NOTE TO USERS

MODELLING AND MAPPING POTENTIAL HOODED WARBLER (*Wilsonia citrina*) HABITAT USING REMOTE SENSING

By

Jonathan Pasher, B.Sc. (Hons.)

A thesis submitted to

the Faculty of Graduate Studies and Research

in partial fulfillment of the requirements of the degree of

Master of Science

Department of Geography and Environmental Studies

Carleton University

Ottawa, Ontario

June 2005

Your file Votre référence
ISBN: 0-494-10144-X
Our file Notre référence
ISBN: 0-494-10144-X

ABSTRACT

Conserving potential habitat for threatened or endangered species is critical to their survival. This research investigated canopy gap characteristics associated with hooded warbler (*Wilsonia citrina*) nest sites in southern Ontario, Canada, and linked them to remote sensing analysis of canopy heterogeneity.

Using hemispherical photography, it was shown that there was significantly less canopy cover at nest sites compared to non-nest sites, that the maximum gap size at nest sites was significantly larger, and that non-nest sites had approximately the same number of gaps as nest sites, but they lacked the large canopy gap that was typically present at nest sites.

In high and medium resolution satellite imagery, many highly significant differences were found between nest and non-nest sites in terms of spectral, textural, and sub-pixel image fractions, and this information was used to produce habitat/non-habitat maps for the hooded warbler. The success rate was high for the focal forest (St. Williams), where most known nest sites occur (70 – 80% accuracy for nest habitat in Ikonos and Landsat classification). However, maps generated for all of southern Ontario suffered from high errors of commission for the habitat class.

These results suggest that while continuous habitat mapping using medium resolution imagery and the methods employed in this research may not be suitable across this region, classification of targeted forests using high or medium resolution imagery can provide spatially explicit maps of habitat and non-habitat. Regional mapping, either as a whole or in subsets, may be possible given more training data sampled from the major potential forests in order to provide more local training data.

ACKNOWLEDGEMENTS

This thesis is dedicated to my older brother Daniel Tsur Pasher. Dan, your thirst for knowledge has always inspired me to strive for the highest level of education possible. Even though you were not able to achieve your own academic goals, I know how proud of me you are for achieving mine. We may not have always been close, but I will always admire you, and look up to you for being my older brother. I will always think of you as I continue my intellectual journey.

Many thanks to Doug and Kathy for your supervision and guidance. I greatly appreciate your patience over these past two years, and hope that you have enjoyed working on this project as much as I have. I trust it has been the start to many more future bridges between your two worlds!

Funding for this research was provided through Dr. King's NSERC grant, along with generous funding supplied by Dr. Lindsay through the National Wildlife Research Centre (Canadian Wildlife Service, Environment Canada). Many thanks to the Canadian Foundation for Innovation, along with the private donors, who provided me with such a wonderful environment to work in.

Thank you Hazel, as well as others in the Geography Department, for helping me ensure that all my applications were the best that they could be. I owe much of my scholarship success to you. Your support has been much appreciated.

The fieldwork for this project could never have been completed without the most generous help of Stephanie Melles and Brad McLeod. Your love for nature was certainly inspiring, so thank you for sharing it with us! Thanks as well to Debbie Badzinski and her

colleagues at Bird Studies Canada for providing help in the field along with their data and expertise.

Robby, I know sharing a shack with a bunch of Kentucky fishermen and visiting endless trailer parks was not your idea of a holiday, but thanks so much for all your help in the field and I hope you enjoyed your summer as much as I did! The quest for the hooded warbler was certainly never without adventure. Our "whiteboard" chats have been essential for getting this thesis done, so thanks for listening and providing your ideas and opinions.

Thanks Evan for all of your forest and remote sensing wisdom. I hope you didn't mind spending time bouncing ideas around with me and providing me guidance on those long winter days when things seemed to be moving backwards. I know one day you will make a fantastic scientist and a wonderful supervisor for up and coming researchers.

A big thank you to the Clocktower Consulting Group, and the boys from CCRS. What is being a Masters student without a few visits to the pub for Tuesday night "discussions"?

Thanks to Bruce Summers, for introducing me to the wonderful world of geography and geomatics. Your passion for teaching certainly left an impression on me. I hope others are lucky enough to be taught and inspired by you.

As well, thanks to JC (Dr. John Clarke) for giving me such an excellent research opportunity and being a source of endless support, advice, and encouragement. I know I let you down time after time when I turn my back on historical geography, but still have a spot on my coffee table polished and waiting for the long awaited sequel to LPE.

All my other friends and family, thank you for always showing interest in my work, and encouraging me to continue my education.

Mum and Dad, thanks for signing me up for university and not letting me watch television all day! Guess I needed a little nudge to get going! You two have provided me so much over the years, but mostly the ability to think for myself and be independent and for that I cannot thank you enough.

Lastly and most importantly, thanks so much to my best friend Tina. Your constant reminders how wonderful and exciting life truly can be are inspirational to say the least. These past two years have been full of excitement as well as very difficult times and I could never have done this without your love, support, and patience. I admire you for your dedication and passion, and I know you will be proud of me for putting my all into this work.

TABLE OF CONTENTS

ABSTRACT ..... **ii**
ACKNOWLEDGEMENTS ..... **iii**
TABLE OF CONTENTS ..... **vi**
LIST OF TABLES ..... **ix**
LIST OF FIGURES ..... **xi**

1.0 INTRODUCTION **1**
1.1 Objectives 4
1.2 Thesis Structure 5

PART A
Habitat Field Study

2.0 INTRODUCTION TO PART A **7**

3.0 BACKGROUND **7**
3.1 The Canadian Species at Risk Act 7
3.2 Critical Habitat 9
3.3 The Hooded Warbler (*Wilsonia citrina*) 10
3.3.1 Population Status in Canada 11
3.3.2 Habitat Selection 14
3.3.3 Recovery 18

4.0 METHODS **20**
4.1 Study Area 20
4.2 Field Sites 23
4.3 Hemispherical Photograph Acquisition and Processing 25
4.4 Comparing Nest and Non-nest Forest Structure: Statistical Analysis of Forest Cover and Canopy Gaps 27
4.4.1 Forest Cover Analysis 27
4.4.2 Analysis of the Gap Size Distribution 28

5.0 RESULTS **30**
5.1 Hemispherical Photographs 30
5.2 Statistical Analysis of Forest Cover and Canopy Gaps 32
5.2.1 Forest Cover Analysis 32
5.2.2 Analysis of the Gap Size Distribution 35

6.0 DISCUSSION **37**
6.1 Significant Findings 37
6.2 Research Limitations and Recommendations 39

PART B
Image Analysis and Habitat Mapping

7.0 INTRODUCTION TO PART B 44

8.0 BACKGROUND..... 45
8.1 Habitat Mapping 45
8.1.1 Habitat Mapping and Modelling for the Hooded Warbler..... 48
8.2 Linear Spectral Unmixing 50

9.0 METHODS 57
9.1 Compilation of Hooded Warbler Point Location Database..... 57
9.2 Mapping With Ikonos Imagery..... 59
9.2.1 Study Area 59
9.2.2 Field Data Collection 59
9.2.3 Image Acquisition and Initial Processing 60
9.2.4 Extraction of Spectral and Textural Information..... 61
9.2.5 Linear Spectral Unmixing..... 64
9.2.5.1 Endmember Selection..... 65
9.2.6 Comparison of Canopy Gaps in Hemispherical Photographs to Shadows in Ikonos Imagery 70
9.2.7 Modelling Canopy Cover Using Ikonos Imagery 70
9.2.8 Comparison of Nest and Non-Nest Sites..... 71
9.2.9 Maximum Likelihood Supervised Classification 71
9.2.10 Binary Logistic Regression 73
9.3 Mapping With Landsat Imagery..... 73
9.3.1 Study Area..... 74
9.3.2 Field Data Collection 74
9.3.3 Image Acquisition and Initial Processing..... 76
9.3.4 Extraction of Spectral and Textural Information 83
9.3.5 Linear Spectral Unmixing 83
9.3.6 Comparison of Nest and Non-Nest Sites..... 83
9.3.7 Maximum Likelihood Classification and Binary Logistic Regression..... 84

10.0 RESULTS..... 86
10.1 Mapping With Ikonos Imagery 86
10.1.1 Endmember Selection and Spectral Unmixing..... 86
10.1.2 Comparison of Canopy Gaps in Hemispherical Photographs to Shadows in Ikonos Imagery 89
10.1.3 Modelling Canopy Cover Using Ikonos Imagery 90
10.1.4 Comparison of Nest and Non-Nest Sites..... 92
10.1.5 Maximum Likelihood Classification 95
10.1.6 Binary Logistic Regression 99

10.2 Mapping With Landsat Imagery 101
10.2.1 Spectral Unmixing..... 101
10.2.2 Comparison of Nest and Non-Nest Sites 102
10.2.3 Maximum Likelihood Classification 107
10.2.4 Binary Logistic Regression 114

11.0 DISCUSSION 114
11.1 Significant Findings 114
11.2 Research Limitations and Recommendations..... 119

12.0 THESIS SUMMARY AND CONCLUSIONS.....122

13.0 REFERENCES 125

APPENDIX 1 TIE POINTS AND GROUND CONTROL POINTS (GCPS) FOR REGISTERING IMAGERY..... 134

APPENDIX 2 TEXTURES CALCULATED FROM THE GREY LEVEL CO-OCCURRENCE MATRIX 139

APPENDIX 3: CALIBRATION TRANSFORMATION APPLIED TO SPECTRAL BANDS (DN VALUES) OF THE SLAVE IMAGE... 140

LIST OF TABLES

Table 1. Mann-Whitney U tests comparing overall percent cover of St. Williams (n = 21) and the pine plantation (n = 12) non-nest sites..... 32

Table 2. Mann-Whitney U tests for percent cover, without pine plantation sites, at nest (n = 24) and non-nest sites (n = 21), focusing on overall coverage and coverage by zenith angles, and azimuth sectors..... 34

Table 3. Kruskal-Wallis test showing no significant difference between percent cover measured in northern direction compared to other directions around nest sites..... 35

Table 4. Mann-Whitney U tests comparing nest (n = 24) and non-nest sites (n = 21) in terms of gap numbers and gap size measured from hemispherical photographs..... 36

Table 5. Landsat 5 TM bands used in this project (Chander et al., 2004)..... 77

Table 6. Average differences of forest validation areas as a percent of the average DN value of the same areas from the master scene. Bolded values show the minimum difference obtained, and therefore the level of calibration that was applied to each band..... 78

Table 7. Mean and standard deviation of DN values for each Landsat spectral band used to determine which bands were well calibrated for forested areas..... 81

Table 8. Ikonos endmembers derived from three endmember selection methods. Cell entries are the mean brightness of each endmember in each spectral band..... 86

Table 9. Pearson correlation matrix for sub-pixel fractions calculated using three endmember selection methods (n = 50). Significance values shown in brackets..... 87

Table 10. Mann-Whitney U tests comparing nest (n = 24) and non-nest (n = 21) sites in terms of spectral variables extracted from Ikonos imagery using a 5 x 5 window..... 92

Table 11. Mann-Whitney U tests comparing nest (n = 24) and non-nest (n = 21) sites in terms of textural variables extracted from Ikonos imagery using a 5 x 5 window..... 93

Table 12. Mann-Whitney U tests comparing nest (n = 24) and non-nest (n = 21) sites in terms of sub-pixel fractions extracted from Ikonos imagery..... 93

Table 13. Principal components created from the Ikonos imagery that showed significant differences between nest and non-nest sites..... 94

Table 14. Factor loadings for Ikonos principal components that showed significant differences ($p \leq 0.05$) between nest and non-nest sites..... 94

Table 15. Distance to habitat measured for 157 validation nest sites in St. Williams Forest 98

Table 16. Distribution of maximum likelihood *a posteriori* probabilities for validation nest sites..... 99

Table 17. Distribution of logistic regression predicted probabilities for validation nest sites..... 101

Table 18. Average DN of the three endmembers used in spectral unmixing of the Landsat imagery..... 101

Table 19. Mann-Whitney U tests comparing nest (n = 279) and non-nest (n = 63) sites in terms of spectral variables extracted from Landsat imager..... 103

Table 20. Mann-Whitney U tests comparing nest (n = 279) and non-nest (n = 63) sites in terms of textural variables extracted from Landsat imagery using a 3 x 3 window..... 104

Table 21. Mann-Whitney U tests comparing nest (n = 279) and non-nest (n = 63) sites in terms of sub-pixel fractions extracted from Landsat imagery..... 104

Table 22. Factors created through a PCA of spectral, textural, and sub-pixel fraction variables that showed significant differences between nest and non-nest sites..... 105

Table 23. Factor loadings for independent variables, extracted from Landsat imagery, using only those showing significant differences ($p \leq 0.05$) between nest and non-nest sites using 106

Table 24. Factor loadings for independent variables, extracted from Landsat imagery, using only those showing significant differences ($p \leq 0.10$) between nest and non-nest sites..... 106

Table 25. Error matrix for training (a) and validation (b) sites for the habitat/ non-habitat map created using the first set of principal components..... 108

Table 26. Error matrix for training (a) and validation (b) sites for the habitat/ non-habitat map created using the second set of principal components..... 108

LIST OF FIGURES

Figure 1. Male hooded warbler, with distinct black hood..... 10

Figure 2. Approximate breeding and wintering habitat of the hooded warbler. (Adapted from Friesen and Stabb, 2001)..... 11

Figure 3. Intense agriculture activity in southern Ontario, leaving fragmented forests 12

Figure 4. Typical hooded warbler nesting site in dense ground vegetation in St. Williams Forest (marked by orange stake in the ground) 15

Figure 5. Forest map showing nest sites and non-nest transects that were used as study sites within the three main forests. Sites within the area labelled as pine plantation were originally used but later discarded from the analysis..... 21

Figure 6. Tripod setup with digital camera and fisheye lens for acquiring hemispherical photographs..... 26

Figure 7. Sectors created using four zenith angles and six azimuth angles for analysis of the hemispherical photographs..... 28

Figure 8. An approximation of the spherical geometry of hemispherical photograph acquisitions using the 7.5° zenith angle..... 30

Figure 9. Examples of raw hemispherical photographs (a) and (c) along with their corresponding binary classified canopy/gap images (b) and (d)..... 31

Figure 10. Three examples of hemispherical photographs taken in the pine plantation showing elongated gaps (a and b), and a large north facing gap at a nest site (c), which was at the edge of a wide trail through the plantation 33

Figure 11. Non-nest area in a pine plantation. Dense needle litter can be seen with very little ground vegetation 33

Figure 12. Understory vegetation in wide trails and access roads within a pine plantation..... 33

Figure 13. Multispectral Ikonos image, shown as a colour infrared composite..... 62

Figure 14. Manual endmember selection from Ikonos imagery. Training areas for soil (brown/orange), shadow (blue), and green photosynthetic vegetation (green) are shown..... 68

Figure 15. Endmember selection from n-dimensional spectral space for soil (brown), shadow (blue), and green photosynthetic vegetation (green) 68

Figure 16. Approximate coverage of Landsat and Ikonos imagery shown along with a forest map, and the nest and non-nest sample locations..... 75

Figure 17. An example of non-nest sites with lack of suitable ground vegetation required for hooded warbler nesting..... 76

Figure 18. Final Landsat mosaic, shown as a colour infrared composite. Temporal differences are clearly visible in agricultural areas, with bare fields in the eastern half of the image and cropped fields in the western half..... 80

Figure 19. Examples of errors of omission (1) and commission (2) in forest classification, shown in a CIR composite (a) and the classified forest map (b)..... 82

Figure 20. Forested areas (green) with locations of nest and non-nest training and validation sites used for Maximum Likelihood supervised classification habitat/non-habitat. Outlined areas show significant forests across Ontario that were examined in more detail..... 85

Figure 21. Ikonos imagery (a) fraction images for GPV (b), soil (c), and shadow (d), created through constrained least squares linear spectral unmixing using endmembers selected manually from the imagery..... 89

Figure 22. Examples of hemispherical photographs and the corresponding Ikonos 1 m x 1 m panchromatic imagery. In the Ikonos images, bright areas are vegetation and dark areas are shadows, while in the hemispherical photographs the white areas are gaps and the dark areas are canopy..... 91

Figure 23. Supervised classification of Ikonos imagery, showing hooded warbler nesting and non-nesting habitat, along with training and validation sites..... 95

Figure 24. Close-up of St. Williams Forest showing areas classified as habitat and non-habitat. Areas outlined in red are classified as non-nesting habitat and that were previously unused for nesting. The black box indicates the location of the subset area shown in Figure 26..... 97

Figure 25. Subset area within St. Williams Forest, showing examples (circled) of validation nest sites classified as non nesting habitat that were located at the boundary of habitat/non-habitat..... 98

Figure 26. Probability curve corresponding to the binary logistic regression model (Equation 3) predicting the presence of hooded warbler nesting habitat... 100

Figure 27. Landsat imagery (a) fraction images for GPV (b), soil (c), and shadow (d), created through constrained least squares linear spectral unmixing using endmembers selected manually from the imagery..... 102

Figure 28. Examples of habitat mapping results for five well known forests across the study area: a. St Williams Forest, Backus Woods and South Walsingham Forest, b. Skunks Misery northeast of Chatham, c. The Pinery Provincial Park along the shore of Lake Huron, d. Dundas Valley near Hamilton, and e. Twelve Mile Creek near St. Catherines..... 109

Figure 29. Close up of St Williams Forest showing classification results created using training data from across the entire study area..... 112

Figure 30. St Williams Forest results created using training data only from the area corresponding to the high resolution Ikonos imagery. Areas outlined in red are classified as non-nesting habitat and were previously unused for nesting..... 113

Figure 31. Demonstration of discrepancies between a vertical photograph and remotely sensed imagery. (A) Vertical photography images small gaps in closed canopy while overhead sensor would image shadow (S) due to differential tree height. (B) A large gap would dominate a vertical photograph while the sensor would detect illuminated ground vegetation..... 115

1.0 INTRODUCTION

The protection of species at risk and the conservation of natural habitats is increasingly becoming a topic of concern in Canada. Human population growth and expansion, including the construction of buildings and roads, as well as the destruction of natural landscapes for resource extraction purposes, have resulted in a decrease in the remaining natural habitats for wild species (Environment Canada, 2003) and this is a major reason for the declining populations of many species. Canada's Species at Risk Act (SARA) follows the country's support for the United Nation's Convention on Biological Diversity, which outlines each country's responsibility to protect biological diversity within their borders.

