall, producer's and user's accuracies) for the classification of SPOT imagery at a segmentation scale of 5, using 30-meter reference data.

Error Matrix

Map Classification: SPOT Segmentation Scale of 5

		Reference Data									
		B	**H**	**MF**	**NF**	**O**	**OF**	**RM**	**WP**	**WPH**	**Totals**
Map Data	**B**	15	4	3	7	6	16	5	2	0	58
	H	4	8	15	4	5	8	8	8	4	64
	MF	2	4	7	2	2	9	9	3	4	42
	NF	0	2	0	15	1	0	3	1	1	23
	O	3	6	5	4	11	10	5	7	5	56
	OF	6	7	4	5	13	11	8	0	4	58
	RM	3	10	9	1	7	7	7	2	4	50
	WP	0	6	7	7	3	4	1	21	2	51
	WPH	1	5	3	2	7	3	12	5	1	39
	Totals	34	52	53	47	55	68	58	49	25	441

1 unclass

Overall Accuracy = **21.8%**

	Producer's Accuracy	**User's Accuracy**
B = beech	44.1%	25.9%
H = eastern hemlock	15.4%	12.5%
MF = mixed forest	13.2%	16.7%
NF = non-forest	31.9%	65.2%
O = oak	20.0%	19.6%
OF = other forest	16.2%	19.0%
RM = red maple	12.1%	14.0%
WP = eastern white pine	42.9%	41.2%
WPH = eastern white pine/eastern hemlock	4.0%	2.6%

Table 6. Error matrix accuracy assessment results (overall, producer's and user's accuracies) for the classification of SPOT imagery at a segmentation scale of 10, using 2-acre reference points data and 6 classification iterations.

Error Matrix

Map Classification: SPOT Segmentation Scale of 10

		Reference Data									
		B	H	MF	NF	O	OF	RM	WP	WPH	Totals
Map Data	B	7	4	0	0	0	2	3	0	0	16
	H	0	0	0	0	0	0	0	1	1	2
	MF	0	1	0	0	0	0	0	0	0	1
	NF	10	27	16	38	18	9	29	25	28	200
	O	0	0	1	0	0	0	0	0	1	2
	OF	33	16	29	7	27	38	9	10	14	183
	RM	0	0	2	0	1	0	2	1	0	6
	WP	0	2	1	2	0	1	1	9	4	20
	WPH	0	0	1	2	3	0	5	2	1	14
	Totals	50	50	50	49	49	50	49	48	49	444
					1unclass	1unclass		1unclass	2unclass	1unlass	

Overall Accuracy = 21.4%

	Producer's Accuracy	**User's Accuracy**
B = beech	14.0%	43.8%
H = eastern hemlock	0.0%	0.0%
MF = mixed forest	0.0%	0.0%
NF = non-forest	77.6%	19.0%
O = oak	0.0%	0.0%
OF = other forest	76.0%	20.8%
RM = red maple	4.1%	33.3%
WP = eastern white pine	18.8%	45.0%
WPH = eastern white pine/eastern hemlock	2.0%	7.1%

Table 7. Error matrix accuracy assessment results (overall, producer's and user's accuracies) for the classification of SPOT imagery at a segmentation scale of 15, using 2-acre reference points data and 6 classification iterations.

Error Matrix

Map Classification: SPOT Segmentation Scale of 15

		Reference Data									
		B	**H**	**MF**	**NF**	**O**	**OF**	**RM**	**WP**	**WPH**	**Totals**
Map Data	**B**	7	4	0	0	0	2	3	0	0	16
	H	0	0	0	0	0	0	0	1	1	2
	MF	0	1	0	0	0	0	0	0	0	1
	NF	10	27	16	38	18	9	29	25	28	200
	O	0	0	1	0	0	0	0	0	1	2
	OF	33	16	29	7	27	38	9	10	14	183
	RM	0	0	2	0	1	0	2	1	0	6
	WP	0	2	1	2	0	1	1	9	4	20
	WPH	0	0	1	2	3	0	5	2	1	14
	Totals	50	50	50	49	49	50	49	48	49	444
					1unclass	1unclass		1unclass	2unclass	1unclass	

Overall Accuracy = 21.4%

	Producer's Accuracy	**User's Accuracy**
B = beech	14.0%	43.8%
H = eastern hemlock	0.0%	0.0%
MF = mixed forest	0.0%	0.0%
NF = non-forest	77.6%	19.0%
O = oak	0.0%	0.0%
OF = other forest	76.0%	20.8%
RM = red maple	4.1%	33.3%
WP = eastern white pine	18.8%	45.0%
WPH = eastern white pine/eastern hemlock	2.0%	7.1%

Table 8. Error matrix accuracy assessment results (overall, producer's and user's accuracies) for the classification of Landsat TM imagery at a segmentation scale of 5, using 2-acre reference points data and 5 classification iterations.

Error Matrix

Map Classification: Landsat Segmentation Scale of 5

		Reference Data									
		B	**H**	**MF**	**NF**	**O**	**OF**	**RM**	**WP**	**WPH**	**Totals**
Map Data	**B**	22	5	7	0	10	7	2	1	1	55
	H	8	21	5	1	5	4	6	6	14	70
	MF	3	0	8	2	6	5	9	6	13	52
	NF	1	2	0	33	4	2	2	1	2	47
	O	6	2	6	1	11	4	4	4	1	39
	OF	4	4	5	5	5	22	3	1	2	51
	RM	1	7	8	5	3	2	10	5	6	47
	WP	0	4	3	1	1	1	7	20	5	42
	WPH	3	5	7	2	5	3	7	6	6	44
	Totals	48	50	49	50	50	50	50	50	50	447
		2unclass		1unclass							

Overall Accuracy = **34.2%**

	Producer's Accuracy	**User's Accuracy**
B = beech	45.8%	40.0%
H = eastern hemlock	42.0%	30.0%
MF = mixed forest	16.3%	15.4%
NF = non-forest	66.0%	70.2%
O = oak	22.0%	28.2%
OF = other forest	44.0%	43.1%
RM = red maple	20.0%	21.3%
WP = eastern white pine	40.0%	47.6%
WPH = eastern white pine/eastern hemlock	12.0%	13.6%

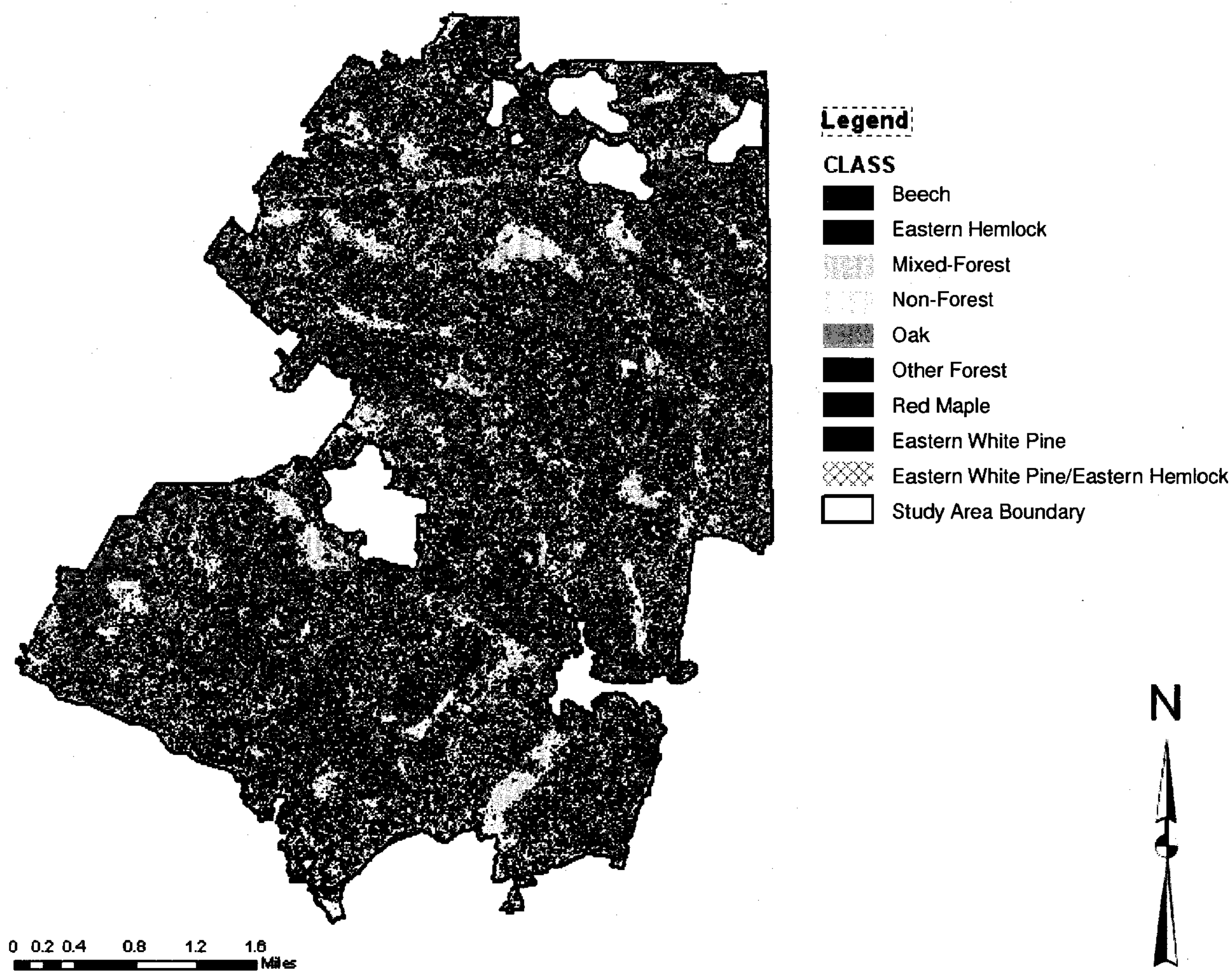


Figure 21. Classified map of IKONOS imagery at a segmentation scale of 10 using 30-meter reference data. Accuracy = 21.1% (see Table 4 for detailed accuracy assessment).

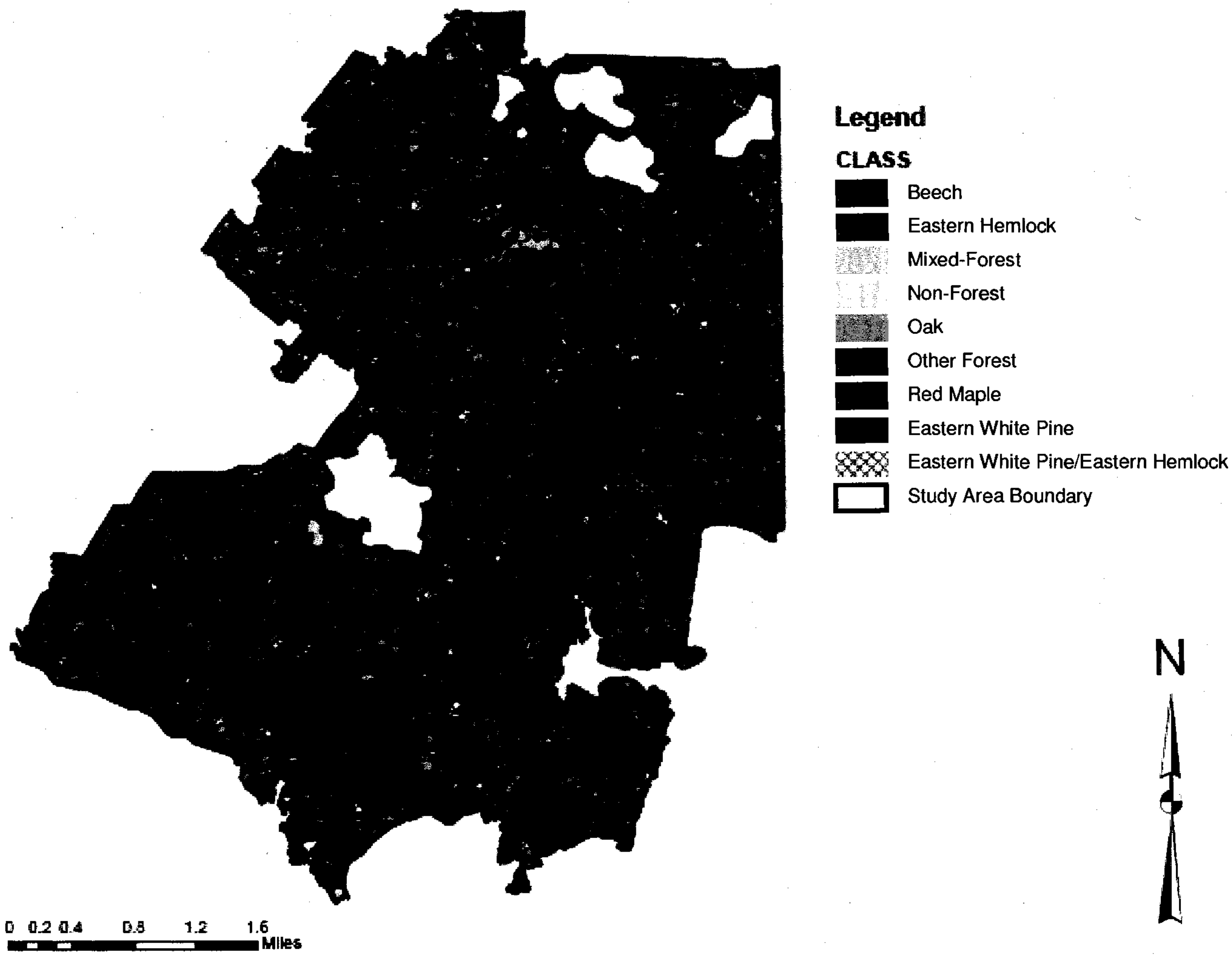


Figure 22. Classified map of SPOT imagery at a segmentation scale of 5 using 30-meter reference data. Accuracy = 21.8% (see Table 5 for detailed accuracy assessment).

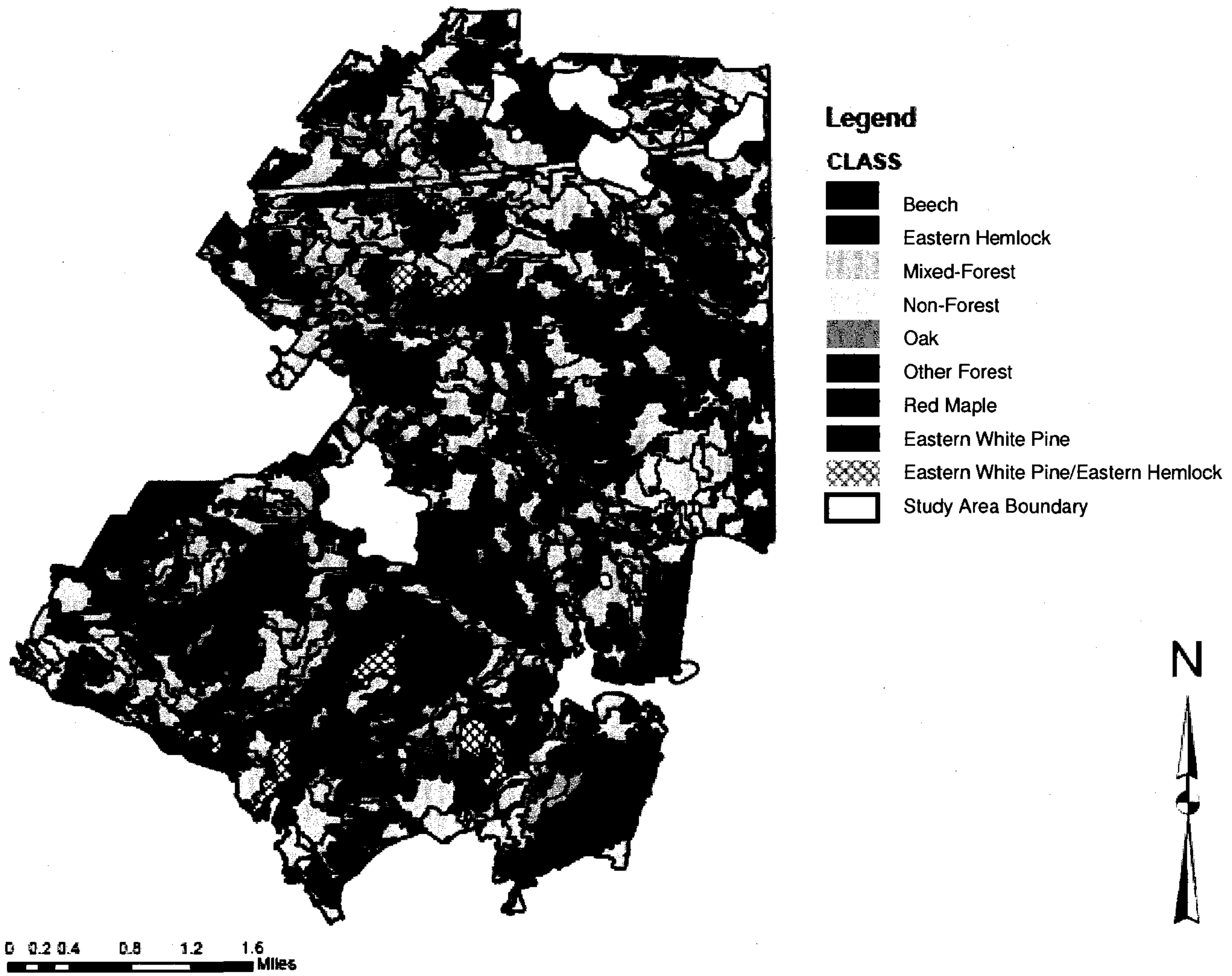


Figure 23. Classified map of SPOT imagery at a segmentation scale of 10 using 2-acre reference data with 6 classification iterations. Accuracy = 21.4% (see Table 6 for detailed accuracy assessment).

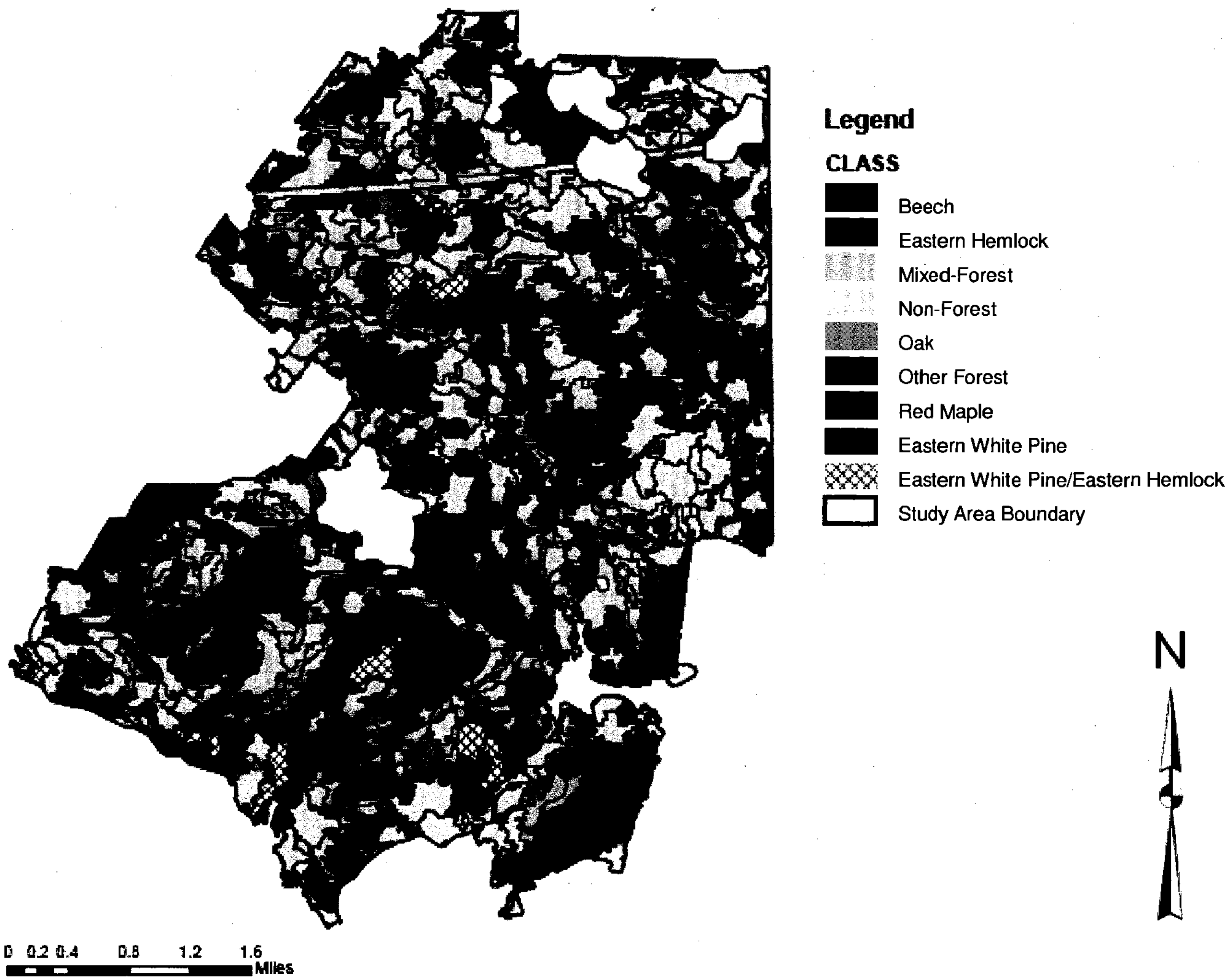


Figure 24. Classified map of SPOT imagery at a segmentation scale of 15 using 2-acre reference data with 6 classification iterations. Accuracy = 21.4% (see Table 7 for detailed accuracy assessment).