The protection of species at risk involves not only enforcement against those persons found to harm the species in any way, but perhaps as important, the protection of their habitat. Whether it is nesting habitat for birds, spawning habitat for fish, or grazing land for caribou, without available habitat species cannot survive. The Species at Risk Act specifies that the habitat for each listed species be properly understood and then mapped, if possible, as part of the recovery efforts towards ensuring their long-term survival. Defining what is considered habitat for a given species involves understanding the species' use of the landscape through a description of the biophysical attributes of habitat that are required by the species, specifically the "geological, vegetative, topographical, climatological, physical, chemical, and biological" characteristics that constitute habitat (Environment Canada, 2004).

Hooded warblers (*Wilsonia citrina*) are currently listed as a threatened species under Canada's Species at Risk Act. They are an example of one of many species that

will disappear as a result of declining habitat if their habitat is not protected. The recovery plan for the hooded warbler requires that their habitat be delineated, but obviously before it can be mapped or modelled, the characteristics of the nesting habitat must first be more clearly understood. There has been a significant amount of previous work done examining nest site characteristics of the hooded warbler (e.g. Allair et al., 2002; Badzinski, 2003), and therefore their nest habitat requirements are fairly well understood. Unfortunately, previous attempts at mapping and modelling potential nesting habitat across southern Ontario (e.g. Pither, 1997; Flaxman; 2004) have perhaps been too simplistic and have not addressed forest canopy gaps, which may be a key forest structural characteristic for locating potential habitat.

Based on previous research by Whittam and McCracken (1999), Bisson and Stutchbury (2000), and Allair et al. (2002), it is fairly well understood that hooded warblers build their nests in dense ground vegetation in Carolinian Forests. While it is not clear if the birds make use of canopy gaps directly, they seem to indirectly require these gaps because the light that penetrates the canopy through the gaps allows the understory vegetation to grow extremely well (Rankin and Tramer, 2002), thereby providing suitable habitat.

This thesis is a collaborative effort between the Canadian Wildlife Service, as part of their responsibility towards protecting species at risk within Canada's borders, and the Geomatics and Landscape Ecology Research Laboratory at Carleton University. The research was initiated to examine the gap size distribution at hooded warbler nesting sites in greater detail than has been achieved in previous studies, and attempt to test advanced

remote sensing methods for mapping and modelling hooded warbler habitat, making use of spatial information relating to canopy heterogeneity.

Previous field research has shown the presence of canopy gaps above and around nesting sites. However measurements were done using ocular (sighting) tubes. This basic method of measuring percent cover is based on repeated observations of the canopy at a site through a piece of PVC pipe with wire cross-hairs (Huffman et al., 1999; Darwin et al., 2004). It does not provide information on gap size distribution, and perhaps is not even a very accurate measure of percent cover. In this thesis the associations of general characteristics of canopy gaps with hooded warbler nesting habitat are examined using more sophisticated hemispherical photographic methods. This information was then used to design a methodology for mapping potential habitat using remote sensing. Research has shown that spectral and textural information in remotely sensed imagery can be used to characterize forest canopy structure (e.g. Pellikka et al., 2000; Peddle et al., 2001; Cosmopoulos and King, 2004) and provide useful information on canopy gaps and forest heterogeneity. Using medium and high spatial resolution satellite imagery the potential exists for mapping canopy heterogeneity related to forest gaps across southern Ontario as indicators of hooded warbler nest habitat.

1.1 Objectives

1. To determine whether forests at hooded warbler nest sites have different overhead gap characteristics than those at non-nest sites.
2. If so, then determine if the satellite image spectral and spatial characteristics associated with nest and non-nest sites are also different.
3. If so, then use that image information to map specific forests of interest and to map forests with similar characteristics throughout the known range of the species in southwestern Ontario.
4. Evaluate the accuracy of maps of predicted nest habitat and non-habitat (where nests have never been found or where there is no ground vegetation for nesting).

Objectives 2-4 will be addressed at a local scale using Ikonos imagery for a forest near Long Point to evaluate the use of such imagery for high resolution spatial analysis of a targeted forest. The objectives were then addressed at a regional scale using Landsat imagery with the goal of producing maps of potential habitat for all of southwestern Ontario.

1.2 Thesis Structure

This thesis is divided into two parts, which each stand alone as unique research ventures, but are strongly linked in that the results from Part A drive the study carried out in Part B.

Part A is a comparison of canopy gaps found at nest sites and non-nest sites based on field measurements. The results from Part A were needed to justify continuing with Part B of this research. If a relation exists, then satellite images, along with methods to quantify canopy heterogeneity, can potentially be used to map nesting habitat.

Part B is the focus of the research and involves habitat mapping using remotely sensed imagery at two different spatial scales. This involved first the production of a local map of potential hooded warbler nest habitat for a targeted forest using high resolution Ikonos imagery. The same procedure was applied to medium resolution Landsat imagery to determine if the same habitat/non-habitat mapping accuracy could be achieved in such a targeted forest. Then, a regional map of potential hooded warbler nesting habitat across southern Ontario using Landsat was produced. Landsat was selected as it provides extensive coverage at very low costs. The spectral and spatial characteristics of the imagery were expected to reflect the canopy gaps measured on the ground. Between crown shadows, indicating canopy gaps, were hypothesized to reduce overall image brightness and increase image texture.

PART A
Habitat Field Study

2.0 INTRODUCTION TO PART A

Previous research into the characteristics of hooded warbler nesting habitat have shown various relations to forest structure, including a potential relation to overhead canopy gaps. Part A of this thesis addresses this specific habitat characteristic using quantitative analysis to determine if this relation does exist. This was accomplished through a field study comparing the canopy cover and gaps found at nest non-nest sites.

3.0 BACKGROUND

3.1 The Canadian Species at Risk Act

The Canadian government established the Species at Risk Act (SARA) on June 5th, 2003, nine years after it was originally tabled (Anonymous, 2003). In addition to prohibiting "killing, harming, harassing, capturing or taking individuals, possessing individuals or their parts, or destroying the residence of individuals", this legislation lays out a framework for actively protecting wild species within Canada, by preventing further extinctions through the recovery of endangered, threatened, extirpated and special concern species, as well as their habitat (Nadeau, 2003). A species is classified as endangered if it is facing imminent extirpation (no longer exists in the wild in Canada, but does exist elsewhere) or extinction (no longer exists on the planet), while the threatened classification is reserved for those species that are likely to become endangered if the situation is not changed. Those species listed as being of special concern have been found to be particularly sensitive to human activities or natural events. The Species at Risk Act defines critical habitat as being "the habitat that is necessary for the survival or recovery" of the species (SARA, 2005).

The Species at Risk Act relies on the Committee on the Status of Endangered Wildlife in Canada (COSEWIC), an independent body of experts set-up in order to determine the national status of wild species in Canada, based on scientific research along with Aboriginal traditional knowledge. COSEWIC was originally established in 1977 and is composed of government, academic, and private sector representatives. Under the new Species at Risk Act, for the first time the COSEWIC list has been given legal status as the official listing of species that are extinct, extirpated, endangered, threatened, and of special concern (COSEWIC, 2003). While recommendations are made by COSEWIC, the Minister of the Environment decides whether or not the species will be included on the SARA list.

Following the investigation into the status of a potential species at risk, if the species is found to be at risk, but not yet extinct or extirpated, COSEWIC will designate the species as being endangered, threatened, or of special concern. The classification of a species must be reviewed once every ten years, or at any time that the status of the species is believed to have changed significantly (Environment Canada, 2003).

According to SARA, once a species has been placed on the List of Wildlife Species at Risk, a recovery strategy and action plan must be established, which sets out goals and objectives in terms of population size and distribution, along with approaches for recovering the species, and identifying critical habitat, if possible. While the recovery strategies provide timelines, the action plans provide specific measures that will be undertaken in order to achieve population goals, as well as identifying ways to protect critical habitat and monitor the species over time (Nadeau, 2003). In the past these recovery plans have been put together and monitored by recovery teams, which are

composed of a group of people/organizations with an interest in the specific species. For example, in the case of the hooded warbler, the recovery team includes scientists from Environment Canada, Bird Studies Canada (NGO), researchers affiliated with universities, and private consultants. Unfortunately, due to the vast number of species added to the list over the years, this approach may not be possible in the future.

3.2 Critical Habitat

Since a species' population can decline simply based on a loss of habitat, it is extremely important to identify the amount and location of habitat that is critical to a species survival, and make every attempt to protect it. Locating the critical habitat within the species' range is an important part of protecting and recovering a species at risk, and producing a map of this habitat is an obvious solution in many cases, although in others it is not possible. While all critical habitat characteristics may be complex and difficult to completely understand, certain aspects of a species' critical habitat can be studied and mapped. For example, in the case of bird species, nesting habitat is generally studied and understood compared to the bird's territory, which is perhaps more complex, but is still included as a component of the species' "critical" habitat.

The habitat requirements of a species can be very simple, but are often extremely specific and complex. Taulman and Smith (2002) studied various bird species across North America and compiled a detailed list of criteria that describe the desired habitat of, for example, Kentucky warblers (*Oporornis formosus*) from numerous studies in the published literature. This specific bird "prefers to nest on the ground...in moist deciduous

forests with an abundant understory component...forest interior is the preferred habitat with minimum tract requirements of about 45 ha".

As part of the Species at Risk Act, it is important to understand the habitat of each species, particularly in order to map out currently used habitat, as well as map and model potential habitat.

3.3 The Hooded warbler (*Wilsonia citrina*)

The hooded warbler (*Wilsonia citrina*) (Figure 1), currently listed as a threatened species, is a small neotropical migrant, meaning that while it breeds in North America, it spends its nonbreeding season south of the United States (Hayes, 1995). It winters in areas of Mexico, Central America and the Caribbean, and breeds throughout the eastern United States and north into the Carolinian Forest Zone in southern Ontario (Figure 2) (Whittam et al., 2002; Bisson and Stutchbury, 2000).

Figure 1. Male hooded warbler, with distinct black hood.

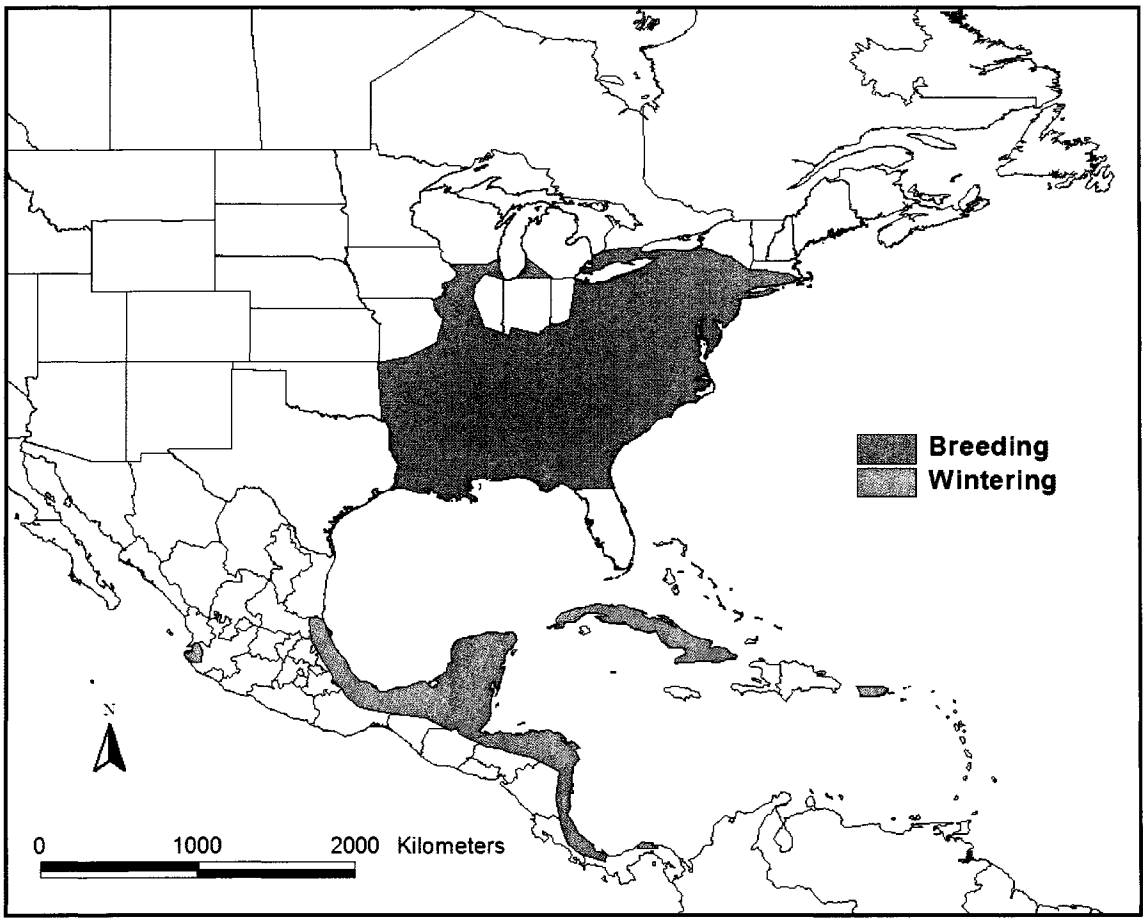


Figure 2. Approximate breeding and wintering habitat of the hooded warbler. (Adapted from Friesen and Stabb, 2001)

3.3.1 Population Status in Canada

Hooded warblers nest across the most southern parts of Ontario, predominantly in the Carolinian Forest Zone (Badzinski, 2003). The Carolinian Forests are unique to southern Ontario, but are simply northern extensions of similar forests in the United States. They are dominated by American beech (*Fagus grandifolia*) and sugar maple (*Acer saccharum*), and contain many other deciduous species found elsewhere in Canada. However, it is the many rare plants, found only in this ecosystem, including such species as Kentucky coffee tree (*Gymnocladus dioicus*), black gum (*Nyssa* sylvatica), black

walnut *(Juglans nigra)*, tulip tree (*Liriodendron tulipifera*) and sassafras (*Sassafras albidum*) among others, which characterize the ecosystem. They also contain many animals that are not found anywhere else in Canada, including the southern flying squirrel (*Glaucomys volans*), fowlers toad (*Bufo fowleri*) and rare bird species such as the hooded warbler (*Wilsonia citrina*), prothonatory warbler (*Protonotoria citrea*), and Acadian flycatcher (*Empidonax virescens*) (Carolinian Canada, 2005).