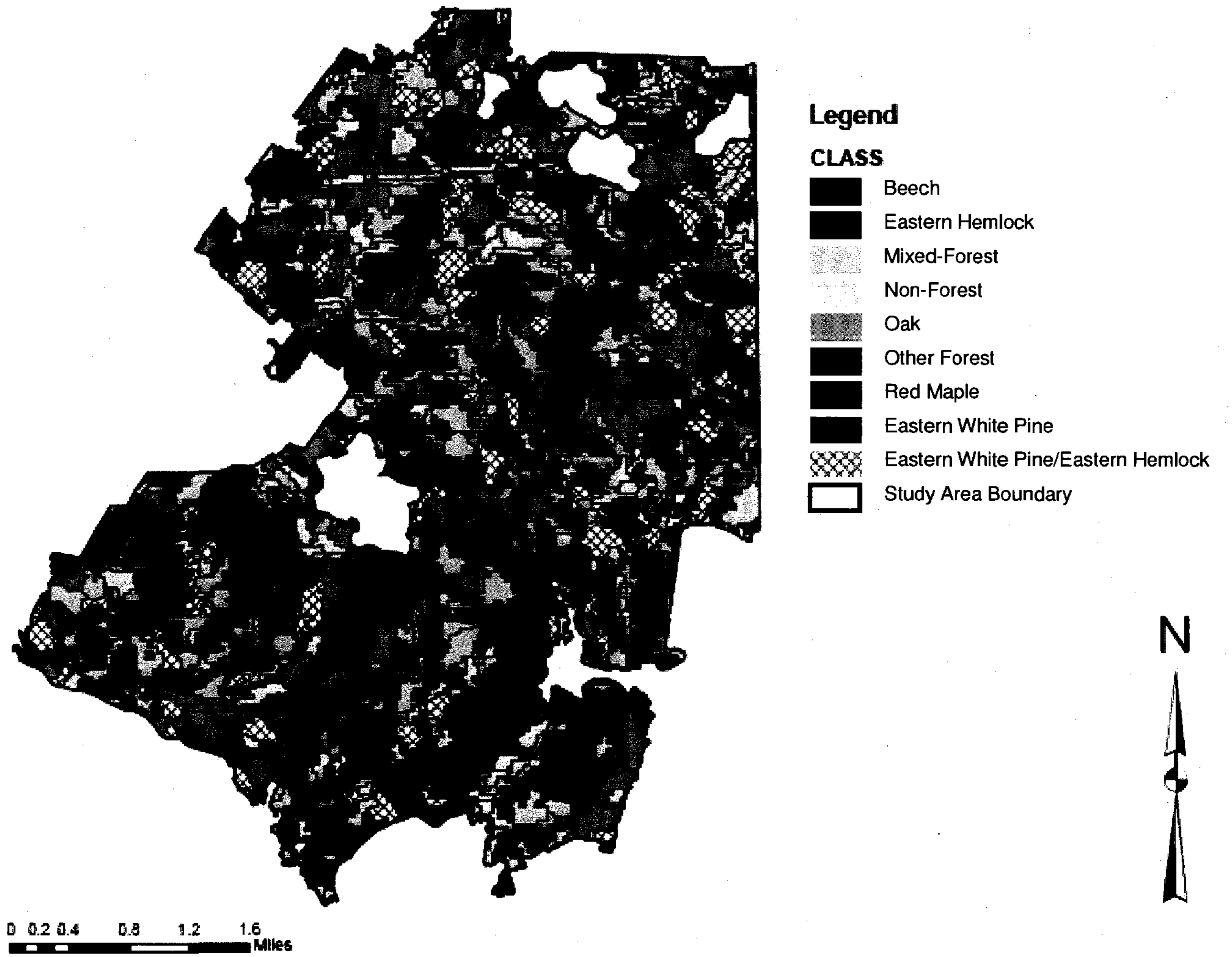


Figure 25. Classified map of Landsat TM imagery at a segmentation scale of 5 using 2-acre reference data with 5 classification iterations. Accuracy = 34.2% (see Table 8 for detailed accuracy assessment).

CHAPTER IV

DISCUSSION AND CONCLUSIONS

Although the maps generated by this research are ultimately unreliable for use in the field due to the low overall accuracies, some distinct conclusions can be reached as a result of this research.

Segmentation and Classification

This study provides insight into the two aspects of object-oriented image classification: segmentation and classification. Visual inspections of the segmentation results verified that the Definiens software performed accurate segment generation, regardless of scale or imagery used (excepting the instance of individual tree classification). That is to say that the actual delineation of image-objects (such as a stand) was performed satisfactorily: object boundaries were placed similarly to how they would be placed by a manual photointerpretation.

Although the first part of the classification process, segmentation, was well-executed, accuracies were low for all images' classification results. These low accuracies would indicate that error lies in the second aspect of object-oriented image classification or when the actual labeling of segmented object occurs. As high segmentation accuracy but low overall classification accuracy occurred with all imagery, it can be assumed the Definiens Professional software sufficiently segments an image, but may not be effective to classify the created segments. More sophisticated algorithms, better suited imagery and/or a different methodology may be required to adequately classify segmented imagery.

Spectral Resolution versus Spatial Resolution

The difficulty of species classification is supported by low accuracies of previous research (both object-oriented and pixel-based) within the same study area (Pugh 1997, Lennartz 2004) and is likely attributed to the level of species detail desired and, in some cases, the increased spectral variation inherent in higher resolution imagery. Increased spatial resolution leads to the detection of shadows and minute shading variations, which increases the apparent stand complexity and makes classification more difficult, as between class spectral confusion is increased. Compounded with the intrinsic structural complexity of New England forests, the increase in spectral variation makes species level classification challenging with broadband satellite sensor data, like that used in this study.

That being said, it is still important to remember the trade-off that exists between spectral resolution and spatial resolution. As the imagery with the lowest spatial resolution, but the highest spectral resolution yielded the best classification accuracies and was the only classification to be significantly better than random, this research would suggest that spectral resolution is more important than spatial resolution when employing object-oriented image classification of forest stands, as it better captures the natural spectral variation within those stands.

A promising source of higher spectral resolution lies with the implementation of hyperspectral imagery in object-oriented classifications. Research from those using higher spectral resolution imagery supports this conclusion. For example, Cochrane (2000) used spectrometer data comprised of 512 wavelength bands between 350 and 1050nm and automated classification to correctly discriminate 11 target tree species 94% of the time. These resulted validate Cochrane's (2000) conclusions that the NIR spectrum captures the most spectral variation. Although Cochrane's (2000) research utilized remote sensing at the leaf scale, it substantiates the hypothesis that hyperspectral vegetative reflectance can accurately be applied to species classification.

Similarly, Clark et al. (2005) were able to use spectrometry to accurately (100%) classify leaves of seven species of trees. Additionally, they achieved, when classifying 1.6-meter, 30 band hyperspectral imagery, a 92% overall accuracy classifying the same seven species. Clark et al.'s (2005) research further indicates that the integration of hyperspectral imagery with object-oriented classification could improve overall accuracy. Interestingly, NIR again proved to be the most valuable wavelength spectrum (Clark et al. 2005).

However, it is important to note that, although the seven separate species stands were classified accurately, individual trees were not. The use of hyperspectral data to delineate and classify individual canopies has been less successful. Delineation of canopy crowns using 1m spatial resolution hyperspectral aerial imagery in Definiens, however, achieved 70% classification accuracies (which varied based on canopy density) in an Australian mixed species forest (Bunting and Lucas 2005). As this thesis suggests, Bunting and Lucas' (2005) research supports the hypothesis that, although hyperspectral data performs well in stand scale classifications, higher spatial resolution imagery is needed to identify individual trees.

This study utilized some of the highest spatial resolution satellite imagery available (IKONOS). Given that the segmentation in this study did not reliably delineate individual tree canopies and manual delineation of canopies using IKONOS data resulted in a 65% overestimation of canopy coverage in the Amazon (Asner et al. 2002), it is likely that currently available satellite imagery resolutions are spatially inadequate to delineate individual trees. Larsen 2007 has also suggested that satellite imagery lacks the spatial resolution necessary to accurately perform detailed land cover classification .

As the currently available satellite imagery spatial resolutions are repeatedly too coarse for individual canopy delineation and the use of higher spectral resolution imagery improves imagery classification, a need for increased spatial *and* spectral resolution onboard satellite sensors is apparent for the classification of individual trees.

Accuracy Assessment

In addition to improving the methodology necessary to attain higher accuracies using object-oriented classification, it is necessary to also improve upon the techniques to correctly assess segmentation accuracy. Object-oriented classification requires an understanding of not only pixel registration, but also an understanding of segment registration, specifically where reference and training points are located within individual segments. Further development is needed to effectively determine the accuracy of segment placement and points within those segments. Currently, it is often necessary to study accuracy on a per-object basis to thoroughly understand the relation between the imagery and the software segment delineation (Yu et al. 2008).

As was the case with this research, field survey plots and reference points rarely match the segmented image objects (Yu et al. 2008). Generating the segments before collecting field data is a possible improvement to the methodology, allowing field observers to locate the center of segments to gather reference and training data. Having segment locations before field work would improve sample quality (i.e., they would be more representative of the segment) and eliminate the possibility of multiple reference points per segment. However, there is no methodology currently in place to determine the accuracy of segment placement (e.g., stand or canopy delineation).

It is also important to note that some of the accuracies in this study appear to be artificially inflated. A comparison between field knowledge and a visual examination of the two SPOT land cover maps (scales 10 and 15, both using 2-

acre reference data) revealed little agreement with actual stand and landscape patterns observed on the ground in the field (i.e., image classification yielded far too much other forest). Although the accuracy analysis was completed in accordance with standard error matrix practices, including stratified random sampling of points, it is likely that the assessment is biased toward abundant areas. The correct classification of the other forest and the non-forest categories likely boosted the overall accuracy assessment.

Sources of Additional Error

In addition to improper assessment technique, low accuracies can also be the result of error accumulated throughout part or the entirety of the land cover classification process. For the purposes of registration between ground data and imagery, continuous and homogenous samples (30-meter x 30-meter) were collected to serve as reference and training data. However, consistently sampling within homogenous areas can result in biased results (Plourde and Congalton 2003). Image pixels may cover several classes, which is not represented by homogenous sampling schemes.

Previously compiled reference data (30-meter and 2-acre) were deemed appropriate for use in this study, to supplement the field data collected, as it provided unequalled wall-to-wall study area coverage. However, the reference data were collected roughly ten years previous to the commencement of this study. While some landscape change likely occurred, much of the study area is within a state park, used primarily for recreation with little to no active forest management. Although development in the area is not prevalent, it could have affected the accuracy of the reference data. Natural succession change (e.g., regeneration of a field into forested area) undoubtedly occurred throughout the study area, but it is questionable as to what magnitude of change is necessary to elicit detection in an accuracy analysis.

Observer bias is a probable source of error in any research situation. However, the use of both existing and newly created reference and training data likely compounded this bias. Although the categories within the classification scheme were identical between the various datasets, there was room for observer opinion to influence category assignment. For example, observer estimations of canopy cover are likely to vary (e.g. 20% oak or 30% oak) between individuals. Although this may not always result in differing classifications, a better defined classification scheme would eliminate much of the ambiguity associated with observations. A suggested modification might be to determine stand composition based upon measured basal area or DBH values. However, as the previously existing data did not specify these classification parameters, including this protocol in the future would not increase similarity between historical data, but rather would increase future consistency.

Advantages of Object-Oriented Image Processing

Although the methodology of object-oriented analysis needs improvement, object-oriented analysis provides a good solution to the frequent problem of classifying objects that is associated with high resolution imagery. Consider a high resolution image of a forested landscape. Individual trees may span multiple pixels, and, although these pixels all represent the same tree, there is inherent variation among them. Segmentation before classification allows pixels to be grouped into an object (i.e., the tree) and allows the analysis and classification of a continuous group of pixels, rather than individual pixels (Yu et al. 2006). This grouping produces more visually pleasing maps, as the process mimics the delineation process made by the human's brain. Given its advantages, future research in image segmentation could promise for forest classification.

Future Research

As object-oriented image analysis has demonstrated an ability to map vegetation, although not as accurately as traditional photo-interpretation (Mathieu et al. 2007), it is worth investigating means of improvement using the data and technologies currently available. This is especially true considering that the automated methodology provides total enumeration and is less expensive while simultaneously more efficient than manual photo-interpretation. As this study produced low accuracies, regardless of imagery used, standard methodological improvements should be fine-tuned to a higher quality on one image, before attempting to distinguish what imagery is the best. Based on the results of this research and the available literature, some suggested methodologies are presented here.

Classification and Regression Tree (CART) Analysis: This research utilized the sample editor function of Definiens Professional for segment classification, as opposed to the rule-based classification method. Further incorporation of statistical methods into the classification process using the rule-based approach could be accomplished through preliminary analysis of training data using a classification and regression tree (CART) approach.. A rule-based segment classification at the stand scale, using Landsat 7 data comparable to that used in this study, resulted in an 83% overall accuracy (Lucas et al. 2007).

A CART analysis statistically determines the most important parameters or attributes to be used in classification, based on training point attributes. For example, a CART analysis would allow the researcher to determine which bands of imagery are most important to the classification process. CART can also provide rule-based classification guidelines, which, once incorporated into the rule-based classification of Definiens, have been shown to be more effective than the sample-based classification method (Gao et al. 2007). CART also has the power to determine the usefulness of ancillary data in classification, which has the potential to eliminate excess data, reducing overall classification cost and processing time.

Ancillary Data: Although this project incorporated three different image data sources, classification was based solely on the properties associated with these respective images. That is, properties associated with bands and derived bands were used. Research has suggested that incorporating ancillary data, such as

LIDAR, wetlands or soils data, can improve classification accuracies (Lu et al. 2008). Inclusion of additional data would allow for a better understanding of each class' properties beyond spectral and textural information and integrate the power of GIS modeling. Modeling of individual canopy shapes in three dimensions has also been suggested as a means of distinguishing between individual trees and guiding their classification (Larsen 2007). However, while this methodology may improve classification it may be impractical because of added cost and time.

Classification Simplification: Research indicates that the accuracy of image-object classification could possibly be improved through the modification of the classification scheme (although this may detract from the original intent of a classification). Simplification of the classes (e.g., coniferous vs. deciduous rather than species level classification) could yield an improvement in overall classification accuracy. Yu et al. (2006) achieved accuracies of 58%, a substantial increase over their original results, by simplifying an individual species classification to a landscape level, non-species specific class scheme.

Simplification also showed a noted improvement with the 1992 National Land Cover Data set that used Landsat TM data to classify the land cover of the United States. Two classification schemes were developed: Level I and Level II. Level I contained nine categories and distinguished major land cover types (e.g., water from forest from agriculture). Level II used 21 categories to further distinguish between cover types (e.g., open water from ice snow from deciduous forest from coniferous forest). Using 1573 reference points, Level I achieved an overall accuracy of 80%, while Level II achieved only 47% overall accuracy (Environmental Protection Agency 1992). Although this project utilized ancillary data, the more detailed classification scheme did not achieve usable accuracies, meaning that the map would be unreliable for field use. However, the simplified classification scheme produced an impressive 80% accuracy.

Although this research demonstrated that object-oriented image analysis is not reliable for discriminating tree species, regardless of scale, with the currently available satellite image resolutions, it did provide some insight regarding procedural improvements. Still, past and current research (Lennartz 2004, Gao et al. 2007) have demonstrated that object-oriented classification is superior to the traditional per-pixel classification method, especially using high spatial resolution satellite imagery. As spatial and spectral imagery resolution continues to improve both spatially and spectrally, further development and perfection of object-oriented image analysis is a necessary step to understanding and translating data into a useful form.

Conclusions

Although this study provided no conclusive evidence as to which of the three satellite images used was "best" for mapping tree species in New Hampshire, the results did provide several insights and conclusions.

1. Segmentation works well in Definiens software, while classification does not. Better methodology, be it software, algorithms or data, is needed.
2. As the highest spectral resolution, but lowest spatial resolution imagery was best for classifying stands, spectral resolution may be more important than spatial resolution for stand and landscape scale classification.
3. Higher spatial resolution is needed to delineate individual tree canopies, but it is likely that high spectral resolutions will be needed to classify them.

The research presented in this thesis was focused on identifying the best imagery for use at three given scales, based upon the accuracies of the resulting classifications. As IKONOS, SPOT and Landsat TM data yielded similarly poor accuracies at the desired levels of detail, perhaps the research focus should shift to identifying to optimal methodology, in lieu of both spectrally and spatially high resolution imagery. It is also important to recognize that many of the methods utilized in this study are beyond the financial and technical grasp of an "everyday" forester. The usefulness of these processes, as they now are, is also fairly limited due to the time required to perform them. However, this research points to possible ways to improve results to an accurate, "useful" level. Once this level is attained, the process of automated species classification at the stand individual tree scale could be fine-tuned, with software parameters standardized to obtain optimal results with the push of a single button. This would allow foresters to quickly, easily and consistently classify their stands, not only providing data about species absence/presence, but also their spatial distribution. This type of data could be the cornerstone for timber management plans and timber inventories and would be more efficient than the current practices of timber cruising. Therefore, attention should be paid to developing efficient and cost effective methods to allow for the use of these methodologies beyond the research arena into the areas of professional forestry. In the future, the usefulness of the most accurate classification methods may be limited by the cost to those that need them. From a forest management perspective, this research is promising but needs either technology or methodology improvements before a useable product can be attained.

LIST OF REFERENCES

Addink, E.A., S.M. de Jong and E.J. Pebesma. 2007. *The importance of scale in object-based mapping of vegetation parameters with hyperspectral imagery.* Photogrammetric Engineering and Remote Sensing 73(8): 905-912.

Asner, G.P., M. Palace, M. Keller, R. Pereira Jr, J.N.M. Silva and J.C. Zweede. 2002. *Estimating canopy structure in an Amazon Forest from laser range finder and IKONOS satellite observations.* Association of Tropical Biology 34(4): 483-492.

Bauer, M.E., T.E. Burk, A.R. Ek, P.R. Coppin, S.D. Lime, T.A. Walsh, D.K. Walters, W. Befort and D.F. Heinzen. 1994. *Satellite inventory of Minnesota forest resources.* Photogrammetric Engineering and Remote Sensing 60(3): 287-298.

Benz, U.C., P. Hofmann, G. Willhauck, I. Lingenfelder, M. Heynen. 2004. *Multi-resolution, object-orientated fuzzy analysis of remote sensing data for GIS-ready information.* Photogrammetry and Remote Sensing 58: 239-258.

Boyd, D.D., G.M. Foody and W.J. Ripple. 2002. *Evaluation of approaches for forest cover estimation in the Pacific Northwest, USA, using remote sensing.* Applied Geography 22: 376-392.

Brogaard, S. and R. Ólafsdóttir. 1997. *Ground-truths or Ground-lies? Environmental sampling for remote sensing application exemplified by vegetation cover data.* Lund eRep. Phys. Geogr., No. 1. http://www.natgeo.lu.se/Elibrary/LeRPG/1/LeRPG1Article.htm.

Bruzzone, L., Roli, F., & Serpico, S. B. (1995). *An extension to multiclass cases of the Jeffries–Matusita distance.* IEEE Transactions on Geoscience and Remote Sensing, 33, 1318–1321.

Bunting, P and R. Lucas. 2005. *The delineation of tree crowns in Australian mixed species forests using hyperspectral Compact Airborne Spectrographic Imager (CASI) data.*

Czarnecki, C. 2006. Unpublished Data. Master of Science Thesis Research. University of New Hampshire.

Chuvieco, E. and R.G. Congalton. 1988. *Using cluster analysis to improve the selection of training statistics in classifying remotely sensed data.* Photogrammetric Engineering and Remote Sensing 54(9): 1275-1281.

Clark, M.L., D.A. Roberts and D.B. Clark. 2005. *Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales.* Remote Sensing of Environment 96: 375-398.

Cochrane, M.A. 2000. *Using vegetation reflectance variability for species level classification of hyperspectral data.* International Journal of Remote Sensing 21(10): 2075-2087.

Cohen, Jacob (1960). *A coefficient of agreement for nominal scales.* Educational and Psychological Measurement (20): 37–46.

Cole, D.N. 1995. *Experimental trampling of vegetation. II. Predictors of resistance and resilience.* Journal of Applied Ecology 32: 215-224

Congalton, R.G. 1991. *A review of assessing the accuracy of classifications of remotely sensed data.* Remote Sensing of Environment 37: 35-46.

Congalton, R.G. and K. Green. 1993. *A practical look at the sources of confusion in error matrix generation.* Photogrammetric Engineering and Remote Sensing 59 (5): 641-644.

Congalton, R.G. and K. Green. 1999. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, Lewis Publishers, Boca Raton, Florida. 131 pp.

Congalton, R.G. and R.A. Mead. 1983. *A review of three discrete multivariate analysis techniques used in assessing the accuracy of remotely sensed data from error matrices.* IEEE Transactions on Geoscience and Remote Sensing 24(1): 169-174.

Dai, X. and S. Khorram. 1998. *The effects of image misregistration on the accuracy of remotely sensed change detection.* IEEE Transactions on Geoscience and Remote Sensing 36(5): 1566-1577.

Deckert, C. and P.V. Bolstad. 1996. *Forest canopy, terrain and distance effects on global positioning system point accuracy.* Photogrammetric Engineering and Remote Sensing 62 (3): 317-321.

Degraaf, R.M. and M. Yamasaki. 2001. *New England Wildlife.* University Press of New England, Hanover, New Hampshire. 482pp.

Definiens AG. 2006. *Definiens Professional 5 User Guide.* Munich.

Edwards, T.C., Jr., G.G. Moisen and D.R. Cutler. 1998. *Assessing map uncertainty in remotely-sensed, ecoregion-scale cover-maps.* Remote Sensing and Environment 63: 73-83.

Environmental Protection Agency. 1992. NLCD (National Land Cover Data). URL: http://www.epa.gov/mrlc/nlcd.html. (last date accessed: 15 May 2008).

Eyre, F.H. 1980. *Forest Cover Types of the United States and Canada.* Society of American Foresters. Washington, D.C.

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.

Foody, G.M. 1999. *The continuum of classification fuzziness in thematic mapping.* Photogrammetric Engineering and Remote Sensing 65 (4): 443-451.

Foody, G.M. 2002. *Status of land cover classification accuracy assessment.* Remote Sensing of Environment 80: 185-201.

Freedman, J. 1949. *Stratigraphy and Structure of the Mt. Pawtuckaway Quadrangle, Southeastern New Hampshire.* Geological Society of America Bulletin. pp. 449-492.

Gao, Y., J.F. Mas, I. Niemeyer, P.R. Marpu and J.L. Palacio. 2007. *Object-based image analysis for mapping land-cover in a forest area.* Proceedings from the 5th International Symposium on Spatial Data Quality. ITC, Enschede, The Netherlands.

GeoEye. 2007. GeoEye Imagery Products: IKONOS. URL: http://www.geoeye.com/products/imagery/ikonos/precision_plus.htm. (last date accessed: 15 May 2008).

Hay, G.J., K.O. Niemann and G.F. McLean. 1996. *Object-specific image-texture analysis of H-resolution forest imagery.* Remote Sensing of Environment 55: 108-122.

Irons, J.R., B.L. Markham, R.F. Nelson, D.L. Toll and D.L. Williams. 1985. *The effects of spatial resolution on the classification of Thematic Mapper data.* International Journal of Remote Sensing 6(8): 1385-1403.

Jacquin, A., L. Misakova and M. Gay. 2007. *A hybrid object-based classification approach for mapping urban sprawl in periurban environment*, Landscape Urban Planning, doi:10.1016/j.landurbplan.2007.07.006

Jensen, J.R. 2005. *Introductory Digital Image Processing: A remote sensing perspective 3rd ed.* Pearson Prentice Hall, Upper Saddle River, NJ.

Joyce, A.T. 1978. *Procedures for Gathering Ground Truth Information for a Supervised Approach to a Computer-Implemented Land Cover Classification of Landsat-Acquired Multispectral Scanner Data.* NASA Reference Publication 1015. Houston, TX: National Aeronautics and Space Administration, 43 p.

Larsen, M. 2007. *Single tree species classification with a hypothetical multi-spectral satellite.* Remote Sensing of the Environment 110: 523-532

Leica Geosystems. 2005. *ERDAS Field Guide.* GIS and Mapping Division, Atlanta.

Lennartz, S.P. 2004. *Pixel versus segment-based image processing techniques for forest cover type mapping using IKONOS imagery.* Master of Science Thesis. University of New Hampshire.

Liu, C., P. Frazier and L. Kumar. 2007. *Comparative assessment of the measures of thematic classification accuracy.* Remote Sensing of Environment 107: 606-616.