Over the years, the Carolinian forests in Ontario have become highly fragmented as a result of intense agricultural activity and urban expansion (Figure 3), and currently less than thirty-percent (measured in 1994) of the original forest cover remains in most counties (Allair et al., 2002).

Figure 3. Intense agriculture activity in southern Ontario, leaving fragmented forests.

While hooded warblers are abundant in the United States, due to the small population and disappearing breeding habitat, they were originally classified as a threatened species in Canada in 1994, and remained in this category following a re-assessment in 2000 (Page and Cadman, 1994; Friesen et al., 2000). Populations in Canada were estimated to be between 80 and 176 breeding pairs in 1992, but this increased to between 145 and 300 breeding pairs by 1997 (Friesen et al., 2000). Researchers have seen an overall increase in the number of breeding pairs since the intensive surveys began (Whittam and McCracken, 1999; Whittam et al., 2002). The most recent estimates showed approximately 240 pairs in 2002 and 280 pairs in 2003, approximately half of the 500 breeding pairs recommended to ensure a stable population (RENEW, 2004). Although some of the increased estimates may have been a result of an increased population, the intensive surveys across southern Ontario, implemented by the hooded warbler recovery team in 1997 (Allair et al., 2002) clearly contributed to such a dramatic increase. The abundance of hooded warblers in the United States does bring into question why they are listed as a threatened species in Canada given the thriving population to the south, and how can priority, in terms of time and money, be directed towards such a species. First of all, based on Canada's support for the United Nation's Convention on Biological Diversity, we are legally obligated to conserve "the full range of native biological diversity" within Canada (Friesen et al., 2000). As well, since hooded warblers are extremely vulnerable to human land use activities, efforts ensuring the recovery of the species should benefit many other forest species with similar habitat needs (Friesen et al., 2000). The Carolinian Forests in southern Ontario contain over 125 species listed by COSEWIC as being at risk of extinction, which is almost one-third of all

species at risk in Canada (Carolinian Canada, 2005). Finally, Friesen et al. (2000) point out that although the Canadian population of hooded warblers are not reproductively isolated, the population at the northern edge of the breeding range may possess important genetic or behavioural variations, compared with the more southern population. At the same time, one must keep in mind that there are opportunities for gene exchange and recruitment from bordering Great Lake states.

3.3.2 Habitat Selection

Hooded warbler nesting habitat characteristics have been widely studied and have become fairly well defined and understood. Whittam and McCracken (1999) cite U.S. research from the 1980s and early 1990s showing that hooded warblers avoid forest edges and colonize sites with a dense layer of shrubs (Figure 4) in mature forests. As well, they reference the research of others who found nesting sites to be associated with forest gaps.

Whittam and McCracken (1999) measured habitat characteristics surrounding nest sites, territory sites, as well as control sites (searched and found not to be used by hooded warblers) in St. Williams and South Walsingham Forests, in Norfolk County. They examined canopy cover (using a sighting tube), approximate canopy height, gap information, nest height, and dominant shrubs present at each site. Their results showed nests to be constructed in ground vegetation, at an average height of 0.60 m, which corresponded to later work by Allair et al. (2002) that showed nests to be constructed between 0.55 m and 0.61 m over a few years. The various shrubs commonly found around nesting sites mainly species of *Rubus spp.* As well, they found that although there

was no difference in canopy cover between nesting and non-nesting sites, there were larger gaps and more variability found within the hooded warbler's territory. They concluded that hooded warblers perhaps had a preference for small gaps around their nesting sites but did tolerate larger gaps within their territories, perhaps due to "increased foraging possibilities, a warmer microclimate, and increased visibility" compared to more contiguous forest. They did however point out that the larger gaps at the control sites, resulting in a denser ground cover, may have made it more difficult to find nests in these areas. Therefore, these apparent "control sites" may have in fact been occupied nesting habitat.

Figure 4. Typical hooded warbler nesting site in dense ground vegetation in St. Williams Forest (marked by orange stake in the ground).

Later research done by Bisson and Stutchbury (2000) showed further evidence that hooded warbler nesting sites had denser ground vegetation and lower canopy cover compared to random sites. Canopy cover was measured using a sighting tube (B. Stutchbury, *pers. comm.*, 2005).

The presence of dense ground vegetation has been attributed to the presence of canopy gaps or low canopy cover. Gaps in the canopy allow increased light to reach the understory, promoting denser ground vegetation than in areas with a more closed canopy. Rankin and Tramer's (2002) investigation into understory succession in current and former canopy gaps showed that some species, including *Rubus allegheniensis* (common blackberry), which is a very common species surrounding hooded warbler nests, were completely confined to gaps in their study area, and were in fact gap specialists. Their research showed that this species in the understory actually peaked in thickness during the fourth and fifth years following the formation of a gap, but over time they disappeared once woody vegetation overtook the understory. In general, their results showed that understory plant species exhibited higher total percent cover in gaps compared to beneath canopy.

Many studies have shown that hooded warblers actually prefer to nest in managed forests. Some cutting practices produce the necessary gaps in the canopy, allowing increased light to reach the understory, thereby producing suitable nesting habitat. Baker and Lacki (1997) studied the effects of different logging practices on bird populations, including the hooded warbler in Kentucky forests. They did preharvest surveys, then applied four different silvicultural prescriptions, followed by two years of postharvest surveys. Some forests were not harvested, while others were clearcut. As well, they used

two different levels of selective logging, leaving 7 m^2 residual basal area per hectare in some forests, and 3.5 m^2 residual basal area per hectare in others. Their results showed that for the two years following harvest, hooded warblers were more abundant in the selectively logged forests compared to the no-harvest and clearcut forests.

Annand and Thompson (1997) found similar results when looking at hooded warbler presence in southern Missouri hardwood forests in relation to different logging practices. Their results showed hooded warblers to be more abundant in forests that were selectively logged compared to clearcut forests and mature stands.

Rodewald and Smith (1998) also looked at hooded warblers nesting in Arkansas oak-hickory forests. They applied understory and overstory treatment to different forests, sometimes including heavy thinning of the understory vegetation in order to eliminate competition for desired tree species. Their experiments showed that hooded warblers made heavy use of the understory structure, and that the dense shrub layers created by the canopy openings were important habitat features, more so than the canopy openings themselves.

Whittam and McCracken (1999) summarized many of these earlier findings, and discussed the recent logging history of St. Williams Forest, one of the focal forests of this thesis, as well as the potential correlation between selective logging in the early to mid-1990s with the arrival of hooded warblers to the forest in 1995. It is clear that the gaps created by logging provided nesting habitat for the birds. They also pointed out that in the neighbouring South Walsingham Forest selective logging had not occurred since the mid 1980s, and interestingly by the late 1990s the population of hooded warblers was

declining, suggesting "that there is an optimal and limited time after cutting when a forest gap is suitable for hooded warblers".

Whittam et al. (2002) investigated the relation between selectively logged sites and hooded warbler nesting sites in St. Williams and South Walsingham. They measured habitat characteristics surrounding nesting and control sites (250 m x 250 m grid overlaid over St. Williams and a transect through South Walsingham). Their results showed that the hooded warblers chose nest sites with dense ground vegetation, and as well with more cut stumps around nests compared to control sites. They also found that the nesting sites were more likely to be associated with canopy gaps produced by the recent selective logging. Whittam et al. did point out that hooded warblers are not always found in logged areas, since they do have a niche in uncut mature forests where gaps are the result of the periodic death of trees, from wind, disease, or senescence. In their conclusions the authors of this research suggest that as a benefit to hooded warblers and in order to "potentially remove them from Canada's list of Species at Risk", forest managers across southwestern Ontario should incorporate single-tree or small-group-selection logging in order to create small canopy gaps, and at the same time leave the resulting shrub layer untouched. Unfortunately, using this type of management it is impossible to leave the understory vegetation completely undisturbed, however the least damage possible would obviously be preferable.

3.3.3 Recovery

A national recovery plan for the hooded warbler and Acadian flycatcher (*Empidonax virescens*), another migratory songbird sharing similar habitat, outlined the

current status of the species as well as recovery strategies and an implementation schedule (Friesen et al., 2000).

The main recovery goal for the hooded warbler is to increase the breeding population to 500 pairs, which is believed to be a self-sustaining level that would no longer be considered threatened. The recovery plan provides detailed strategies involving determining current population status and distribution, determining habitat requirements, protecting and enhancing breeding habitat, managing woodlands, and developing professional and public support through education and incentives. Ongoing research, mainly through Bird Studies Canada, investigating hooded warbler productivity and nesting characteristics (Badzinski, 2003), along with habitat selection (Whittam and McCracken, 1999; Allair et al., 2002) is currently the best hope for achieving these goals.

The Environment Canada website for the Species at Risk Act provides a listing of recent recovery activities for the hooded warbler (Environment Canada, 2003). Many of these efforts involve preventing the destruction of core breeding habitat for development purposes through the purchase of land from private owners, as well as through educating communities and local landowners.

As previously mentioned, a crucial part of protecting species at risk is the protection of their habitat. This research project falls under the recovery initiative that involves determining habitat requirements, and more specifically the initiative to map and monitor occupied and potential breeding habitat. The recovery plan states that this is to be carried out using aerial photography and satellite imagery combined with information on the species' habitat requirements (Friesen et al., 2000). Additional information and methods for mapping potential habitat for the hooded warbler would greatly help the

recovery team in protecting currently unprotected suitable habitat, and as well potentially provide target areas for the discovery of existing, but undiscovered, breeding pairs (J. McCracken, *pers. comm.*, 2004).

4.0 METHODS

4.1 Study Area

While hooded warblers are found in pockets across the Carolinian Forests of southern Ontario, the largest population exists in and around St. Williams Forestry Station, in Norfolk County along the north shore of Lake Erie (Figure 5) (Whittam and McCracken, 1999). This area also contains some of the largest remaining woodlots in southern Ontario, including South Walsingham Forest and Backus Woods. High spatial resolution satellite imagery was acquired over this area, and a map showing the exact location of the area is shown in Part B of this thesis (Ikonos Coverage in Figure 13).

Managed by the Ontario Ministry of Natural Resources (OMNR), the St. Williams Forestry Station, which is approximately 1200 ha, is recognized as an Area of Natural and Scientific Interest (Whittam and McCracken, 1999). It was originally prairie and oak savannah, with mainly aeolian fine sand soils that have formed dunes and ridges in some parts of the forest (Carolinian Canada, 2005). Most of the native Carolinian species were harvested in order to provide agricultural land. However, following failed agricultural attempts in the 1800s, mainly due to the highly unstable and eroding soils, most farms were abandoned and in 1908 the province of Ontario established the first provincial forestry station there in order to "provide nursery stock...and demonstrate the feasibility of reclaiming lands by reforestation" (OMNR, 2003). The planting of white

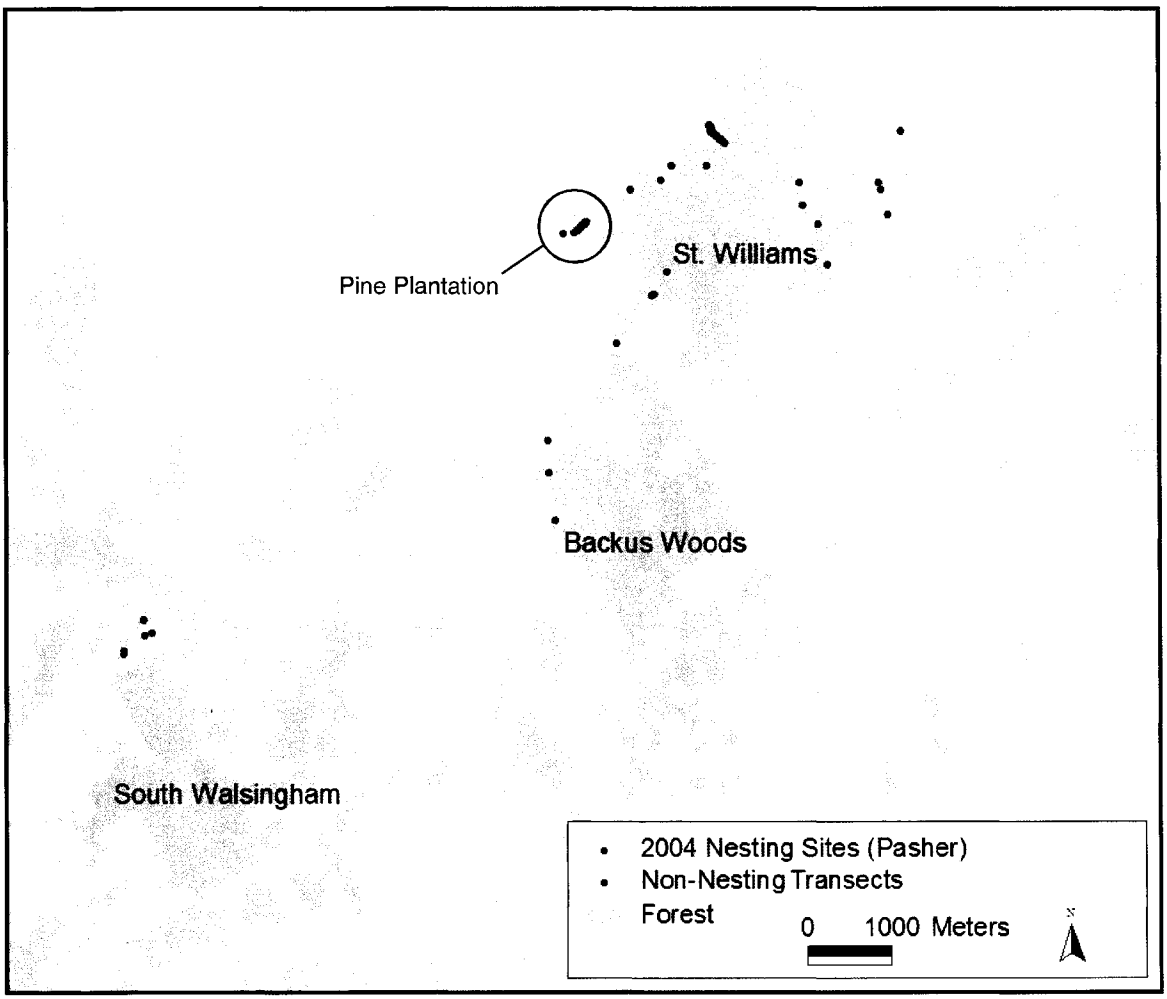


Figure 5. Forest map showing nest sites and non-nest transects that were used as study sites within the three main forests. Sites within the area labelled as pine plantation were originally used but later discarded from the analysis.

pine (*Pinus strobus*), red pine (*Pinus resinosa*), Scots pine (*Pinus sylvestris*) and jack pine (*Pinus* banksiana) was successful in stabilizing the soils in the failed agricultural fields. As the pine species grew, some of the native species remained mainly due to the presence of fire roads and selective logging, which permitted enough light to reach the forest floor (OMNR, 2003).

The Forestry Station was officially closed in 1998, following over eighty years of logging the planted pines for production of utility poles and saw logs (Whittam et al.,

2002). Most of the forest is currently protected in an attempt to conserve the "767 species of vascular plants, 139 breeding birds, 39 mammals, 13 amphibians, 12 reptiles, and a substantial, but incompletely known list of insects", many of which are considered at risk in Canada (OMNR, 2003). The forest is also used extensively for recreational activities such as hiking, bird watching, hunting, horseback riding, mountain biking and off road riding with motorcycles and ATVs (OMNR, 2003).

St. Williams is currently made up of approximately 26% deciduous, 29% coniferous, and 45% mixed forest. Deciduous trees include red oak (*Quercus rubra*), white oak (*Quercus alba*), black oak (*Quercus velutina*), American beech (*Fagus grandifolia*), large-toothed aspen (*Populus grandidentata*), trembling aspen (*Populus tremuloides*), black cherry (*Prunus serotina*), basswood (*Tilia americana*), white ash (*Fraxinus americana*), silver maple (*Acer saccharrinum*), sugar maple (*Acer saccharum*), and red maple (*Acer rubrum*). Some of the common understory shrubs found in the forest include blackberry (*Rubus* spp.), witch hazel (*Hamamelis virginiana*), American hazel (*Corylus americana*), red-berried elder (*Sambucus pubens*), eastern flowering dogwood (*Cornus florida*), cherry (*Prunus* spp.), poison ivy (*Toxicodendron radicans* s.l.), spice bush (*Lindera benzoin*), maple-leaved viburnum (*Viburnum acerifolium*), grape (*Vitis* spp.), wood fern (*Dryopteris* spp.), bracken fern (*Pteridium aquilinum*), dwarf chinquapin oak (*Quercus prinoides*), and hawthorn (*Crataegus* spp.) (Whittam et al., 2002).