Lo, C.P. and J. Choi. 2004. *A hybrid approach to urban land use/cover mapping using Landsat 7 Enhanced Thematic Mapper Plus (ETM+) images.* International Journal of Remote Sensing 25(14): 2687-2700.

Lu, D., M. Batistella, E. Moran and E.E. Miranda. 2008. *A Comparative Study of Landsat TM and SPOT HRG Images for Vegetation Classification in the Brazilian Amazon.* Photogrammetric Engineering and Remote Sensing 74(3): 311-321.

Lucus, R., A. Rowlands, A. Brown, S. Keyworth and P. Bunting. 2007. *Rule-based classification of multi-temporal satellite imagery for habitat and agricultural land cover mapping.* Journal of Photogrammetry and Remote Sensing 62: 165-185.

Lunetta, R.S., R.G. Congalton, L.K. Fenstermaker, J.R. Jensen, K.C. McGuire and L.R. Tinney. 1991. *Remote sensing and geographic information system data integration: error sources and research issues.* Photogrammetric Engineering and Remote Sensing 57(6): 677-687.

Manning, R. and P.L. Cormier. 1980. *Trends in the temporal distribution of park use.* Proceedings of National Outdoor Recreation Trends Symposium. 20-23 April. Durham, NH.

Martin, M.E., S.D. Newman, J.D. Aber and R.G. Congalton. 1998. *Determining forest species composition using high spectral resolution remote sensing data.* Remote Sensing of Environment 65: 249-254.

Mathieu, R., J. Aryal and A.K. Chong. 2007. *Object-based classification of Ikonos imagery for mapping large-scale vegetation communities in urban areas.* Sensors 2007 7: 2860-2880.

McCabe, M.F. and E.F. Wood. 2006. *Scale influences on the remote estimation of evapotranspiration using multiple satellite sensors.* 2006 Remote Sensing of Environment 105.

McCloy, K.R. and P.K. Bocher. 2007. *Optimizing image resolution to maximize the accuracy of hard classification.* Photogrammetric Engineering and Remote Sensing 73(8): 893–903.

Miguel-Ayanz, J.S. and G.S. Biging. 1997. *Comparison of single stage and multi-stage classification approaches for cover type mapping with TM and SPOT data.* Remote Sensing of Environment 59: 92-104.

NH DRED (New Hampshire Department of Resources and Economic Development). 1996. *New Hampshire Forest Resources Plan.* Forest Resources Plan Steering Committee and New Hampshire Department of Resources and Economic Development and Division of Forests and Lands.

New Hampshire State Climate Office. 2008. *Climatic Averages for Selected New Hampshire Cites and Towns: Durham.* University of New Hampshire. URL: http://www.unh.edu/stateclimatologist/. (last date accessed: 15 May 2008).

Plourde, L.C. 2000. *Important Factors in Assessing the Accuracy of Remotely Sensed Forest Vegetation Maps: A Comparison of Sampling and Total Enumeration and Their Effects on Data Accuracy.* Master of Science Thesis. University of New Hampshire.

Plourde, L. and R.G. Congalton. 2003. *Sampling method and sample placement: how do they affect the accuracy of remotely sensed maps?* Photogrammetric Engineering and Remote Sensing 69(3): 289-297.

Piedallu, C. and J.C. Gégout. 2005. *Effects of forest environment and survey protocol on GPS accuracy.* Photogrammetric Engineering and Remote Sensing 71 (9): 1071-1078.

Pontius, R.G. 2000. *Quantification Error Versus Location Error in Comparison of Categorical Maps.* Photogrammetric Engineering and Remote Sensing 66 (8): 1011-1016.

Pugh, S.A. 1997. *Applying Spatial Autocorrelation Analysis to Evaluate Error in New England Forest Cover Type Maps Derived From Landsat Thematic Mapper Data.* Master's Thesis. University of New Hampshire, Durham.

Pugh, S.A. and R.G. Congalton. 2001. *Applying spatial autocorrelation analysis to evaluate error in New England forest-cover-type maps derived from Landsat Thematic Mapper data.* Photogrammetric Engineering and Remote Sensing 67(5): 613-620.

Rahman, A.F., J.A. Gamon, D.A. Sims and M. Schmidts. 2003. Optimum *pixel size for hyperspectral studies of ecosystem function in southern California chaparral and grassland.* Remote Sensing of Environment 84: 192-207

Rocchini, D. and A. Di Rita. 2005. *Relief effects on aerial photos geometric correction. Applied* Geography 25: 159-168.

Rubens, A.F., C.A. Vettorazzi and G.A. Sarries. 2002. *Evaluation of the accuracy of positioning a GPS receiver operating under different vegetation covers.* 6(2): 325-331.

Sader, S.A., D. Ahl and W.S. Liou. 1995. *Accuracy of Landsat-TM and GIS rule-based methods for forest wetland classification in Maine.* Remote Sensing of Environment 53: 133-144.

Sawaya, K.E., L.G. Olmanson, N.J. Heinert, P.L. Brezonik and M.E. Bauer. 2002. *Extending satellite remote sensing to local scales: land and water resource monitoring using high-resolution imagery.* Remote Sensing of Environment 88: 144-156.

Schriever, J. 1992. *Using Multi-Temporal LANDSAT Thematic Mapper Data to Map Forest Cover-Types in New Hampshire.* Master of Science Thesis. University of New Hampshire.

Society For the Protection of New Hampshire Forests. 2007. *New Hampshire Conservation/Public Lands Layer available from the Complex Systems Research Center.* NH Granit. URL: http://www.granit.unh.edu/data/downloadfreedata/category/databycategory.html. (last date accessed: 15 May 2008).

Sperduto, D.D. 1995. *Natural Community Systems of New Hampshire.* New Hampshire Natural Heritage Bureau and The Nature Conservancy. Department of Resources and Economic Development – Division of Lands and Forests, Concord, NH.

Sperduto, D.D and W.F. Nichols. 1994. *Natural Communities of New Hampshire.* New Hampshire Natural Heritage Bureau, Concord, NH. UNH Cooperative Extension, Durham, NH.

SPOT Image. 2007. Spot 5 Supermode Technical Sheet. URL: http://www.spot.com/web/SICORP/449-sicorp-spot-images.php. (last date accessed: 15 May 2008).

Stehman, S.V. and R.L. Czaplewski. 1998. *Design and analysis for thematic map accuracy assessment: fundamental principles.* Remote Sensing of Environment 64: 331-334.

Story, M. and R.G. Congalton. 1986. *Accuracy assessment: a user's perspective.* Photogrammetric Engineering and Remote Sensing 52(3): 397-399.

Stuckens, J., P.R. Coppin and M.E. Bauer. 2000. *Integrating contextual information with per-pixel classification for improved land cover classification.* Remote Sensing of Environment 71: 282-296.

Toll, D.L. 1985. *Effect of Landsat Thematic Mapper sensor parameters on land cover classification.* Remote Sensing of Environment 17: 129-140.

University of Minnesota. 2008. Department of Natural Resources. DNR Garmin Application. URL: http://www.dnr.state.mn.us/mis/gis/tools/arcview/extensions/DNRGarmin/DNRGarmin.html. (last date accessed: 15 May 2008).

USDA (United States Department of Agriculture). 2006. *Soil Survey Geographic SSURGO database for Rockingham County, New Hampshire.* NH GRANIT. URL: http://www.granit.sr.unh.edu/cgi-bin/nhsearch?dset=soils/soil015. (last date accessed: 15 May 2008).

USGS and NASA. 4 April 2006. *Landsat: A Global Land-Observation Project. Landsat Data Sheet.* URL: http://landsat.usgs.gov/project_facts/project_description.php. (last date accessed: 15 May 2008).

Verbyla, D.L. and S.H. Boles. 2000. *Bias in land cover change estimates due to misregistration.* International Journal of Remote Sensing 21(18): 3553-3560.

Vitousek, P.M. 1994. *Beyond Global Warming: Ecology and Global Change.* Ecology 75 (7): 1861-1876.

Warner, T.A., J.Y. Lee and J.B. McGraw. 1998. *Delineation and identification of individual trees in the eastern deciduous forest.* Proceedings of the Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry, Victoria, British Columbia, Canada. February 10-12, 1998. Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Victoria, British Columbia, 81-91.

Yu, Q., P. Gong, N. Clinton, G. Bigin, M. Kelly and D. Schirokauer. 2006. *Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery.* Photogrammetric Engineering and Remote Sensing 72(7): 799-811.

Yu, Q., P. Gong, Y.Q. Tian, R. Pu and J. Yang. 2008. *Factors affecting spatial variation of classification uncertainty in an image object-based vegetation mapping.* Photogrammetric Engineering and Remote Sensing 74(8): 1007-1018

APPENDICES

APPENDIX A: Classification System Guidelines and Forest Type Definition

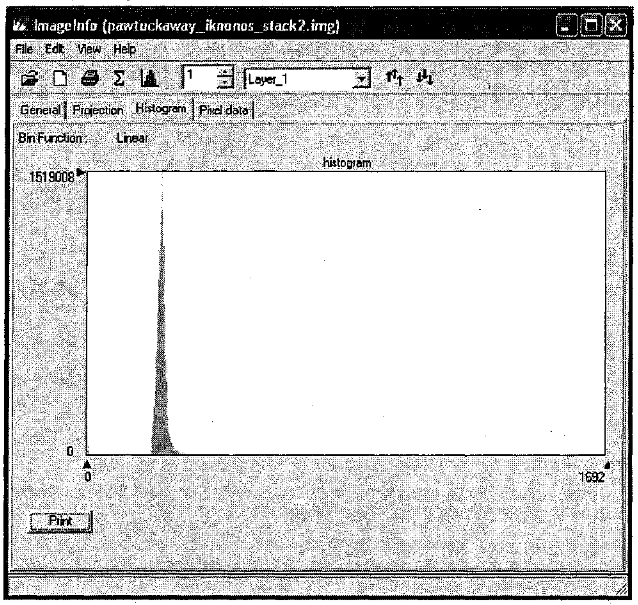
- Beech (B)
  - Description: A stand primarily or completely comprised of American beech (*Fagus grandifolia*)
  - Classify as B when beech is at least 70% of the stand.
- Oak (O)
  - Description: A stand primarily or completely comprised of northern red oak (*Quercus rubra*) or white oak (*Quercus alba*)
  - Classify as O when either red or white oak is at least 70% of the stand.
- Red Maple (RM)
  - Description: A stand primarily or completely comprised of maple species (*Acer spp.*) most likely red maple (*Acer rubrum*) or sugar maple (*Acer saccharum*). Although labeled Red Maple, either species is acceptable.
  - Classify as RM when either red or sugar maple is at least 70% of the stand.
- Other Forest (OF)
  - Description: A stand primarily comprised of deciduous species, but not dominated by beech, oak or maple. This stand may be comprised of any combination of beech, oak or maple, but may also be comprised or dominated by birch (*Betula spp.*), shagbark hickory (*Carya ovata*) or hop hornbeam (*Oystra virginiana*).
  - Classify as OF when:
    1. the stand is at least 70% deciduous species
    2. the stand is not at least 70% of single species B, O or RM
- Hemlock (H)
  - Description: A stand primarily or completely comprised of eastern hemlock (*Tsuga Canadensis*)
  - Classify as H when hemlock is at least 70% of the stand
- White Pine (WP)
  - Description: A stand primarily or completely comprised of eastern white pine (*Pinus strobus*), red pine (*Pinus resinosa*), pitch pine (*Pinus rigida*) or eastern red cedar (*Juniperus virginiana*)

- Classify as WP when white pine or above species is at least 70% of the stand
- White Pine/Hemlock (WPH)
  - Description: A stand primarily or completely comprised of a mixture of eastern hemlock (*Tsuga canadensis*) or eastern white pine (*Pinus strobus*)
  - Classify as WPH when:
    1. The stand is coniferous
    2. The stand is not at least 70% of single species H or WPH
    3. The stand is at least 30% of hemlock **and** 30% of white pine
    4. When combined, hemlock and pine comprise at least 70% of the stand
- Mixed Forest (MF)
  - Description: A stand comprised of a mixture of deciduous and coniferous species
  - Classify as MF when:
    1. The stand is less than 70% of deciduous species
    2. The stand is less than 70% of coniferous species **or** is not classifiable as H, WP or WPH
    3. The stand is comprised of tree species
- Non-forest (NF)
  - Description: A "stand area" that is not forested (e.g. marsh, wetland, open field, rock, regeneration)
  - Classify as NF when less than 30% of the area is forested.