South Walsingham Forest was not used for logging as extensively as St. Williams and is therefore dominated by deciduous trees, with an approximate composition of 75% deciduous, 3% coniferous, and 22% mixed forest. This forest contains very much the same vegetation species as St. Williams (Whittam et al., 2002), with 573 species of

plants, 102 breeding birds, 23 mammals, and 22 amphibians and reptiles (Elliott et al., 1999).

The third major forest found within this focus area is Backus Woods. It covers approximately 491 ha and is perhaps the best remaining example of Carolinian forest in Canada (Allair et al., 2002). Most of Backus is considered undisturbed old growth forest, a rarity amongst Carolinian woodlots. Backus Woods, like the other forests in the area, sits on eolian deposited sand dunes, with some dunes in the forest reaching up to 25 m in height (OMNR, 2004). The uplands in the forest contain large amounts of American beech, sugar maple, red maple, red oak, along with extremely large concentrations of tulip tree. Backus also has large swampy areas in the lowlands, which contain the highest concentrations of black gum known anywhere in Canada (Allair et al., 2002). There are several small pine plantations in the forest, however most areas have not been logged in over twenty-five years, with some of the areas having been left for over one hundred years (Allair et al., 2002).

4.2 Field Sites

The field data for this research were collected in July 2004 to correspond to full leaf-on conditions which were desired in order to reduce any temporal differences between field and image data (Part B of this thesis). As well, by this time, hooded warblers had returned to the area and in most cases had finished nesting and had moved on (D. Badzinski, *pers. comm.,* 2004), allowing access to new nests without disturbing the birds. The goal was to visit as many hooded warbler nest sites as possible within this local area to maximize the sample size for analysis of cover characteristics in relation to

non-nest sites. It was originally planned to include both 2003 and 2004 nest sites, since the canopy would likely not have changed significantly from one year to the next. Unfortunately, locations for the 2003 nest sites were not available. Experts from BSC and University of Toronto provided help in finding the 2004 nest sites in and around St. Williams Forest. Overall this provided a good distribution of twenty-five nest sites across the forests.

In addition to nest sites, non-nest areas were visited. The majority of the forests in the area provided suitable nesting habitat for the hooded warbler, so finding specific areas that were unsuitable for nesting was difficult (D. Badzinski, *pers. comm.,* 2004). The only definite non-nest area was found in the northern section of St. Williams forest. No nest sites had ever been found in this region, so twenty-one sample locations were selected along a 300 m transect through the middle of the area (Figure 5). They were selected and staked every 10 m up to 100 m, and then every 20 m up to 300 m. Varying distances between sample sites were used because the spatial autocorrelation of the forest was originally going to be examined as well.

Since hooded warbler nests are constructed in ground vegetation, it was obvious that any area of the forest that did not have ground vegetation could be considered non-nesting areas. The only other place where this existed was a pine plantation near St. Williams forest. A 200 m transect was set up through the plantation (Figure 5), with sites staked every 10 m up to 50 m, then every 20 m up to 150 m, with a last site at 200 m.

4.3 Hemispherical Photograph Acquisition and Processing

Objective 1 was to determine if nest and non-nest sites differed in total forest cover or in some other characteristics of the gap size distribution. To do this, hemispherical photographs were taken at each site and analyzed as described below.

Hemispherical photography has proven to be a useful method for measuring forest cover (Jennings et al., 1999; Pellikka et al., 2000). It was selected over the Li-Cor LAI-2000 Plant Canopy Analyzer (Li-Cor Inc., 1991), which is also a common method of calculating forest cover at a single location (Cosmopoulos and King, 2004; Seed and King, 2003), due to the ability of hemispherical photographs to provide permanent images of the canopy, which is very useful for visual examination of the gaps (Chen et al. 1997). Hemispherical photographs also allow for simple estimation of cover at different zenith angles simultaneously (Martens et al., 1993), which was of interest to this research.

Photographs were taken at each nest and non-nest site under diffuse sky conditions (dawn, dusk, and on days with low uniform cloud cover) in order to minimize direct scattering of sunlight off leaves and trunks (Olthof et al., 2003). A Nikon Coolpix 990 digital camera was used with a Nikon Fisheye FC-E8 lens. The exposures were manually set based on light meter readings of background sky in each photograph. The camera was mounted on a tripod at a height of 1.5 m within 0.5 m of the nest (Figure 6). It was then levelled and pointed north in order to ensure that each photograph had the same orientation.

A k-means unsupervised clustering algorithm was used to convert each photograph from a raw 24-bit colour TIF file into a binary canopy/gap image. Eight

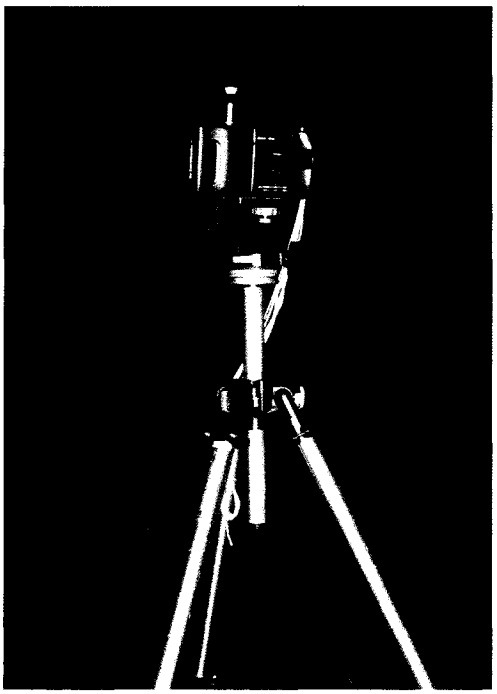


Figure 6. Tripod setup with digital camera and fisheye lens for acquiring hemispherical photographs.

clusters were extracted from the three bands of each photograph, and manually aggregated into either the "canopy" or the "gap" class, based on visual assessment of the imagery. Testing was done with different numbers of clusters, from four up to twelve, and it was found that eight provided an appropriate number to differentiate between canopy and gaps. Some authors point out that manual thresholding of hemispherical photographs can result in significant amounts of error, and therefore various automatic thresholding techniques have been developed (Ishida, 2004; Nobis and Hunziker, 2005). Recent work by M. Lindsay (*unpublished,* 2005) in similar deciduous dominated forests, however, has shown little difference in cover measured using automatic and manual thresholding and the correlation between the two methods was very high ($r = 0.92 - 0.99$, $p < 0.01$ for a range of zenith angles tested).

As a result of this, it was felt that the above mentioned semi-automatic thresholding technique provided sufficiently robust results, as opposed to subjectively selecting intensity thresholds for one or all three of the image bands (Frazer et al., 2001), or relying on a fully automated algorithm.

Hemiview Canopy Analysis Software v2.1 (Delta-T Devices, 1999) was used to calculate forest cover from each binary image by calculating the gap fraction (proportion of visible sky) for two zenith angles. Zenith angles of 7.5° and 22.5 were calculated to be consistent with the other methods such as the LAI-2000 Plant Canopy Analyzer, which uses 7° and 22° as its smallest angles (Seed and King, 2003). Only the 7.5° and 22.5° zenith angles were used because the analysis was focused on forest cover directly overhead and not obliquely. The photographs were also divided into six azimuth sectors, with 0° pointing North (bottom of photograph) to determine if forest cover in specific directions was related more strongly to nest site occurrence than in others (Figure 7).

4.4 Comparing Nest and Non-nest Forest Structure: Statistical Analysis of Forest Cover and Canopy Gaps

4.4.1 Forest Cover Analysis

The percent cover estimates at nest sites and non-nest sites were compared using non-parametric Mann-Whitney U tests. This statistical test shows whether two populations are significantly different based on overall ranking of the combined cases (SPSS, 2003), and was used in preference over a parametric Student's t-test for consistency throughout the thesis since many variables tested were non-parametric.

Cover was compared between nest and non-nest sites for each zenith-azimuth angle combination in Figure 7.

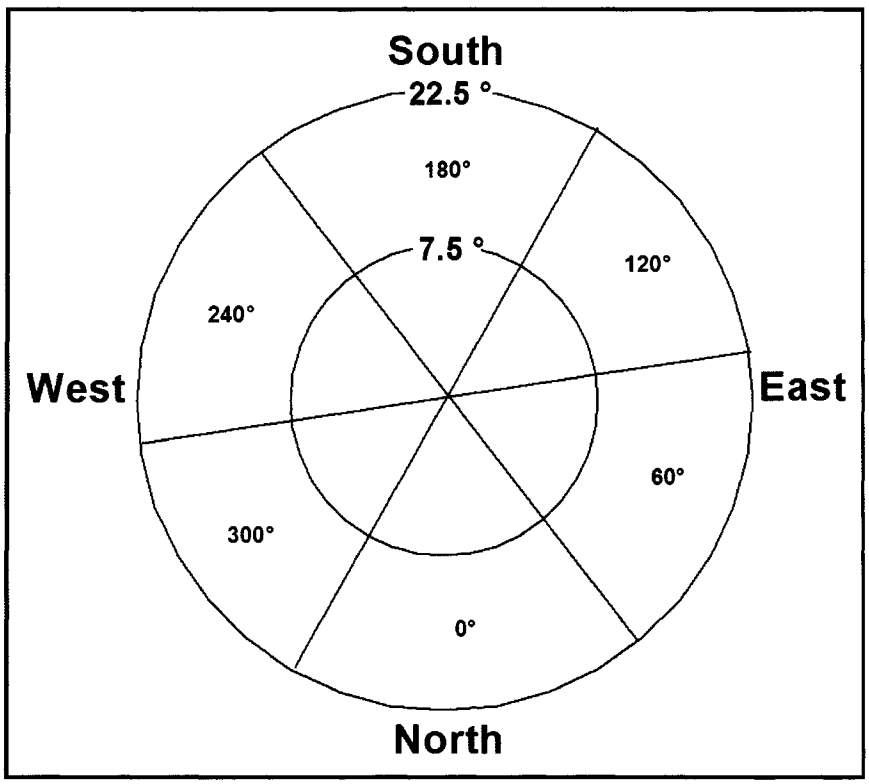


Figure 7. Sectors created using four zenith angles and six azimuth angles for analysis of the hemispherical photographs.

4.4.2 Analysis of the Gap Size Distribution

The size of overhead (7.5° zenith) gaps, the number of gaps and the average and maximum gap size at each site (7.5° and 22.5° zenith) were calculated and compared between nest and non-nest sites using Mann-Whitney U tests. Gap areas were determined by counting the number of pixels in each gap. In a more qualitative analysis, the histograms of the distributions of gap sizes for nest and non-nest sites were also compared.

While the actual gap sizes used by hooded warblers would definitely be of interest, especially with regards to forest management and selective logging, hemispherical photographs do not lend themselves very well to calculating the true size of the gaps represented in the images. Hemispherical photographs are projections of the hemisphere onto a plane (Jonckheere et al., 2004) and therefore areal distortions increase from nadir outwards. For this reason, the actual size of canopy gaps measured with the hemispherical photographs could not be accurately calculated, except using the smallest zenith angle, and even then it was an approximation.

Runkle (1992) provided guidelines for field measurements of canopy gap by approximating the gap area using ellipses. His methods were meant for calculating the size of a single gap at a site, while hemispherical photographs provide details on the entire range of gap sizes at a site. In order to calculate the approximate size of gaps from the hemispherical photographs, the areal coverage of each pixel in the image was needed. Given an approximate tree height in St. Williams of 23 m (Bisson and Stutchbury, 2000), the distance from the camera to the top of the canopy was approximately 21.5 m (Figure 8). For a zenith angle of 7.5°, the radius of the base of the cone of the field-of-view was calculated to be approximately 2.83 m giving an area of 25.17 m^2. With 45 763 pixels in the 7.5° zenith field of view, each pixel represented 0.00055 m^2. This value was used to approximate the average, standard deviation, and maximum gap sizes and allow comparison between nest and non-nest sites. To be conservative, since the 22.5° zenith did have some distortion problems, only the 7.5° zenith angle was used in order to calculate true gap areas, however gap sizes measured in pixels were determined for both angles. The maximum distortion is found at the edge of the field of view as a result of

this distance (the hypotenuse) from the lens being longer than the nadir distance. For the 7.5° zenith angle the hypotenuse is only 1.008X greater than at nadir, however, for the 22.5° zenith angle it is 1.08X greater. While the distortion is not consistent, there could be up to an 8% error in areal calculations if this angle was used.

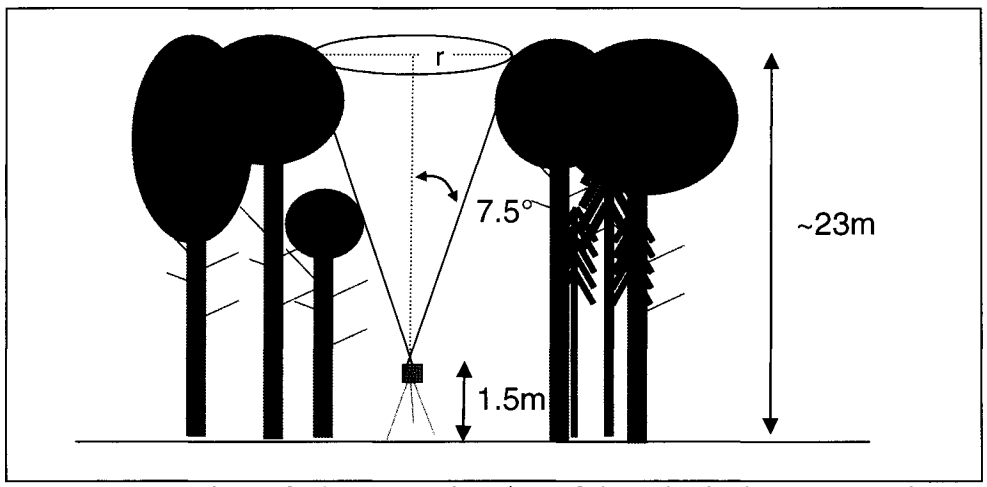


Figure 8. An approximation of the spherical geometry of hemispherical photograph acquisitions using the 7.5° zenith angle.

5.0 RESULTS

5.1 Hemispherical Photographs

Properly exposed photographs showed the canopy as completely black, while the gaps in the canopy, or open sky, appeared much lighter, but ranged from very pale blue to a deep blue/purple, depending on how much diffuse sunlight was available when the photograph was taken. The clear definition between canopy and open sky made thresholding each photograph into a binary canopy vs. gap image quite simple. Figure 9 shows two examples of raw hemispherical photos along with their corresponding classified canopy/gap images. The gap fraction was calculated for each of the azimuth

and zenith divisions using the Canopy Analysis Software, and these values were converted to percent cover for use in the research (1 - Gap Fraction).