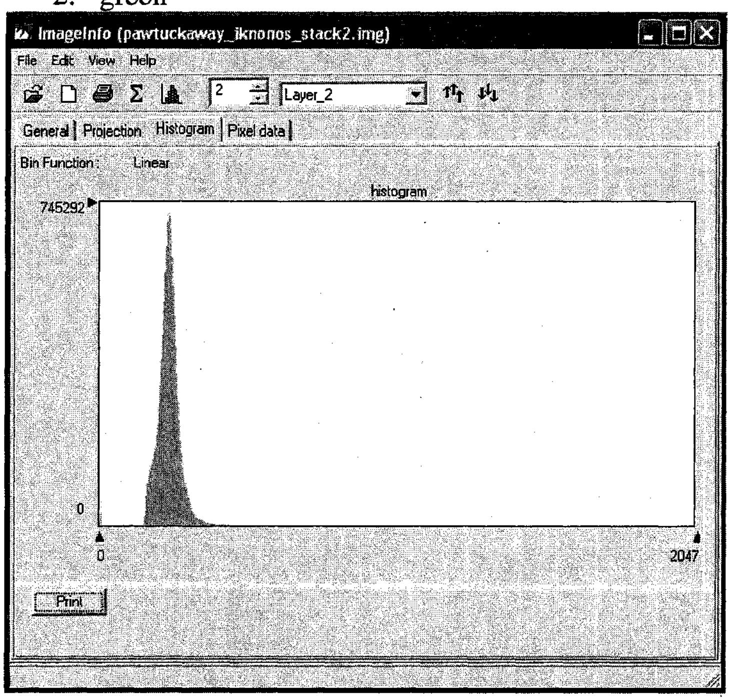
APPENDIX B: Histograms for IKONOS, SPOT and Landsat TM image bands

IKONOS Band:

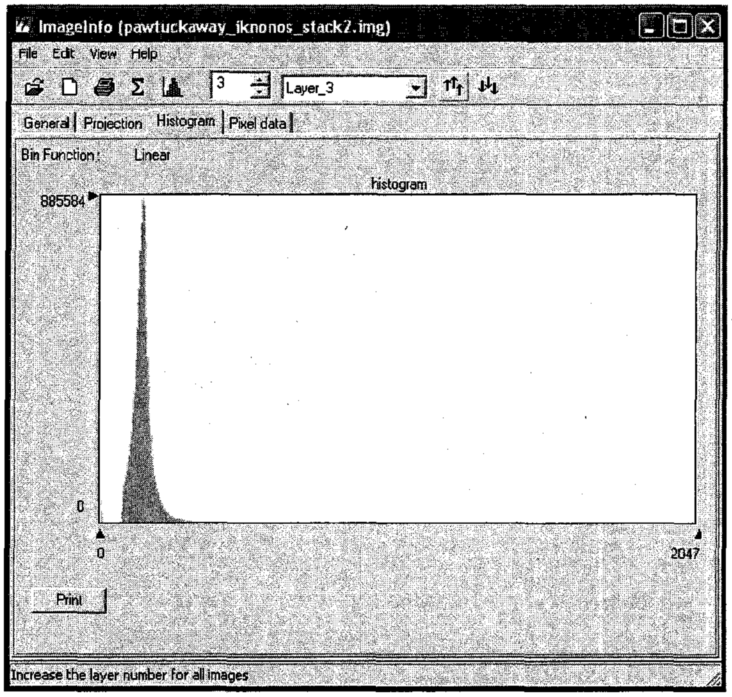
1. blue



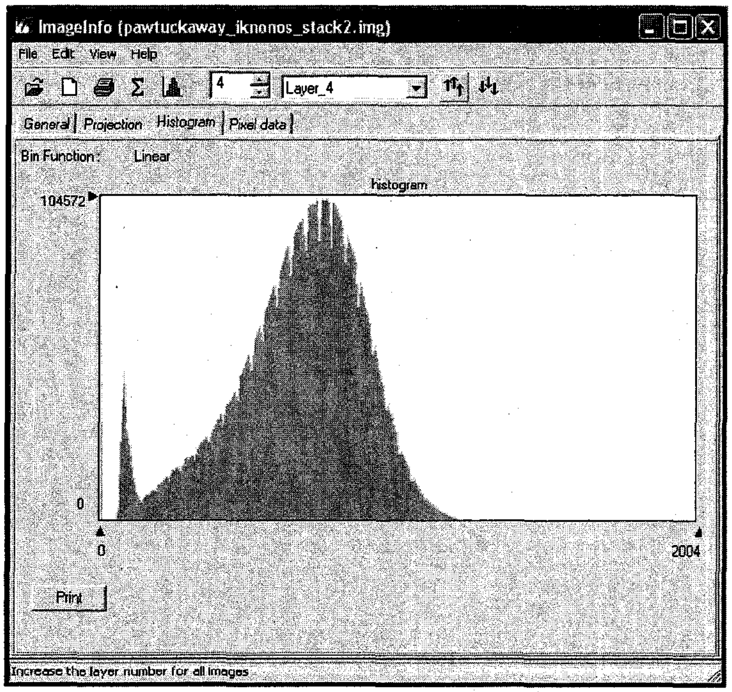
2. green



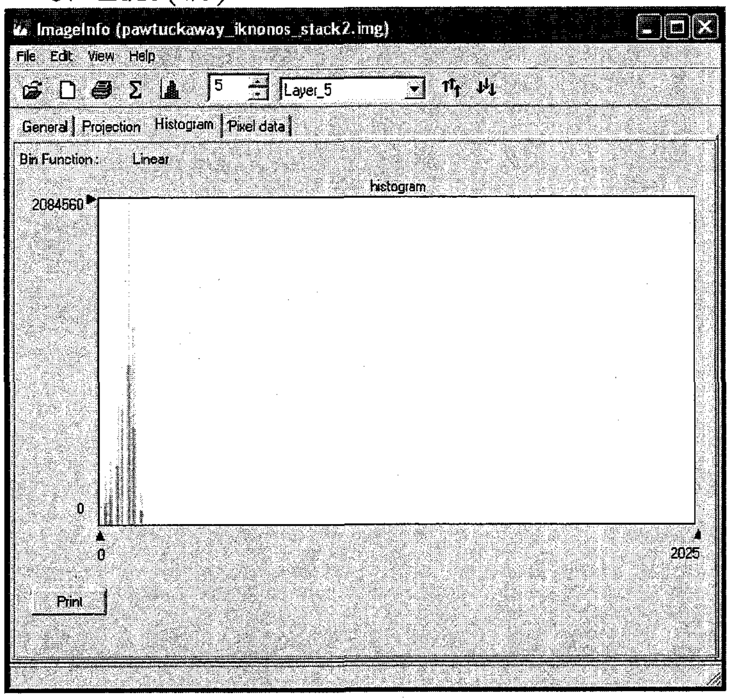
3. red


ImageInfo (pawtuckaway_iknonos_stack2.img)
File Edit View Help
3
Layer_3
General Projection Histogram Pixel data
Bin Function: Linear
histogram
885584
0
0
2047
Print
Increase the layer number for all images

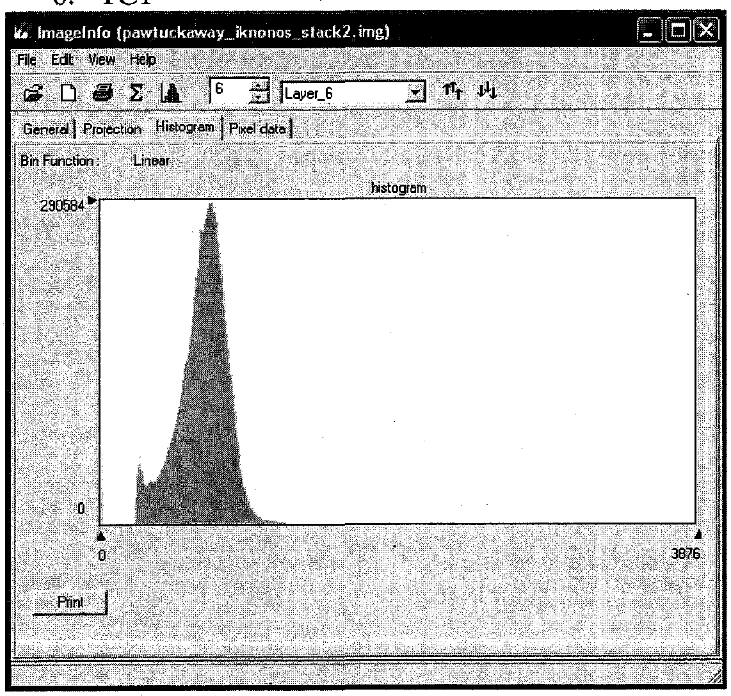
4. NIR


ImageInfo (pawtuckaway_iknonos_stack2.img)
File Edit View Help
4
Layer_4
General Projection Histogram Pixel data
Bin Function: Linear
histogram
104572
0
0
2004
Print
Increase the layer number for all images

5. IR/R (4/3)

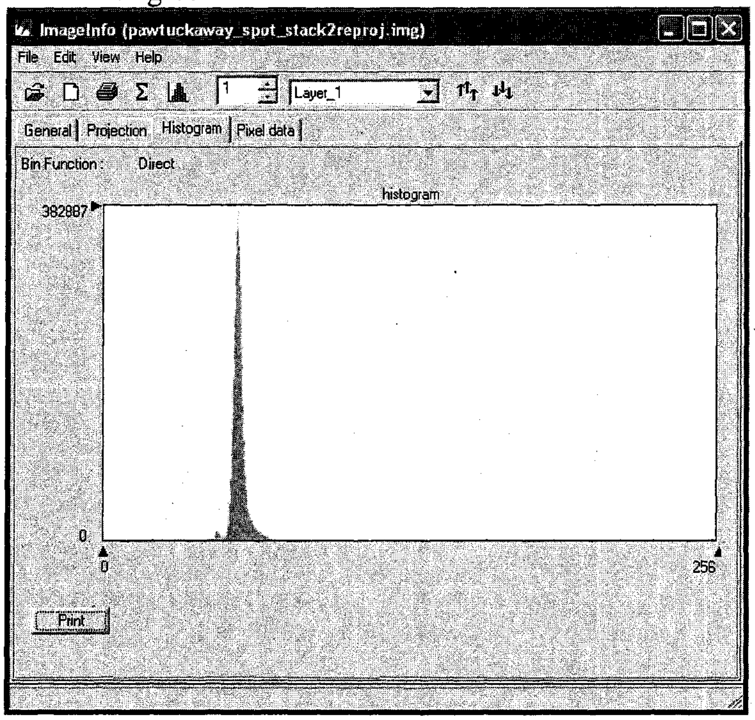

ImageInfo (pawtuckaway_iknonos_stack2.img)
File Edit View Help
5
Layer_5
General Projection Histogram Pixel data
Bin Function: Linear
histogram
2084560
0
0
2025
Print