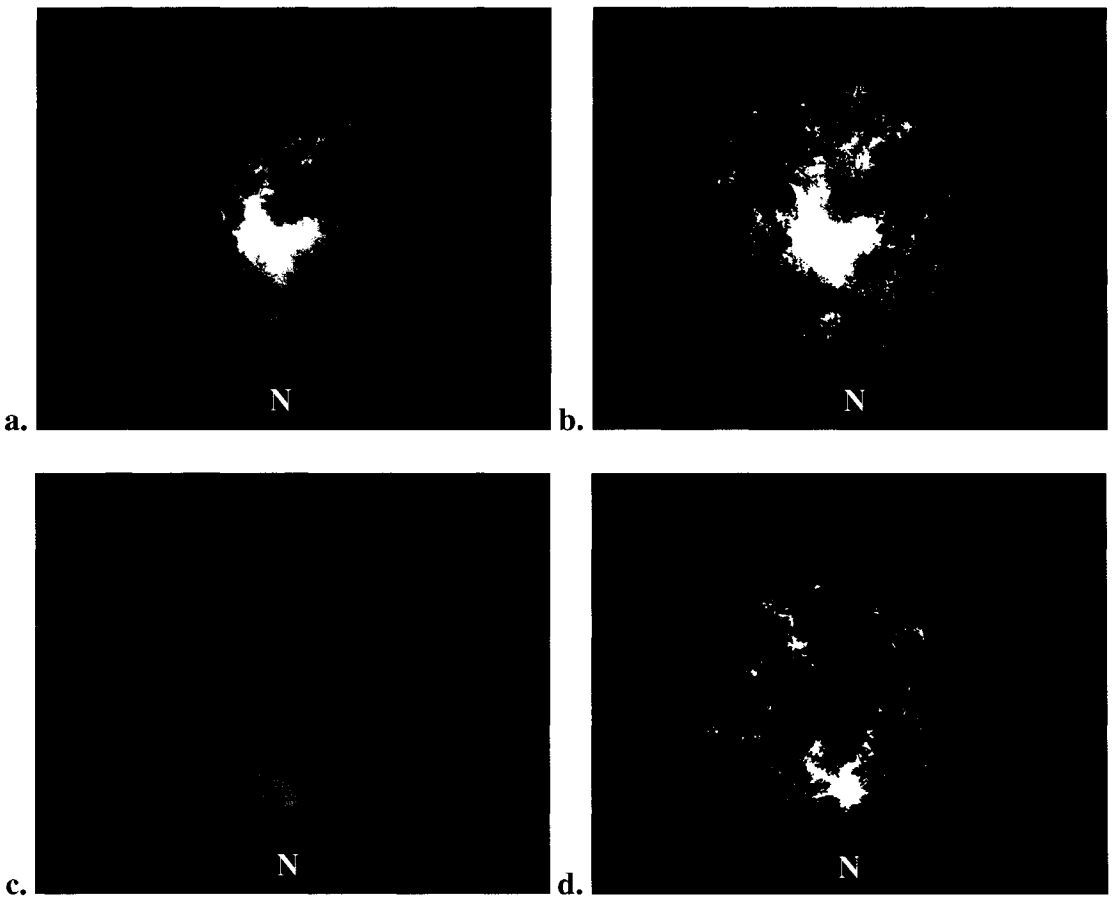


Figure 9. Examples of raw hemispherical photographs (a) and (c) along with their corresponding binary classified canopy/gap images (b) and (d).

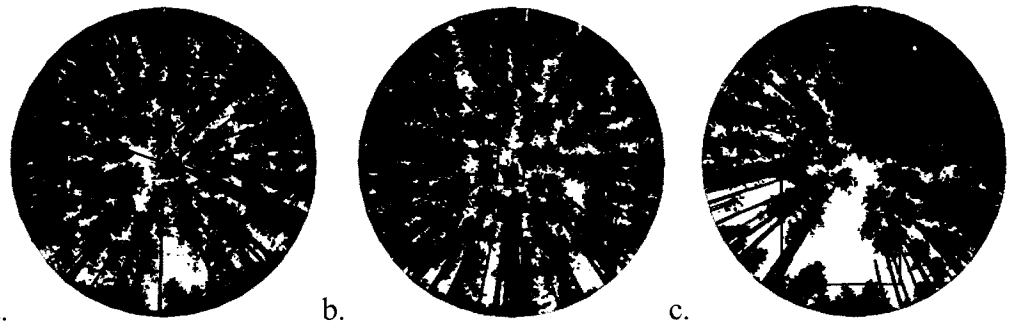
5.2 Statistical Analysis of Forest Cover and Canopy Gaps

5.2.1 Forest Cover Analysis

When examining the hemispherical photographs from the different sites it was noticed that there were large differences between non-nest sites in St. Williams and the pine plantation. The non-nest sites in the pine plantation contained many more gaps compared to St. Williams, and the gaps were elongated in many cases (Figure 10), reflecting the row planting pattern. When these two areas were compared to one another, the results (Table 1) showed highly significant differences, with much lower cover in the pine plantation. Given the two areas were clearly very different, and that there was a complete lack of ground vegetation throughout the plantation (Figure 11), except along the wide trails and roads (Figure 12 and Figure 10, (c)), the sites within it were dropped from the analysis. The focus was shifted to differentiating nest from non-nest sites in deciduous and mixed forests in this analysis and in Part B. For mapping purposes, pine plantations could be masked out from an analysis in order to focus solely on deciduous and mixed forests.

Table 1. Mann-Whitney U tests comparing overall percent cover of St. Williams (n = 21) and the pine plantation (n = 12) non-nest sites.

	St. Williams Non-Nest Sites		**Pine Plantation Non-Nest Sites**			
Sector	**Avg. (%)**	**Stdev. (%)**	**Avg. (%)**	**Stdev. (%)**	**U**	**P**
7.5° Zenith	83	10	66	6	19.0	**0.000**
22.5° Zenith	86	7	7	6	13.0	**0.000**



a. b. c.

Figure 10. Three examples of hemispherical photographs taken in the pine plantation showing elongated gaps (a and b), and a large north facing gap at a nest site (c), which was at the edge of a wide trail through the plantation.

Figure 11. Non-nest area in a pine plantation. Dense needle litter can be seen with very little ground vegetation.

Figure 12. Understory vegetation in wide trails and access roads within a pine plantation.

As a result of these observations all further analyses were carried out excluding the pine plantation data, which resulted in a reduction in the number of nest sites by one, and in the non-nest sites to those of the St. Williams transect only. Total percent canopy cover at nest sites was found to be significantly lower than at non-nesting sites for both of the zenith angles (Table 2). This was a result of strong differences in the north/north-eastern azimuths (0° and 60°), as other sectors did not show significant cover differences. Percent cover increased for both nest and non-nest sites as the zenith angle increased from 7.5° to 22.5°. This was expected as at a greater angle, the photographs included more information looking obliquely through the forest, including more tree trunks and sub-canopy branches, rather than vertically into the canopy.

Table 2. Mann-Whitney U tests for percent cover, without pine plantation sites, at nest (n = 24) and non-nest sites (n = 21), focusing on overall coverage and coverage by zenith angles, and azimuth sectors

	Nest Sites		Non-Nest Sites			
Sector	**Avg. (%)**	**Stdev. (%)**	**Avg. (%)**	**Stdev. (%)**	**U**	**P**
7.5° Zenith	68	26	84	10	159.5	**0.035**
22.5° Zenith	72	22	86	7	151.0	**0.022**
7.5° Zenith						
0° Azimuth	57	36	87	12	136.0	**0.008**
60° Azimuth	69	30	86	12	169.5	0.060
120° Azimuth	70	30	82	14	213.5	0.381
180° Azimuth	72	32	78	24	235.5	0.707
240° Azimuth	72	31	86	15	207.0	0.306
300° Azimuth	69	32	84	16	203.0	0.265
22.5° Zenith						
0° Azimuth	64	30	88	8	135.0	**0.008**
60° Azimuth	72	25	86	9	155.0	**0.027**
120° Azimuth	73	25	84	11	194.5	0.191
180° Azimuth	76	24	82	17	219.5	0.460
240° Azimuth	76	27	88	10	204.5	0.280
300° Azimuth	72	28	86	10	198.5	0.223

With significant differences found in the 0° Azimuth between nest and non-nest canopy cover, the question remained whether there was in fact a significant difference between cover in this direction and the other directions around the nest sites. One must keep in mind that both zenith angles tested are very small, and therefore while there is an azimuthal difference, it is still almost vertical. A Kruskal-Wallis one-way ANOVA was used and it was found that there was in fact no significant difference at either zenith angle (Table 3).

Table 3. Kruskal-Wallis test showing no significant difference between percent cover measured in northern direction compared to other directions around nest sites.

	0° Azimuth		Other Azimuths		
	Avg. (%)	Stdev. (%)	Avg. (%)	Stdev. (%)	P
7.5° Zenith	57	36	71	31	0.134
22.5° Zenith	64	30	74	25	0.133

5.2.2 Analysis of the Gap Size Distribution

The distribution of the size of the gaps at all sites, although very difficult to visualize, showed many more very small gaps compared to larger ones. The average size of the gaps was found to be significantly greater at nest sites measured by the 22.5° zenith angle (Table 4). The 7.5° zenith angle showed very different averages for the number of gaps. However, the variability was so great, making differences between nest and non-nest sites insignificant. No significant difference was found between nest and non-nest sites in terms of the average number of gaps present at either of the two zenith angles. The maximum gap size was found to be significantly larger at nest sites for both zenith angles. On average, the largest gap above nest sites was comprised of three to four-times the number of pixels as the largest gap at non-nest sites.

Table 4. Mann-Whitney U tests comparing nest (n = 24) and non-nest sites (n = 21) in terms of gap numbers and gap size measured from hemispherical photographs.

	Nest Sites		Non-Nest Sites			
	Avg.	**Stdev.**	**Avg.**	**Stdev.**	**U**	**p**
7.5° Zenith						
Number of Gaps	61.79	36.50	68.95	21.16	229.5	0.609
Average Size of Gaps (pixels)	2231.98	9278.72	111.54	74.34	189.0	0.152
Maximum Gap Size (pixels)	11957.61	12988.78	4092.46	3481.20	167.0	**0.053**
Maximum as a % of Total Area	26.13	28.38	8.94	7.61	167.0	**0.053**
22.5° Zenith						
Number of Gaps	236.46	107.83	255.71	61.97	233.0	0.665
Average Size of Gaps (pixels)	333.45	514.08	96.70	45.30	147.0	**0.017**
Maximum Gap Size (pixels)	34364.21	42507.34	8208.46	7622.05	135.0	**0.008**
Maximum as a % of Total Area	18.77	23.22	4.48	4.16	135.0	**0.008**

Using 0.00055 m^2 per pixel in order to approximate gap sizes for the 7.5° zenith angle, on average the nest sites had an average gap size of 1.23 m^2 (0 to 25.15 m^2) while the non-nest sites had an average gap size of 0.06 m^2 (0.01 to 0.16 m^2). These numbers, while useful, are obviously greatly influenced by the extremely large number of tiny gaps in the canopy. In terms of maximum gap size, although significant differences were found between the samples of nest and non-nest sites, the largest gap at nest sites ranged from 0 m^2 to 25.15 m^2 (mean = 6.58 m^2) while at non-nest sites the range was 0.08 m^2 to 5.80 m^2 (mean = 2.25 m^2). Thus, these measures are useful as descriptors of overall canopy conditions, but they should not be used to characterise an individual site as potential nest or non-nesting habitat.

6.0 DISCUSSION

6.1 Significant Findings

Prior to this study it appears that hemispherical photographs had yet to be used for quantifying canopy cover and canopy gaps in relation to bird habitat. Weiss et al. (1991) used them to quantify canopy structure and light conditions in relation to overwintering sites for monarch butterflies (*Danaus plexippus*), and Pringle et al. (2003) quantified canopy structure in determining thermal microenvironments available to the Broad-headed snake (*Hoplocephalus bungaroides*). With regards to investigating cover relations to hooded warbler nest site selection, all previous studies (including Whittam and McCracken, 1999; as well as Bisson and Stutchbury, 2000) used sighting tube methods. The research carried out in this thesis provided conclusive evidence that hooded warblers tend to build their nests in ground vegetation in the presence of overhead canopy gaps. The presence of overhead canopy gaps has been shown to affect the succession and success of various understory species, which take advantage of the increased light conditions (Rankin and Tramer, 2002). It is still uncertain whether hooded warblers specifically use the increased light for nesting or foraging purposes, but it is well known that the presence of ground vegetation needed for nest construction is dependent on increased light levels.

The lower canopy cover at nest sites that was found specifically towards the northern region of the sky was unexpected. It had been hypothesized that gaps associated with nest sites would be more to the south, resulting in more ground vegetation where direct sun illumination reaches the forest floor. Instead, the larger gaps in the northern azimuth and directly overhead provide greater diffuse sky illumination. This finding

remains unexplained and requires further research, perhaps specifically into effects of diffuse illumination on ground vegetation growth. It may also be interesting to examine canopy gaps specifically at a zenith angle relating to the sun angle during the growing season in order to examine the relation between understory presence with gaps.

With respect to the number of canopy gaps at the sites, the findings showed that while non-nest sites had approximately the same number of gaps as nest sites, they lacked the large canopy gap that was typically present at nest sites, suggesting that while highly correlated, the amount of cover is not necessarily as important as having a single large gap present.

This research is consistent with previous work (Baker and Lacki, 1997; Annand and Thompson, 1997; Rodewald and Smith, 1998; Whittam and McCracken, 1999) that shows how certain forest management practices are required to ensure the survival of the hooded warbler. Unmanaged forests do have canopy gaps and ground vegetation and therefore provide some of the necessary habitat for hooded warblers. However, it is clear that selectively logged forests, such as the St. Williams Forest, provide extensive nesting habitat. Robinson and Robinson's (1999) research in Illinois showed that creating canopy gaps through selective logging had little effect on the abundance of most bird species that would typically live in mature forests, while it did open up a significant amount of new habitat for gap dependent species such as the hooded warbler. They did, however, point out that the small gaps (0.02 to 0.04 ha) tended to be filled in within five years, and therefore continuous forest management through logging is required to maintain habitat for gap dependent species. Although it is extremely difficult to do, Rodewald and Smith

(1998) suggest that leaving the understory intact as much as possible in selectively logged areas is very important for providing bird habitat.

6.2 Research Limitations and Recommendations

Further research into gap characteristics above hooded warbler nest sites should attempt to involve a larger sample. This can easily be achieved if nest sets are continuously sufficiently marked and located using accurate GPS equipment and would allow researchers to visit previous year's sites. With twenty-four sites properly located from the 2004 field season, more nest sites from the 2005 season could be added to this database in order to increase the sample size.

While locations of previous years' nest sites recorded, the use of the coordinates was risky as their positions were not accurately known. Alternatively, one way to overcome this problem may have been to use "nest areas" instead of using one single location as the nest site. While the precise location of the nests could not be found, a radius relating to the approximate GPS error could have been used to define nesting areas. Hemispherical photographs could then have been taken at various locations throughout these areas and perhaps averaged in order to get measures of canopy cover for each nest.

As well, the number and spatial distribution of non-nesting sites was inadequate. Unfortunately, the local experts were unable to suggest more locations than the one visited in the northern portion of St. Williams Forest, where they were fairly certain hooded warbler nesting habitat did not exist. This transect was set up through a large area that definitely did not provide sufficient ground vegetation for the hooded warblers to

build nests, but finding other locations with a similar lack of ground vegetation proved extremely difficult. Many sites that were visited in the forest, which were thought to not contain the necessary habitat for the birds, were confusing even for experienced hooded warbler researchers (B. McLeod, *pers. comm.*, 2004) and therefore could not be used as definite non-nesting habitat sites.

It was not ideal to have all of the non-nesting habitat sites so close to one another and all in the same area of the forest. The concept of spatial autocorrelation, or spatial dependence, suggests that the observations taken from one area are more likely to be related to observations close by than those further away (Warner and Shank, 1997; Jennings et al., 1999). Jennings et al. (1999) pointed out that for measuring canopy cover, the spatial autocorrelation of crowns and canopy gaps must be taken into consideration, and sample sites must be separated by at least the size of the major spatial feature. Butson and King (1999) found that the range of the semivariogram for crown diameter in two hardwood plots dominated by sugar maple to be 12.2 and 9.6 m, suggesting that a spacing of at least 10 m was sufficient. While spatial autocorrelation and the size of tree crowns were not examined in detail for this area, a sample spacing of at least 10 m was used, which was felt to be adequate to avoid autocorrelation problems. In any case, in order to provide more well distributed non-nesting habitat sites, intense surveys of the forest would be needed that were not possible in this study.

While measurements of canopy cover taken using hemispherical photography are generally quite consistent and accurate, Jennings et al. (1999) suggest that hemispherical photographs are vulnerable to errors as a result of very small holes in the canopy, making thresholding of canopy from sky difficult. Thresholding the photographs can also be

difficult as a result of the exposure settings used for photograph acquisition, which if not set perfectly can cause fuzzy boundaries between canopy and sky. For this research, three photographs were acquired at each site using different exposure settings, and the clearest image was used for the analysis.

Hemispherical photographs taken looking upward do not truly measure gaps as seen from above the canopy. They are line of sight photography from a single location, and therefore provide information regarding the amount of diffuse light that penetrates the canopy at that location. While extremely low branches were pushed out of the way before the photographs were taken, any branches above approximately 2.5 m in height (within reach of photographer) that were not part of tree crowns, greatly influenced the photographs, and resulted in increased canopy cover. Ongoing work by E. Seed (*unpublished*) has shown that hemispherical photographs taken at different heights provide very different pictures of canopy closure as a result of branches. Therefore, while it was found that hooded warbler nest sites had lower canopy cover than non-nesting sites, this observation applies to both overstory and understory closure above about 2.5m. The implications of this with regards to the remote sensing analysis, which views the canopy from above, is discussed in detail in Part B.