6. TC1

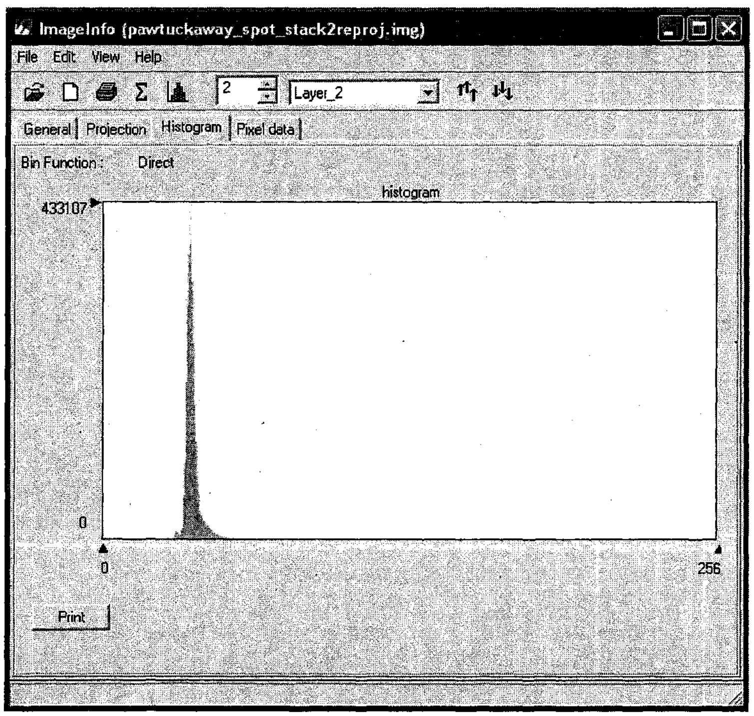

ImageInfo (pawtuckaway_iknonos_stack2.img)
File Edit View Help
6
Layer_6
General Projection Histogram Pixel data
Bin Function: Linear
histogram
290584
0
0
3876
Print

SPOT Band:

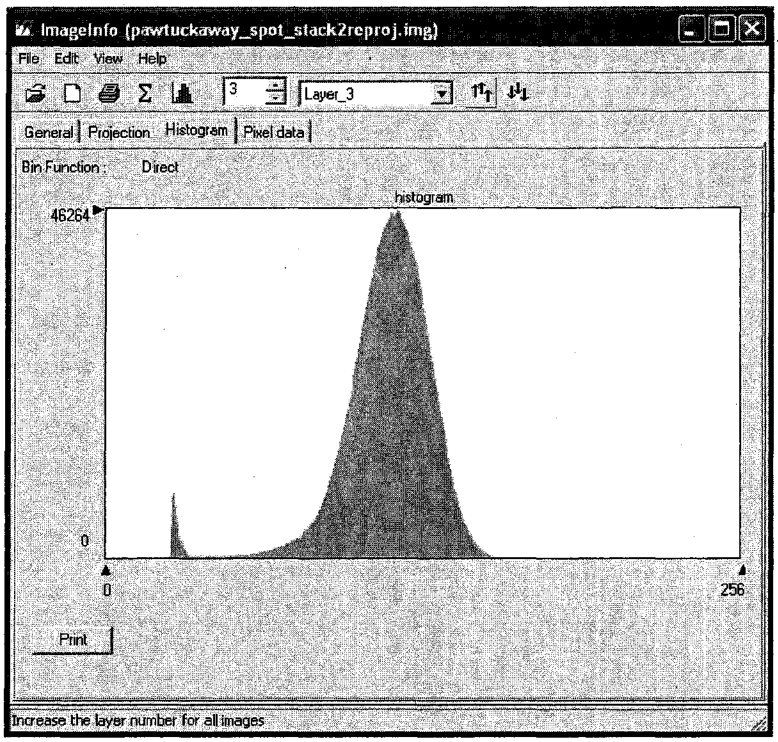
1. green



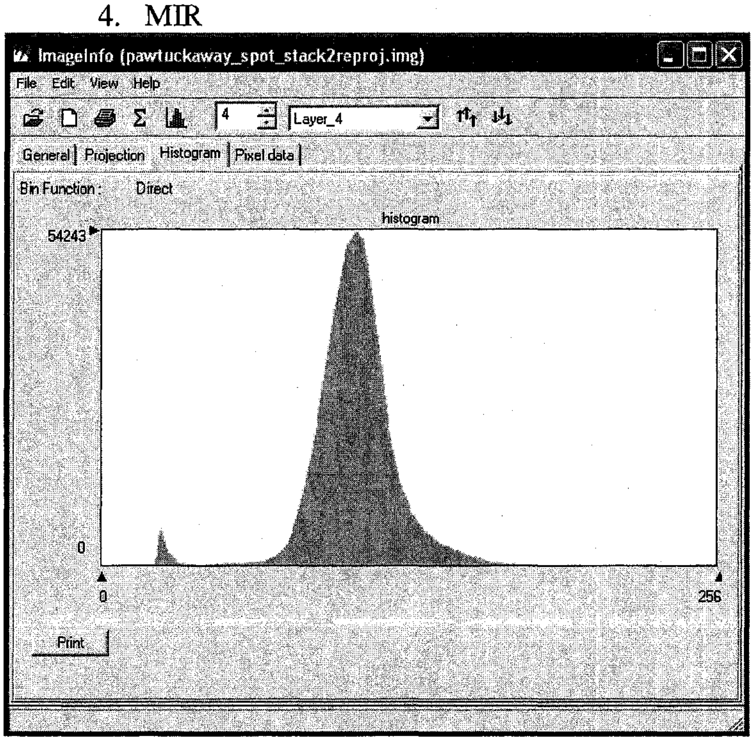
2. red



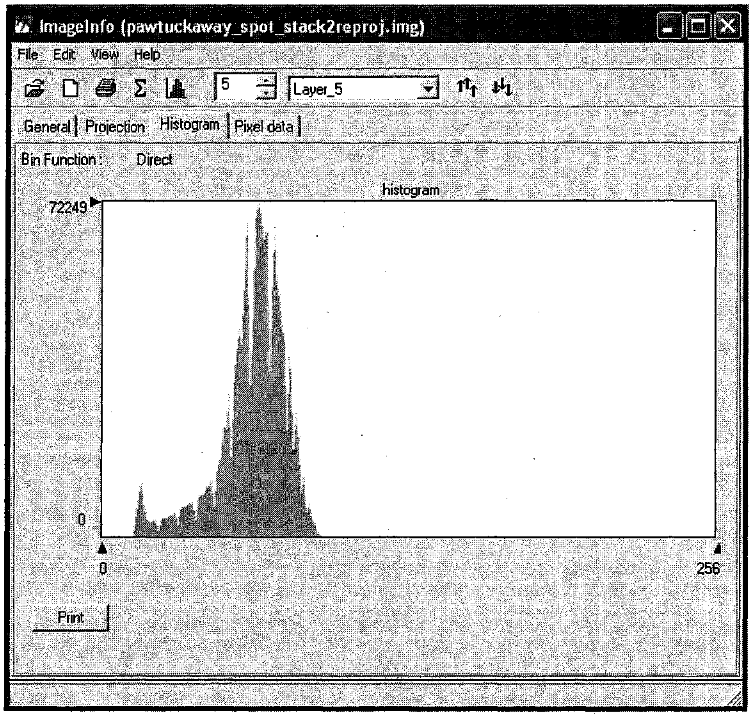
3. NIR


ImageInfo (pawtuckaway_spot_stack2reproj.img)
File Edit View Help
3
Layer_3
General Projection Histogram Pixel data
Bin Function: Direct
histogram
46264
0
0
256
Print
Increase the layer number for all images

4. MIR


ImageInfo (pawtuckaway_spot_stack2reproj.img)
File Edit View Help
4
Layer_4
General Projection Histogram Pixel data
Bin Function: Direct
histogram
54243
0
0
256
Print

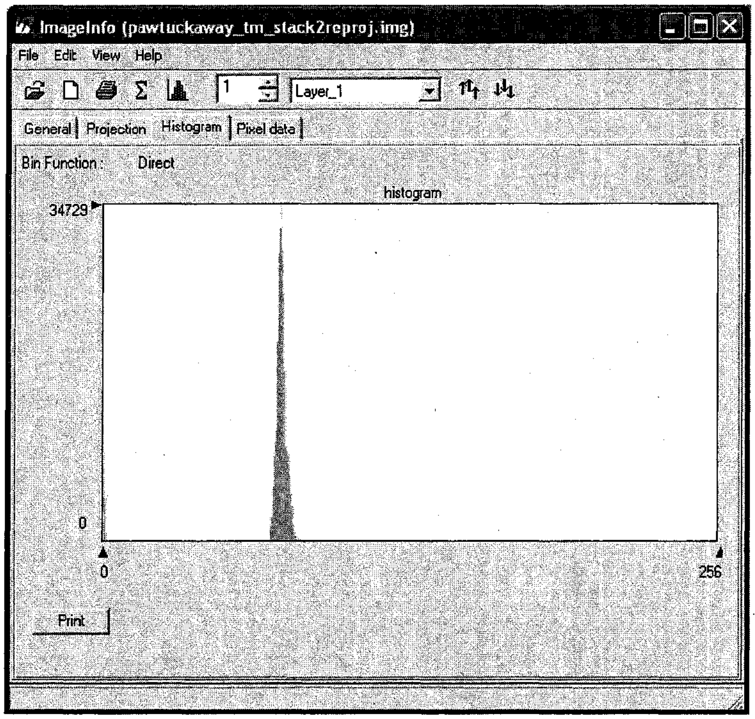
5. IR/R



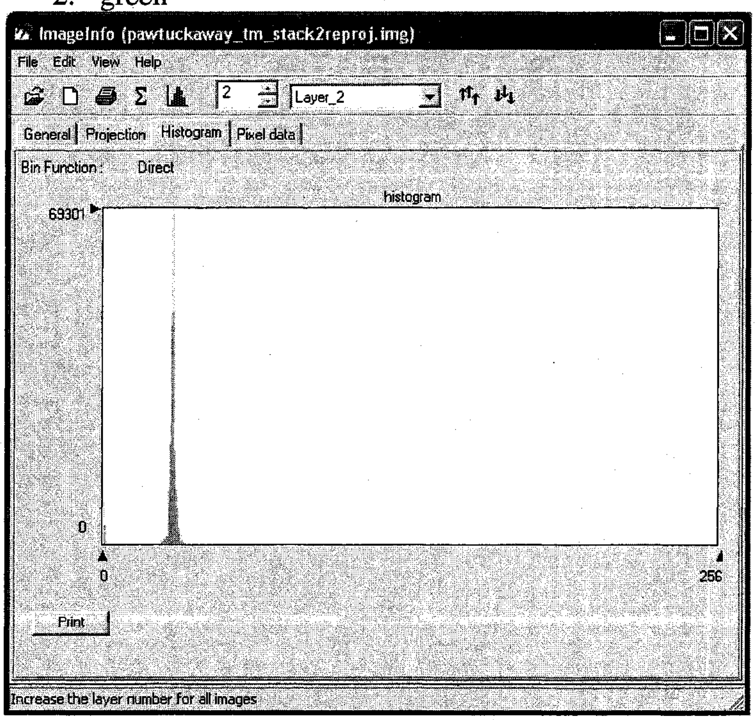
ImageInfo (pawtuckaway_spot_stack2reproj.img)
File Edit View Help
5
Layer_5
General Projection Histogram Pixel data
Bin Function : Direct
histogram
72249
0
0
256
Print

Landsat 5 TM Band:

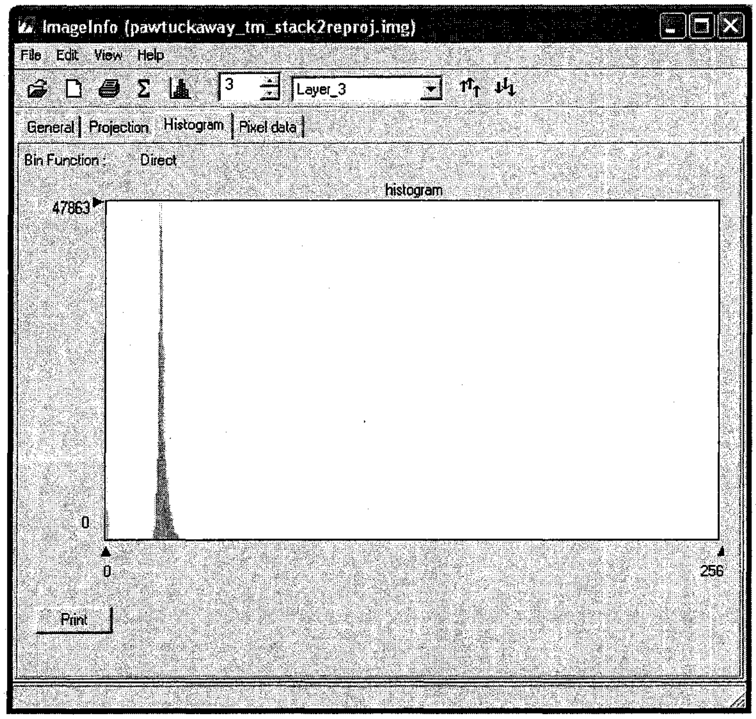
1. blue



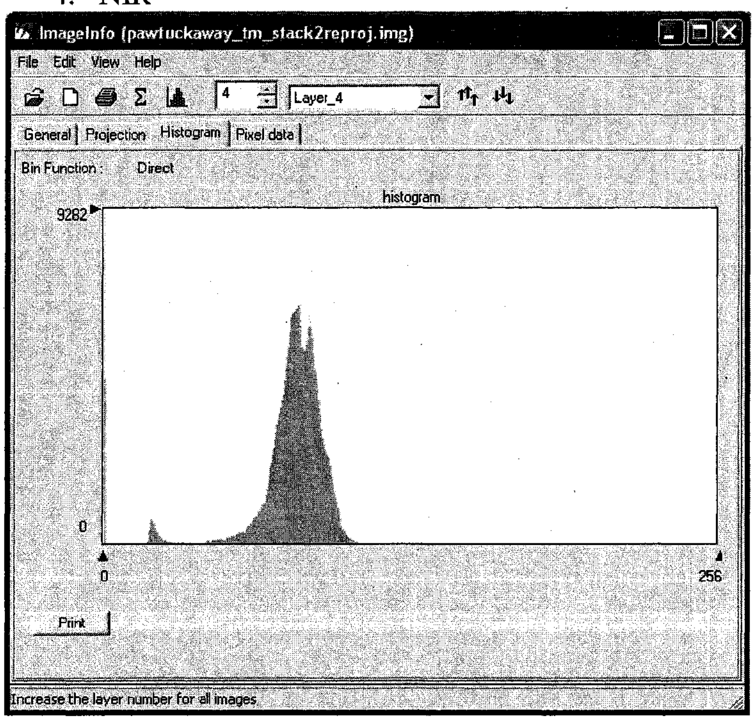
2. green



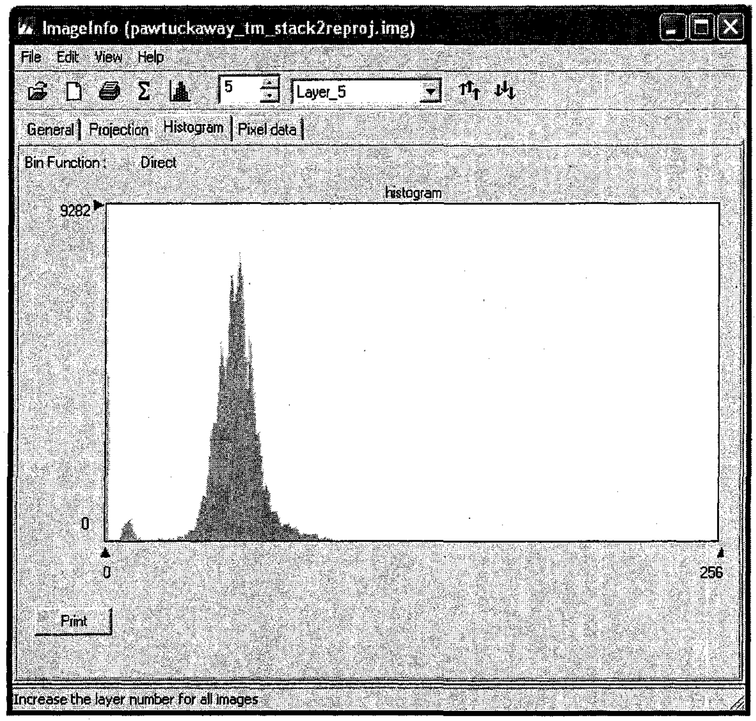
3. red


ImageInfo (pawtuckaway_tm_stack2reproj.img)
File Edit View Help
3
Layer_3
General Projection Histogram Pixel data
Bin Function : Direct
histogram
47863
0
0
256
Print

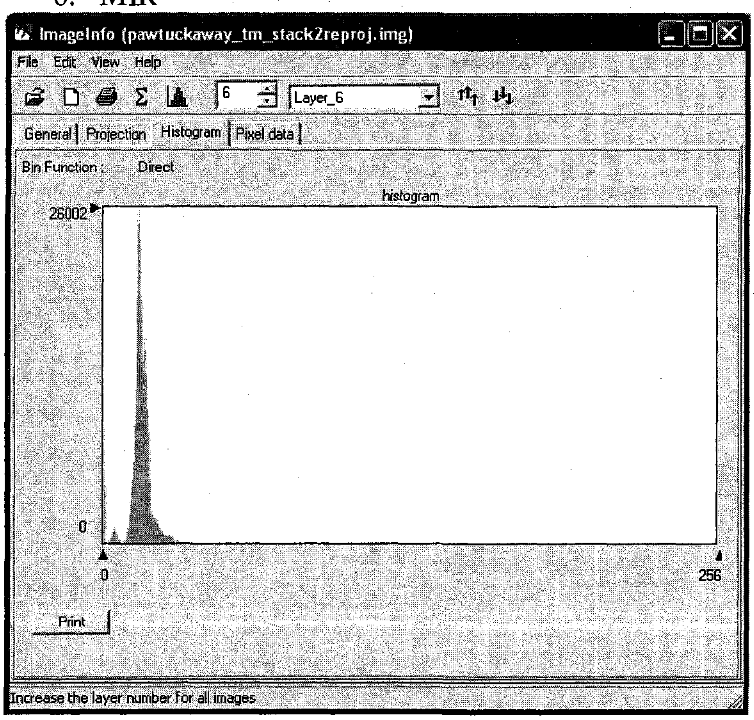
4. NIR


ImageInfo (pawtuckaway_tm_stack2reproj.img)
File Edit View Help
4
Layer_4
General Projection Histogram Pixel data
Bin Function : Direct
histogram
9282
0
0
256
Print
Increase the layer number for all images

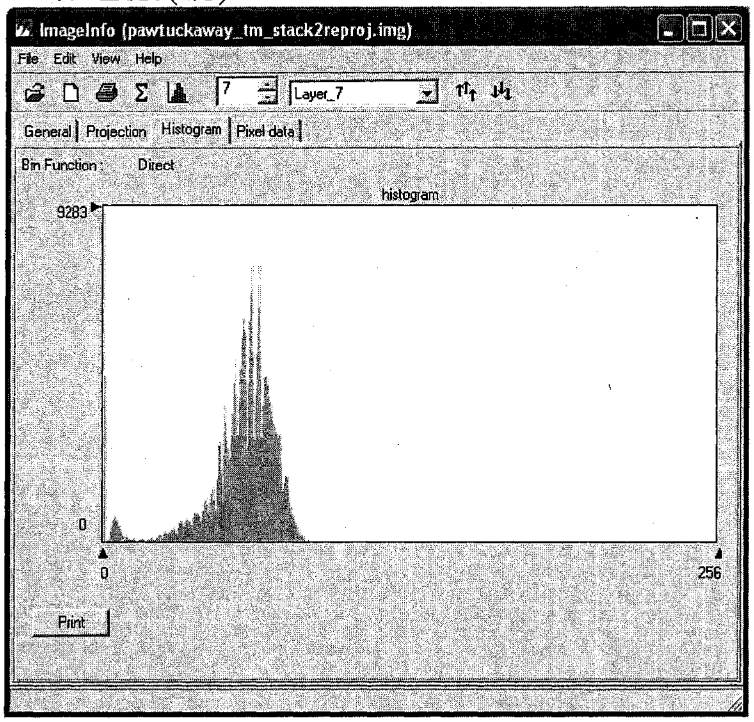
5. MIR


ImageInfo (pawtuckaway_tm_stack2reproj.img)
File Edit View Help
5
Layer_5
General Projection Histogram Pixel data
Bin Function : Direct
histogram
9282
0
0
256
Print
Increase the layer number for all images

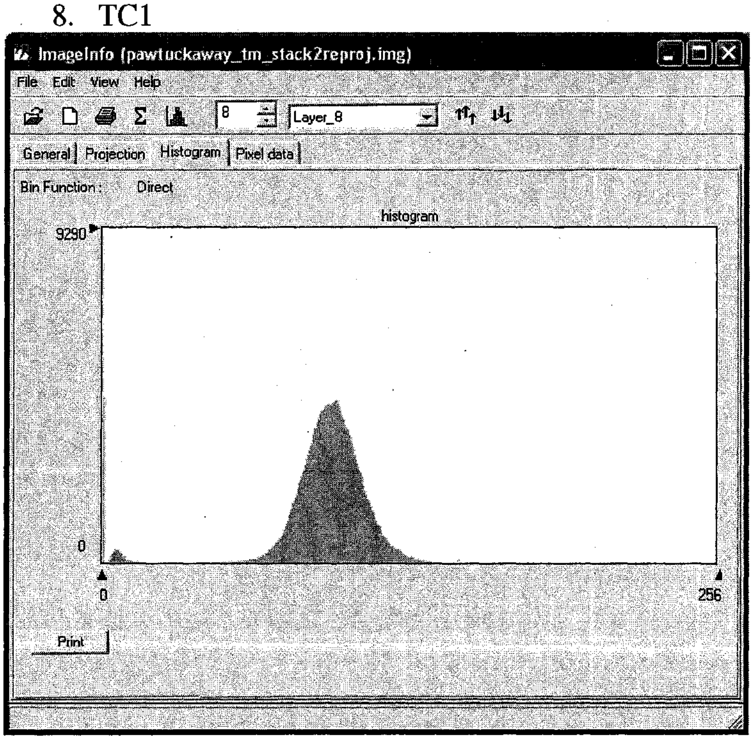
6. MIR


ImageInfo (pawtuckaway_tm_stack2reproj.img)
File Edit View Help
6
Layer_6
General Projection Histogram Pixel data
Bin Function : Direct
histogram
26002
0
0
256
Print
Increase the layer number for all images

7. IR/R (4/3)


ImageInfo (pawtuckaway_tm_stack2reproj.img)
File Edit View Help
7
Layer_7
General Projection Histogram Pixel data
Bin Function : Direct
histogram
9283
0
0
256
Print

8. TC1


ImageInfo (pawtuckaway_tm_stack2reproj.img)
File Edit View Help
8
Layer_8
General Projection Histogram Pixel data
Bin Function : Direct
histogram
9290
0
0
256
Print