Finally, it is important to note that while canopy cover was found to be significantly lower at hooded warbler nest sites, other variables are likely to influence whether or not the necessary ground vegetation will grow at a specific location. Petrone et al. (2004) mention that small scale variations in microtopography (i.e. depressions, slope, aspect) can greatly affect spatial variability of soil moisture within an ecosystem and the associated ground vegetation. For example, sand dunes in St. Williams Forest are

drier where elevated and wetter in depressions, resulting in varied ground vegetation species and amounts. Site conditions, vegetation competition and succession and light conditions all interact to determine where suitable nesting ground vegetation will develop.

PART B
Image Analysis and Habitat Mapping

7.0 INTRODUCTION TO PART B

The results from Part A of this thesis showed that hooded warbler nest sites had lower canopy cover and larger gaps than non-nest sites. With the link made between nest sites and canopy heterogeneity, along with the historical use of remotely sensed imagery for characterizing canopy structure, the potential was evident for using satellite images to map nesting habitat. Part B of this thesis addresses the issue of mapping nesting habitat, which is of great interest to the hooded warbler recovery team. High resolution Ikonos imagery was first used in order to investigate its potential for exploring canopy heterogeneity information in order to map hooded warbler nesting habitat at local scales in high detail. The same procedure was then applied to medium resolution Landsat imagery to determine if the same mapping success could be achieved in such a targeted forest. The Landsat imagery was then tested for identifying habitat across all of southern Ontario. Spectral, textural, and sub-pixel mapping methods are commonly applied in modelling canopy structure (Peddle et al., 1999, Pellikka et al., 2000; Peddle et al., 2001; Cosmopoulos and King, 2004) and therefore all three sources of remote sensing information were used in this research.

Greater canopy heterogeneity associated with nest habitat was expected to result in greater brightness variations in the imagery. Shadows exist in forest canopies as a result of the geometry and spacing of individual tree crowns (Seed and King, 2003), reducing mean canopy reflectance and increasing the spatial variation in reflectance. Image texture, or spatial heterogeneity, of the canopy is expected to be greater and mean image brightness is expected to be reduced where there are larger, or more frequent, gaps compared to more homogeneous canopy. This relation is expected to hold for gaps that

are not large enough to result in direct illumination of the forest floor of understory. For such large gaps, vegetation below the upper canopy can reverse the relation, causing increased brightness and reduced texture, as a result of the sensor directly detecting the illuminated understory. Sub-pixel methods can provide information on the fraction of the pixel that is shadow, which can potentially provide more information on the heterogeneity of the canopy.

8.0 BACKGROUND

8.1 Habitat Mapping

James and McCulloch (2002) suggested that modeling habitat in order to produce maps describing the "presence, abundance, or absence" of species across larger geographic areas is of great importance for managing species at risk. Since habitat models are simplifications of very complex biological systems, they are not expected to be perfectly predictive. However, the ability to develop explicit relations between species and habitat that are quantifiable and testable make habitat models very inviting for attempting to predict species presence (Van Horne, 2002).

There are many methods, and variations of each that exist for mapping suitable habitat for a given species. A few of the more common ones involve the supervised and unsupervised classification of satellite imagery, logistic regression modelling, GIS based multiple-criteria or decision rule analysis, and the Genetic Algorithm for Rule Set Production (GARP) approach.

Bechtel et al. (2004) made use of basic statistical clustering (ISODATA unsupervised classification algorithm) of Landsat pixels in order to attempt to link the

presence of woodland caribou (*Rangifer tarandus*) in northern Alberta to certain spectral classes. They examined the presence of the caribou within each of the twenty-five classes, taking into account the availability of the class simply being the proportion of that class within the study area. This analysis allowed the researchers to create a map showing classes, which were eventually grouped into useful land cover types through visual interpretation and field data, where caribou tended to live, as well as those that they tended to avoid. Although quite a simplistic analysis, for a large species whose movement is strongly related to vegetation cover, the research provided a method for mapping large areas used by the caribou.

The predictive power of regression modelling makes it an obvious choice for mapping potential habitat areas. O^2 Planning and Design (2003) used a common binary logistic regression analysis to map potential whooping crane (*Grus americana*) nesting habitat. A database of known whooping crane nesting sites was divided into a set for modelling and a set for validation. A set of independent variables was calculated from a Landsat-7 scene, including spectral and land cover information, land cover diversity information, and various spatially derived measures thought to be of importance to crane nesting preferences. Forward stepwise regression was used and the best-fit habitat model was then implemented in a GIS to produce a nesting habitat map with nesting probabilities associated with each cell. The independent test set showed that 93% of the known nest sites were in cells with probability values greater than 0.70. The map showed highly probable potential habitat outside of Wood Buffalo National Park, suggesting a need for expansion of park boundaries, or else protection of that habitat.

A similar modelling procedure was carried out by Mladenoff et al. (1995) to assess landscape-scale habitat characteristics of importance to wolves and estimate available habitat in Wisconsin and Michigan. Radio collaring had been done in the region for many years prior to the study resulting in a database of known wolf locations. This study involved five spatial data sets describing human population density, deer density, road density, land ownership, as well as land cover. A set of randomly chosen locations where packs had never been found was tested against known pack locations to attempt to isolate habitat characteristics. This comparison showed that wolf packs preferred mixed forests, little agricultural land, and large lakes. Road density was found to be much lower in pack territories than in nonpack areas, and interestingly deer densities, representing prey, were not different from pack to nonpack areas. All of the variables were entered into a stepwise logistic regression analysis to derive a model predicting wolf pack presence or absence. The best model showed a probability of wolf pack occurrence related to road density, correctly classifying thirteen of fourteen pack areas and twelve of fourteen nonpack areas.

Store and Kangas (2001) suggested a very different method of habitat suitability modelling using a four stage modelling procedure that was completely GIS-based. Habitat factors for the species of interest are defined, followed by assessment of the relative importance of the factors, and their effect on habitat priority. Next the appropriate GIS layers are created, most often from remote sensing based land cover maps. Suitable habitat areas are then modeled by combining the habitat factors, after which a sensitivity analysis demonstrating the effect of different weighting schemes is conducted. The end result of such an analysis is a habitat suitability map, made up of

habitat suitability index values for each map cell. A sensitivity analysis is crucial in this type of analysis as slight changes to some factor weightings can have significant effects on the suitability map.

The GARP approach (Stockwell and Peters, 1999) uses a set of rule based algorithms to find correlations between the presence-absence of a species and environmental variables. The process tests rules using training and validation data, and decides whether to include the rules in the overall model. It is growing in popularity because of its use of artificial-intelligence for combining several modelling methods into one. Many models can be produced, and combined into a single map showing the number of times specific pixels were predicted by the models as "presence" pixels.

8.1.1 Habitat Mapping and Modelling for the Hooded Warbler

Pither (1997) examined the use of satellite imagery and GIS to predict distributions of several bird species in the Carolinian Forest. He attempted to predict hooded warbler presence using spatial variables calculated from forest patches (i.e. patch area, patch shape, fractal dimension, etc...), along with soil types and forest composition. Binary logistic regression showed that for the hooded warbler, patch area as well as soil properties were significant predictors, however the resultant models were quite poor.

Flaxman (2004) criticized Pither's (1997) analysis, and was not surprised by the poor results, suggesting they were a result of the data used. Flaxman stated that using Landsat imagery, with 30 m pixels, could "not directly detect the small gaps required by the hooded warbler". As well, he pointed out that the vegetation classification was not detailed enough, and that ground vegetation or forest structure information was not

included. Instead he supported research done by Mitchell et al. (2001) who found overstory height and the number of understory vines to be more significant predictors of hooded warbler presence in South Carolina.

Flaxman (2004) performed habitat suitability modelling for the hooded warbler across southern Ontario. He used landcover data derived from Landsat imagery, but pointed out the limitations of using such imagery for resolving forest gaps, and went a step further to say that the approach he used, along with the data limitations, was likely to overestimate potential habitat. In any case, the spatial characteristics surrounding nest sites in terms of landcover type, patch size, landcover diversity, and terrain elevation were combined using an additive GIS approach. While the models he produced were able to explain the presence of a high percentage of hooded warbler nesting sites, there were high errors of commission; in other words the suitable habitat was overestimated. Unfortunately, it was not clear whether these were true errors, or simply a result of under sampling and a lack of proper absence data. He concluded by suggesting a need for better vegetation maps and higher resolution satellite imagery in order to better map hooded warbler nesting habitat across Ontario.

It was from the conclusions presented in Flaxman's (2004) report, along with personal communications with the author and members of the hooded warbler recovery team (K. Lindsay and J. McCracken, *pers. comm.*, 2004) that the objectives of this thesis project were defined. The need for higher resolution spatial information for mapping hooded warbler nesting habitat was clear. Therefore, high resolution Ikonos imagery was tested for its ability to provide information on forest canopy heterogeneity at the local scale for mapping and modelling habitat. While this was very useful for targeting small

areas across the landscape, given the extensive range of the species across southern Ontario, the use of high resolution imagery is not ideal due to very high costs. Therefore, the use of medium resolution Landsat imagery was deemed necessary to be able to cover the area. Sub-pixel information was investigated, as it was expected that canopy gaps at nest sites would show an increase in shadow and decrease in sunlit crown within the Landsat pixels. Amongst several methods that have been developed for sub pixel analysis, linear spectral unmixing was applied in this research. Its theory and application are described below.

8.2 Linear Spectral Unmixing

Due to the large extent of the range of the hooded warbler across southern Ontario, medium resolution sensors such as Landsat that provide regional coverage must be used. While these sensors do provide the coverage needed, the lack of spatial resolution has in the past meant a lack of information. Most common mapping analyses using Landsat imagery make use of hard-classifiers, to simply assign a single land cover class to each pixel in the image (Bastin, 1997). While this has proven useful in some cases of habitat mapping, more detailed information, such as canopy heterogeneity, which is important for the hooded warbler, may be lost if each forest pixel is simply classified as forest.

Working with medium and coarse resolution imagery brings about a very serious problem caused by the fact that the pixel size is often much larger than the objects on the ground that are being imaged (Settle and Drake, 1993). Therefore, the individual pixels in a Landsat image contain spectral information from a variety of different cover types that

have all been mixed together, making it difficult to extract information about a single cover type of interest (Oki et al., 2002). Hard-classifiers assign mixed pixels to the most likely or probable class, or leave them unclassified, which can result in a poor representation of reality (Settle and Drake, 1993), especially in ecological studies where many important habitat features might be small, and therefore lost within the mixed pixel (Bastin, 1997).

Soft classification algorithms were developed in order to deal with the classification of mixed pixels (Lillesand and Kiefer, 2000). This emerging field of remote sensing involves various methods of quantifying sub-pixel fractions, such as linear spectral unmixing, fuzzy classification (Bezdek and Pal, 1992; Foody and Cox, 1994; Bastin, 1997) and Maximum Likelihood *a posteriori* probabilities (Foody et al., 1992; Bastin, 1997). Previous exploratory research was carried out in order to assess the ability of each of these three methods for estimating sub-pixel land cover proportions from within Landsat pixels (J. Pasher, *unpublished*). The sub-pixel fractions were compared with a classified Ikonos image in order to determine which method best matched the reference data, and it was found that linear spectral unmixing showed the most potential. It was therefore selected for this thesis as a method for quantifying forest canopy heterogeneity within Landsat imagery.

The process of Linear Spectral Unmixing, or Linear Mixture Modelling, works under the assumption that each photon interacts with only one object before it is reflected back towards the sensor (Bastin, 1997; Van der Meer and de Jong, 2000), as opposed to multiple scattering non-linear unmixing models (Foody et al., 1997, Liu and Wu, 2005). Given the "linear" nature of the procedure, the reflectance values associated with a mixed

pixel are assumed to simply be linear combinations of the reflectance of the constituent ground covers that are within the bounds of that pixel (Klein Gebbinck, 1998; Brown, 2001). Since the reflectance of a coarse resolution pixel is often made up of more than one cover type, or scene component, spectral unmixing can be used to decompose the mixed spectra in order to determine sub-pixel proportions of each (Peddle et al., 2001).

Unmixing can be accomplished, just like any other method of supervised classification, by first selecting training data. Normally, for hard classifiers, such as Maximum Likelihood, training pixels are chosen to represent the variability in each land cover type of interest across the image (Lillesand and Kiefer, 2000), which most often includes many mixed pixels. However, the training classes to be used for spectral unmixing, also known as endmembers, are assumed to represent the reflectance of a pixel that contains nothing but the land cover type, or image component, of interest (Klein Gebbinck, 1998), and therefore must be composed of only the purest pixels of each class.

Endmember reflectance spectra are typically acquired using one of two methods: field spectroradiometry, or by selecting pure pixels from within the image (Peddle and Johnson, 2000). While easier to use with hyperspectral imagery, measurements of forest canopies using field spectroradiometry would be very difficult to relate to multispectral imagery due to the wide bands and therefore lost spectral resolution of the imagery. Since they are the basis of the spectral unmixing procedure, their quality and selection is crucial (Peddle et al., 1999). Calculating endmembers directly from the field would require calibration to match the spectral bands and dynamic range of the imagery being used for unmixing. As well, while it is perhaps easy to collect field based spectra for crops and other fairly uniform land cover types, it is extremely difficult to use in forests due to the

three dimensional complex structure made up of vegetation and shadows. Selecting the endmembers straight from the image can help avoid calibration (Van der Meer and de Jong, 2000). However, it is difficult to find pure samples in a coarse resolution image (i.e. in AVHRR imagery one would need pure examples of 1 km x 1 km pixels). Higher resolution imagery and air photos, along with ground observations, can be useful in finding pure areas in a coarse resolution image, or the endmembers can be selected directly from high spatial resolution imagery, and calibrated to the lower resolution imagery to be used for unmixing.

Most studies make use of manual endmember selection methods, however, researchers at the Canadian Centre for Remote Sensing recently developed an automatic endmember selection algorithm, using an Iterative Error Analysis (IEA) process. This algorithm is now incorporated in PCI Geomatica v.9.0, in a new package of hyperspectral processing tools (PCI Geomatics, 2004). This method performs a series of linear unmixing procedures using endmembers (at the start they are simply arbitrarily defined), which are changed after each iteration as the error is reduced (Levesque, 2001). This continues until a desired error threshold is reached or a specified number of endmembers is reached that satisfy the set parameters. The resulting endmembers can be used directly in a spectral unmixing algorithm. However, the land cover class or a scene component that they represent must first be investigated.

Based on the assumption of linear combinations (Equation 1, Brown, 2001), and a given a set of endmembers, mixed pixels within an image can be unmixed, and the proportions of land covers within the mixed pixel can be solved for using various

decomposition methods. These methods solve the system of linear equations given the various spectral bands and endmembers.

$$DN_\lambda = DN_{1\lambda} * f_1 + ... + DN_{n\lambda} * f_n + E_\lambda, where \sum f_n = 1 \qquad [1]$$

where, DN_λ is the reflectance in band λ, $DN_{n\lambda}$ is the reflectance for endmember n in band λ, f_n is the fraction of the pixel that is of endmember n, and E_λ is the error for the particular band. The error term in this equation incorporates error from an incomplete set of endmembers for a pixel; in other words, there are other spectral components present, as well as errors attributed to non-linear mixing. Although vegetation does produce non-linear scattering of photons due to the multiple reflections and transmissions of light between leaves, bark and soil (Borel and Gerstl, 1994), the high complexity and computational demands involved in non-linear methods have hindered their application (Levesque, 2001). A constrained least squares linear unmixing procedure was used in this research as it is simple to understand and has been commonly assessed in the literature (e.g. Holben and Shimabukuro, 1993; Peddle et al., 2001). Constraints were used in this research specifying that the unmixed fractions must be positive, and the sum of the unmixed fractions for all endmembers must equal one. The least squares unmixing method works by iteratively minimizing the sum of squares of the errors, based on the training pixels used (Holben and Shimabakuro, 1993).

The result of solving such equations, and the goal of spectral unmixing, is a set of fraction images, one for each endmember, as well as an error map. Each pixel in a fraction image provides the sub-pixel proportion of each endmember. This sub-pixel

information can be very useful compared with traditional hard classifiers that only provide a single thematic map with a single class assigned.

Spectral unmixing has generally been used, and works best, with hyperspectral data (Brown, 2001). However, it is quickly becoming a commonly used procedure for multispectral imagery. A limitation of spectral unmixing, cited throughout the literature is that the number of endmembers, and therefore the number of classes to be unmixed, is dependent on the number of spectral bands. Some researchers say that given N spectral bands, the number of endmembers must be less than or equal to N (e.g. Eastman, 2001; Brown, 2001; Pu et al., 2003; Rogers and Kearney, 2004). On the other hand, others say that N+1 endmembers can be used (e.g. Klein Gebbinck, 1998; Levesque and King, 2003). For hyperspectral data, which have many narrow spectral bands (e.g. CASI-3 uses up to 288 Spectral Bands; (ITRES, 2003)) this is not an issue as a huge number of endmembers can be used, resulting in information on many different land covers. With multispectral data such as Landsat, there are only six commonly used bands and therefore the number of endmembers is limited.

Much of the research done using spectral unmixing does not actually focus on unmixing land cover types, but rather radiometric scene components, of which there are usually only a few. It is this area of unmixing that is to be explored for mapping potential hooded warbler habitat, as the vegetation and shadow components within pixels is of great interest. For example, Peddle et al. (1999; 2001) used unmixed fractions of scene components from airborne imagery as predictors of forest biophysical and structural parameters. They expressed boreal forest reflectance as a linear combination of sunlit canopy, sunlit background, and shadow. Selection of endmembers was accomplished

using a field spectroradiometer, along with image data and reflectance models. When compared with ten standard image-based vegetation indices for the prediction of six forest parameters (biomass, net primary productivity, leaf area index, dbh, stem density, and basal fraction), they found that shadow fraction usually had the highest correlation coefficient.

Levesque and King (2003) also used spectral unmixing to investigate sub-pixel fractions from high resolution multispectral imagery in modelling tree crown and canopy health and structure. Unmixed fractions of mine tailings, woody debris, shadow, and vegetation, as well as measures of spatial variation of those fractions were used. Spatial measures of texture and semivariogram range derived from the fraction maps were found to be the best predictors of tree crown and canopy variables, respectively.

While the previous examples make use of high resolution imagery, unmixing has been applied to medium resolution Landsat imagery. For example, Oki et al. (2002) made use of multitemporal Landsat TM data in order to provide better estimation of alder coverage in a wetland region of Japan, compared to a hard classification method. Pu et al. (2003) also used Landsat TM data to unmix soil, grass, and oak crown, in order to calculate oakwood crown closure in a study site in California. These two examples provide evidence of the successful use of linear spectral unmixing for extracting sub-pixel fractions for medium resolution multispectral data. Although the methods were developed for use with hyperspectral data, when the number of endmembers are kept small, spectral unmixing can work reasonably well with Landsat data (Brown, 2001).

9.0 METHODS

This chapter has been divided into two parts in order to reflect the two scales of remote sensing analyses that were carried out for mapping hooded warbler habitat. The order of the analyses is a result of the different scales examined, first looking at field based hemispherical photographs (Part A), then linking the photographs to the high resolution imagery, and then examining the high resolution imagery for habitat mapping for a small target area which reflected the local coverage of the imagery. Following this the analysis was scaled-up to examine the same small target area using medium resolution imagery and finally mapping habitat at the landscape level using the medium resolution imagery.

9.1 Compilation of Hooded Warbler Point Location Database

A database of hooded warbler occurrence across Ontario was compiled from a variety of sources to serve as reference data for model building and validation. It consisted of nest sites visited during fieldwork in July 2004 (n = 25) along with nest sites collected by S. Melles (PhD research, in progress, University of Toronto) over two field seasons (n = 56). Bird Studies Canada provided additional data from three sources. First, nest locations within St. Williams Forest from 2002-2003 were provided from a small database for that forest (n = 89). Second, 2002-2003 hooded warbler records were extracted from their Vulnerable, Threatened, and Endangered (VTE) birds database (n = 338). This database includes positional information, collected through biological surveys, naturalist groups and volunteers, for birds in the "Carolinian Life Zone" in southern Ontario (BSC, 2005). Third, they provided data from the Ontario Breeding Bird Atlas

(OBBA), which was set-up in order to gather breeding distribution data for all bird species in Ontario. This project, running from 2001 to 2005, involves volunteer surveys of assigned 10 km x 10 km areas across southern Ontario in order to compile a listing of species observed in each square. The participants record detailed information when they encounter significant species, such as the hooded warbler (OBBA, 2005). Positional locations of breeding evidence are recorded through various methods. Data from 2002 to 2004 were acquired from the atlas for this research (n = 278).

All five data sources were aggregated and all duplicate records were eliminated, along with a few records that did not specify the method of recording the position, or if they specified a method other than GPS (e.g. approximation from a topographic map). This ensured a high level of positional accuracy and consistency throughout the database.

The OBBA and VTE databases included records that were not specifically nest sites, and in some cases did not indicate breeding evidence. One of the objectives of the atlas is to obtain the strongest possible evidence of breeding, and it provides four levels of evidence: 1. Species observed in breeding season (no evidence of breeding), 2. Possible breeding, 3. Probable breeding, and 4. Confirmed breeding (OBBA, 2003). For the purpose of this research only the highest two levels were used (K. Lindsay, *pers. comm.,* 2004) to provide strong evidence of breeding habitat.

Of the 321 records of probable and confirmed breeding evidence for hooded warblers across southern Ontario (out of 606 unique records), 181 were located within the forest mask created from the Ikonos imagery, while 279 of them were located within a forest mask created from the Landsat imagery (see 9.2.3 and 9.3.3 for forest masking methods for Ikonos and Landsat respectively).

9.2 Mapping with Ikonos Imagery

9.2.1 Study Area

The same 15 km x 13 km area used as the study area in Part A was used for this portion of the research. This area, which contained several forests known to have large hooded warbler populations (St. Williams Forest, South Walsingham Forest, and Backus Woods), was used for testing nesting habitat classification using high resolution Ikonos imagery. The location of this area, relative to the area covered by the Landsat imagery, is shown in Figure 13.

9.2.2 Field Data Collection

The field data consisted mainly of the GPS positional measurements of locations of nest and non-nest sites, which were collected in Part A of this thesis. Unfortunately, within the Ikonos coverage, the only non-nest sites consisted of the St. Williams transect previously discussed.

The GPS measurements were taken at each site using a Trimble GeoXT GPS, which provided real-time sub-meter accuracy (Trimble Navigation Ltd., 2004). As the forest was dense, the GPS external antenna was attached to the top of a 2 m pole, which allowed it to communicate with satellites that were lower on the horizon. For nest sites, the exact location of the nest was desired. However, in order to prevent any damage to the nests or the vegetation in which they were built, the locations were recorded within 0.5 m of the nests, still providing the necessary accuracy to relate the sites to the high resolution satellite imagery.

Additionally, positions and hemispherical photographs of five additional sites with relatively low cover were acquired in the study area. These sites were not considered as nest or non-nest sites, but were simply selected to increase the data range for statistical modelling of percent cover (Section 9.2.7) against image variables.

9.2.3 Image Acquisition and Initial Processing

Cloud and haze-free Ikonos-2 data were acquired July 3, 2004 for a 15 km x 13 km area covering the study area. The acquisition fit very well temporally with the research as the image was acquired during the fieldwork period, allowing for little change in canopy conditions between field and image measurements from the forests.

The data consisted of four multispectral bands (Blue (0.44 – 0.52 μm), Green (0.51 – 0.60 μm), Red (0.63 – 0.70 μm), and NIR (0.76 – 0.85 μm) (Space Imaging, 2004)) each with ground pixel size of 4 m x 4 m, and a panchromatic image with a 1 m x 1 m pixel size. The data were delivered in two overlapping sections since the width of the acquisition was greater than the 11.3 km swath width of the Ikonos sensor (Space Imaging, 2004). Both sections were projected to UTM Zone 17T NAD83 datum.

As the two overlapping sections were slightly offset from one another and had slight reflectance differences, a mosaic was created using fifteen tie points (all tie points and ground control points are shown in Appendix 1) from both the panchromatic sections and multispectral sections independently of one another. For the panchromatic and multispectral mosaics, first order nearest neighbour transformations were used to achieve an RMS error of (0.29, 0.24) and (0.47, 0.27) pixels respectively. PCI's OrthoEngine

(PCI Geomatics, 2004) automatic histogram matching was used in order to overcome the slight difference in brightness between the sections.

The panchromatic mosaic was then georeferenced using a first order nearest neighbour transformation, in order to achieve an RMS error of (0.58, 0.82) pixels, using seventeen GCPs that had been collected at road intersections throughout the image. The multispectral mosaic was georeferenced to the corrected panchromatic mosaic using a first order nearest neighbour transformation with thirty-three tie points, resulting in an RMS error of (0.43, 0.52) pixels. The multispectral Ikonos image is shown in Figure 13.

All non-forest in the Ikonos imagery was masked out using K-means clustering (PCI Geomatics, 2004) with all the spectral bands as inputs and generating forty clusters, which were manually aggregated to produce a forest/non-forest binary mask.

9.2.4 Extraction of Spectral and Textural Information

Image spectral and textural information was used in order to characterize the heterogeneity in the forests, and attempt to show differences between nest and non-nest sites. Spectral information was extracted from the Ikonos data as the average DN in a 5 x 5 (20 m x 20 m) window around each site. This window size was selected because it corresponded to the approximate view area (17.8 m diameter) in the canopy that was covered by the hemispherical photographs within a 22.5° zenith angle.

Figure 13. Multispectral Ikonos image, shown as a colour infrared composite.

Image texture is best described as the change in tone across an image, and is a measure of the overall visual "smoothness" or "coarseness" of the image (Lillesand and Kiefer, 2000). For this research, forest canopies with more gaps or larger gaps that correspond to nest sites (as found in Part A of this research) were expected to have greater texture than non-nest sites with more closed and homogeneous canopy. Many researchers have shown that textural information calculated from the grey level co-occurrence matrix (GLCM) (Haralick et al., 1973) is useful for forest classification. Chan et al. (2003) and Coburn and Roberts (2004) working with Landsat imagery, and Tso and

Olsen (2004) working with Ikonos imagery, all made use of various GLCM textures (e.g. Homogeneity, Contrast, Dissimilarity, Entropy, Angular Second Moment, and Correlation) along with spectral information to improve classifications. The GLCM is derived by tabulating the frequency of all possible grey level pairs for pixels separated by a specified distance within a specified window. Pixels in more homogeneous areas have more similar DN values, compared to more heterogeneous areas. For this study, the six GLCM textures listed above were calculated for the Ikonos (5 x 5 window) imagery using sample pixels adjacent to each other in all directions (the imagery showed no systematic directional gradients). Textures were not calculated for the blue band as it was not thought to contain as useful canopy heterogeneity information such as the green, red, and NIR bands might provide due to the very narrow range of values. Appendix 2 gives their formulation. Homogeneity is a measure of the degree of local homogeneity, while Contrast is the opposite of Homogeneity in that it measures the amount of local variation. Dissimilarity is a very similar measure to Contrast. Entropy measures the organization within the co-occurrence matrix in terms of the presence or absence of patterns. Angular Second Moment is the opposite of Entropy in that it measures the uniformity of grey level values. Finally, Correlation measures the linear dependency of grey levels in neighbouring pixels (Haralick et al. 1973, Hall-Beyer, 2004, Jensen, 2005).

In addition, the average DN values in a 5 x 5 pixel window, along with textures calculated from the same sized window were extracted from the Ikonos panchromatic image.

9.2.5 Linear Spectral Unmixing

A constrained least squares linear spectral unmixing algorithm (PCI Geomatics, 2004) was used as a means of accessing sub-pixel land cover information. Details of spectral unmixing are given in Section 8.2. With everything except forest areas masked out in the imagery, the selected endmembers had to account for all the sub-pixel fraction components that would be present within forest pixels. Land covers such as agricultural crops, water, urban, and any other cover type not expected in forests were ignored. While the number of endmembers was limited to six by the number of input spectral bands, only those representing vegetation, soil, and shadow were extracted.

Often when carrying out linear mixture modelling in forest environments four endmembers are used, consisting of green photosynthetic vegetation (GPV), non-photosynthetic vegetation (i.e. bark and branches), bare soil, and shadow (Asner et al., 2003). For this research, the first, and most important, endmember used was green photosynthetic vegetation. While many coniferous trees were found within the forests, only deciduous trees were used to provide endmember spectra since their pixels were at a well defined vertex in spectral space (explained further below), while coniferous trees were found to be positioned in the gradient between the deciduous and shadow vertices, showing that they were a mixture of the two spectra. This was thought to be a result of their inherent intercrown shading caused by their conical shape.

Originally it was proposed to make use of the potentially different spectral signatures of the ground vegetation in order to isolate patches within the forest that may indicate potential nest habitat (e.g., as Linderman et al. (2004) did using Landsat TM imagery for estimating understory bamboo presence in forests in western China).

However, in this imagery, ground vegetation, which was generally leafy deciduous shrubs mixed with the same tree species as in the ovestory, could not be differentiated from the canopy.

Significant openings in the canopy can result in direct ground illumination (confirmed in the field) that may be visible to the sensor. Most of the ground surface was covered with leaf litter, bare soil or woody debris, but pure pixels could only be found in the bare soil of agricultural fields, so the second end member extracted was termed 'soil'.

The third and final endmember used, was shadow. Due to the inherent structure of trees a large amount of shadow exists (Seed and King, 2003; Asner and Warner, 2003; Asner et al., 2003), which was of specific interest as gaps were thought to increase the amount of shadow and image heterogeneity.

Non-photosynthetic vegetation was not used as an endmember as it was almost impossible to train for by manually selecting pixels from the imagery (see next section) since pure pixels of woody debris or branches do not exist. Had a field spectroradiometer been used, then potentially non-photosynthetic vegetation could have been included in the analysis, which may have helped provide information on canopy heterogeneity by detecting trees with more visible branches, indicating more gaps and increased sunlight penetration to the forest floor.

9.2.5.1 Endmember Selection

As outlined in the background section, there are many different methods of selecting endmembers for use with unmixing algorithms. For this research, three different methods were compared: manually delineating pure areas of each endmember within the

imagery, selecting pure areas directly from n-dimensional spectral space (scatterplots), and automatic endmember selection, based on an iterative error analysis algorithm (Staenz et al., 1998). These methods were tested using the Ikonos imagery because it was easier to examine the spectral characteristics present in the higher resolution imagery and therefore develop a better understanding of the methods and decide which performed best.

Manually delineating pure endmember areas was done by zooming in at various locations and selecting pixels that were thought to only contain the spectra of a single image component. For green photosynthetic vegetation (GPV) this entailed selecting examples of very bright red pixels in a colour infrared composite thought to be homogeneous canopy, and that could most likely be categorized as sunlit canopy (Peddle and Johnson, 2000; Peddle et al., 2001). Pure pixels of soil were delineated within the large bare agricultural fields in the image. Lastly, as pure pixels of shadow would not exist in the data (Pellikka et al., 2000), the shadow endmember was created by selecting dark pixels, with little to no reflectance in any spectral bands that were located in water bodies (Holben and Shimabakuro, 1993; Braun and Herold, 2003; Twele and Barbosa, 2004). Figure 14 shows examples of selected pixels for each of the three endmembers. The average values for the endmembers in each spectral band were then used to feed the unmixing algorithm.

Selecting endmembers from the n-dimensional spectral space was done by viewing different combinations of spectral bands in 2D scatterplots. Figure 15 shows an example of a generalized scatterplot of the NIR band versus the red band for the whole Ikonos scene. This method is based on the idea that the purest pixels of land cover types

or scene components will be located at the extreme vertices or boundaries of the spectral space (Van der Meer and de Jong, 2000). Through interactive visualization of the three vertices, the three endmembers of interest were defined. Given that GPV was known to reflect highly in the NIR portion of the spectrum and low in the red portion of the spectrum, it was no surprise that the purest pixels of GPV were clustered at the lower-right vertex of the triangle. The second vertex, located at the lower left, where little NIR or red radiance was reflected, corresponded to pixels within deep non-reflecting water bodies, which were used to represent shadow. The third vertex corresponded to pure soil pixels found in bare fields with moderate reflectance of NIR and red wavelengths. All other land cover types and scene components were therefore considered to be mixtures of these three endmembers and were located in between the vertices of the spectral space triangle. Again, the average values for the endmembers in each spectral band were then used to feed the unmixing algorithm. The portion of spectral space selected for deriving each endmember was selected somewhat arbitrarily in order to provide a good sample size, but not to contain too many mixed pixels. By interactively visualizing the selected pixels in the actual image, it was determined when too much of the spectral space was selected, thereby including truly mixed pixels.

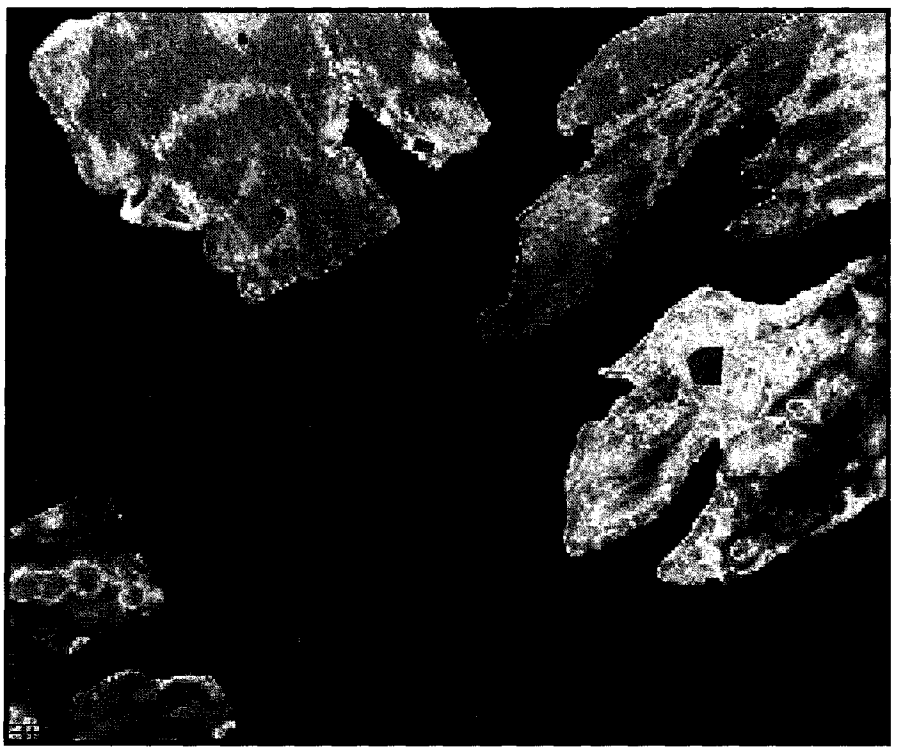


Figure 14. Manual endmember selection from Ikonos imagery. Training areas for soil (brown/orange), shadow (blue), and green photosynthetic vegetation (green) are shown.

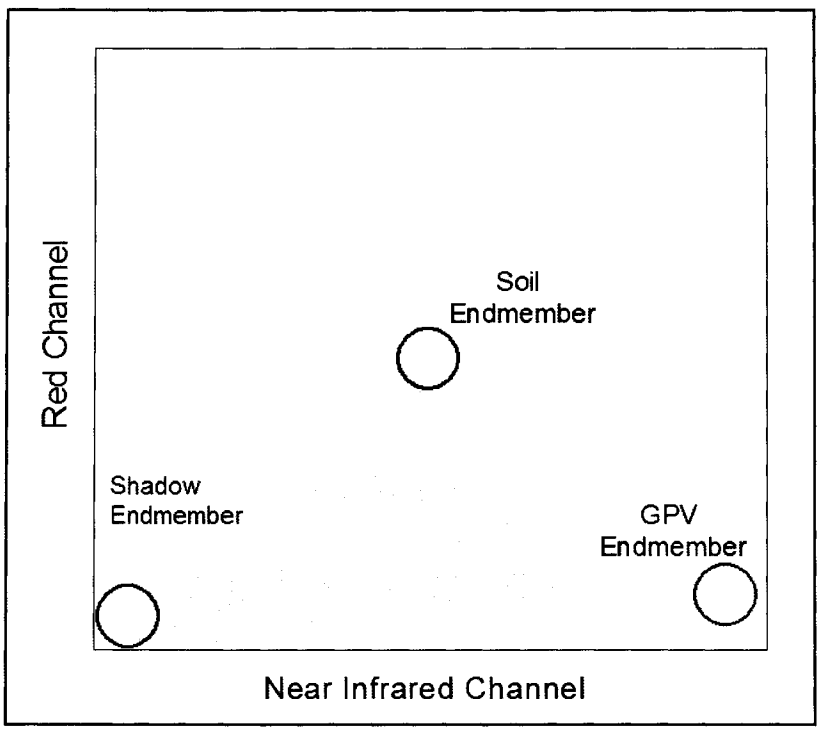


Figure 15. Endmember selection from n-dimensional spectral space for soil (brown), shadow (blue), and green photosynthetic vegetation (green) are seen.

Automatic endmember selection was carried out using the PCI Geomatica algorithm (PCI Geomatics, 2004), which is based on the Iterative Error Analysis (IEA) procedure, built into CCRS's Imaging Spectrometer Data Analysis System (ISDAS) (Staenz et al., 1998). The algorithm was run and 4 endmembers were specified, in order to see what endmembers could be identified from within the imagery. Once the endmembers were created and run through the spectral unmixing algorithm, those fraction images that showed seemingly useful information were named and used for further analysis. Only the first two endmembers could be identified following the unmixing. While it was not certain as to what these endmembers represented, they were originally loosely defined as being shadow (first automatically generated endmember) and green photosynthetic vegetation (second automatically generated endmember).

While endmembers were derived from the entire image, in order to ensure pure samples of each existed (i.e. water bodies and agricultural fields), only the forested areas underneath the forest mask were unmixed.

Sub-pixel fraction values, derived using the three endmember selection methods, were then extracted at each of the nest (n = 24) and non-nest (n = 21) sites, as well as at the five sites added with specifically low percent cover. Pearson correlations were calculated between the various sub-pixel fractions in order to examine the relations between the end member selection methods.

9.2.6 Comparison of Canopy Gaps in Hemispherical Photographs to Shadows in Ikonos Imagery

As an initial investigation into the ability of satellite imagery for characterizing canopy gaps, the high resolution Ikonos imagery was compared to the hemispherical photographs to see if the gaps in the hemispherical photographs were imaged as shadows in the remote sensing imagery.

Several of the hemispherical photographs were visually compared to the 1 m x 1 m pixel size panchromatic Ikonos image, which was the highest resolution imagery available. Sites with very large gaps, medium gaps, and almost no gaps, as seen in the hemispherical photographs were examined. A 20 m diameter circle surrounding the sites in the imagery was used, which corresponded to the approximate 2D view-area in the canopy of the hemispherical photographs (see Section 9.2.4).

9.2.7 Modelling Canopy Cover Using Ikonos Imagery

Potentially a map of forest cover could also have been used to identify nest areas. Given that the hemispherical photographs provided simple binary images of canopy and sky, it was hypothesized that either the sub-pixel fractions or the spectral and spatial information in the Ikonos imagery would be related to the percent cover measured by the hemispherical photographs. Pearson correlations and multiple linear regression analysis were used to test these relations. Percent cover within the 22.5° zenith angle was used since it was found to provide the most useful information in Part A of this thesis.

9.2.8 Comparison of Nest and Non-Nest Sites

The spectral, textural, and image fraction variables extracted from the Ikonos imagery were examined for differences between the 24 nest and 21 non-nest samples within St. Williams Forest. Due to lower canopy cover and greater numbers and amounts of shadows at nest sites, it was expected that: 1. the forest canopy associated with nest sites would show lower reflectance throughout the visible and near-infrared wavelengths, 2. texture would be greater at nest sites, reflecting the more heterogeneous canopy, and 3. sub-pixel fraction GPV would be lower while the shadow fraction would be higher. Comparisons were made using Mann-Whitney U tests, in order to keep the statistical tests consistent throughout the thesis.

The variables that did not show significant differences were dropped from further analysis. Those that were found to be significantly different ($p \leq 0.05$) were entered into a Principal Components Analysis in order to create a set of uncorrelated factors, which explained a large proportion of the variance, to be used for modelling of potential habitat. The factors created by these analyses were explored in terms of the percent variance each represented, as well as their factor loadings, and a final set of PCs was selected for habitat mapping. These components were used as input variables for the Maximum Likelihood classification, as well as the logistic regression modelling.

9.2.9 Maximum Likelihood Supervised Classification

A Maximum Likelihood supervised classification (MLC) algorithm was used as an initial method to create a map of potential nesting and non-nesting habitat. MLC classifies each pixel in the image into the class to which it most likely belongs based on

probability density functions computed for each class from the training data (Lillesand and Kiefer, 2000). In this case, two classes were used so each pixel in the masked forest image was classified as either nesting habitat or non-nesting habitat based on the training data used.

Training data were extracted from 5 x 5 pixel windows centred on each site in order to incorporate any local variability in the imagery due to positional or radiometric errors. As well, the final maps produced were filtered with a 5 x 5 pixel mode filter to remove noise in the classification, and to indicate the majority class within this area. For example, if a nest site was located within a pixel classified as non-habitat, but five or more of the eight surrounding pixels were classified as habitat, that central pixel would be changed to become a habitat pixel as well. Spatial autocorrelation in the data at this scale would be expected, justifying this processing step.

The *a posteriori* class probabilities, representing the probability of each pixel in each class, were saved and analyzed further (Bastin, 1997). Validation was conducted using 157 independent nest sites that were overlaid on the classified maps to see if the sites were classified as habitat or non-habitat. The non-nest transect used for training was unfortunately the only area conclusively identified by a Bird Studies Canada expert as being a known non-nesting area. Consequently, instead of quantitative error analysis, qualitative analysis of the patterns of the classification in relation to nest site locations was conducted in order to assess the potential for identifying nesting and non-nesting habitat.

9.2.10 Binary Logistic Regression

Following the methods of Mladenoff et al. (1995) for mapping wolf habitat, as well as O^2 Planning and Design (2003) for mapping whooping crane nesting habitat, binary logistic regression was tested for creating maps of potential hooded warbler nesting habitat. Binary logistic regression is used to examine the relation between a binary dependent variable (i.e. presence/absence, dead/alive, etc.) and a set of independent predictor variables (Trexler and Travis, 1993). In this case, the models created could then be used to predict the probability that a pixel in the image would contain suitable habitat for hooded warbler nests.

The image principal components previously created were used as independent variables with nest sites (presence) and non-nesting areas (absence) used as the dependent variable. Stepwise binary logistic regressions were run with a 95% significance level. Collinearity between independent variables was not an issue since the variables were principal eigenvectors and were therefore completely uncorrelated with one another. Models were evaluated based on the goodness-of-fit as determined by the Nagelkerke r^2, which, as in linear regression modelling, provides a measure of the amount of variance of the dependent variable that was explained by the model (SPSS, 2004). As well, the model classification accuracy was examined, which summarizes the correct and incorrect predictions of presence/absence for the cases used (the model input or training data).

9.3 Mapping with Landsat Imagery

This section outlines the methods used for mapping hooded warbler nesting habitat across their known range in southern Ontario. Landsat imagery was investigated

following the high resolution local analysis using Ikonos imagery, due to the fact that it provides extensive coverage at very low cost.

9.3.1 Study Area

In order to examine the full geographic extent of potential hooded warbler nesting habitat, it was necessary to examine forests across southern Ontario, from Toronto west to Lake Huron and south to Windsor (Figure 16). The area was covered by two Landsat scenes with the exception of a small portion south of Lake St. Clair.

9.3.2 Field Data Collection

With an extensive database of nest sites already compiled (as described in section 9.1), examples of non-nest sites across the region were also desired. They were selected based on the fact that they a) were in forests not specifically marked as private property, b) were in forest patches large enough to be distinctly visible in the satellite imagery, and c) contained no ground vegetation that could serve as potential nest habitat (Figure 17). Due to the lack of forest across much of the area, and because most of the remaining forests were located at the backs of agricultural fields and were therefore impossible to access without permission, only twenty-one forests were visited. They were evenly spread out across the region (Figure 16) and at least 2 sites were selected per forest.

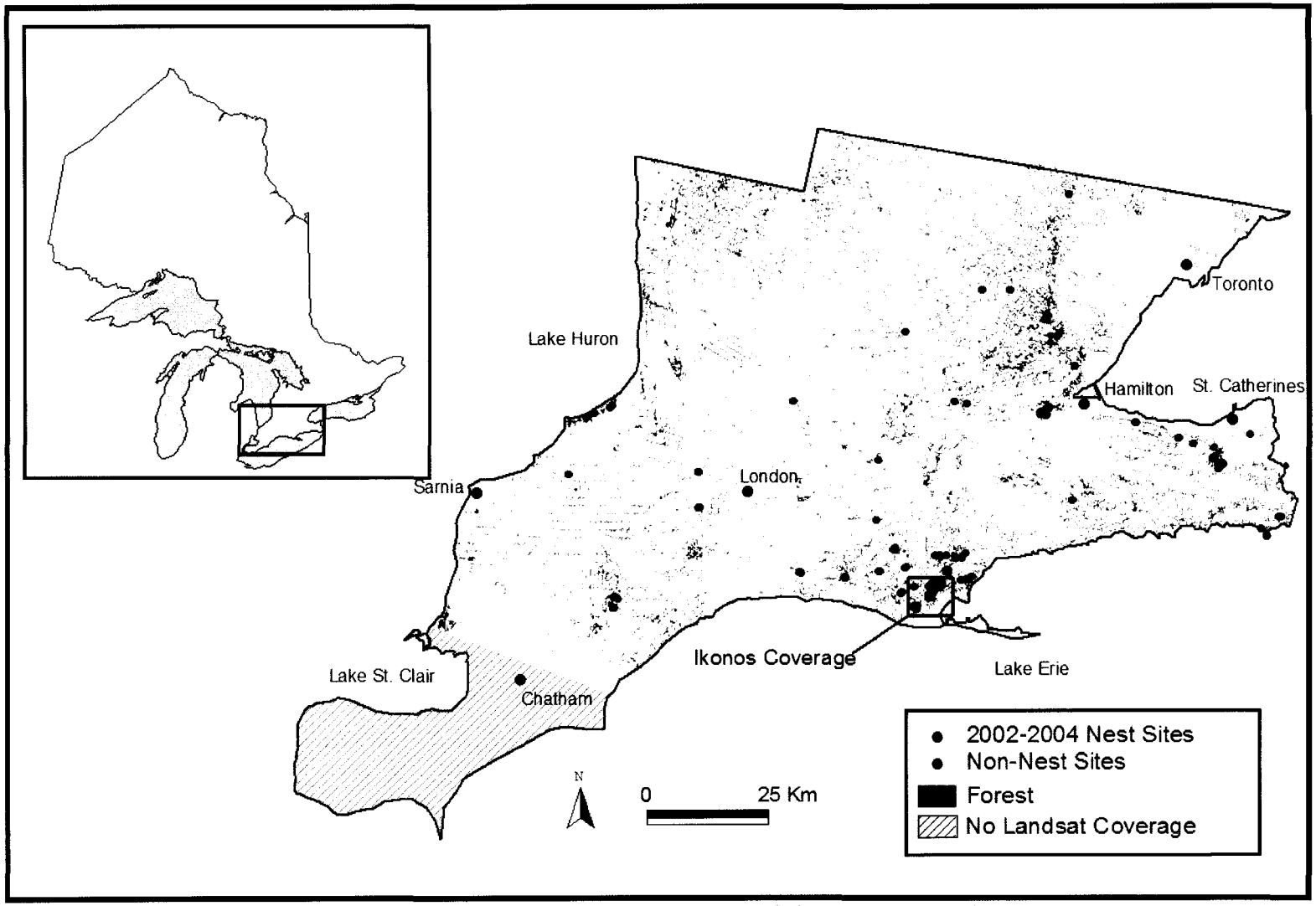


Figure 16. Approximate coverage of Landsat and Ikonos imagery shown along with a forest map, and the nest and non-nest sample locations.

Figure 17. An example of non-nest sites with lack of suitable ground vegetation required for hooded warbler nesting.

9.3.3 Image Acquisition and Initial Processing

Landsat 7 provides coverage of approximately 185 km x 185 km per scene with a spatial resolution of 28.5 m x 28.5 m. Unfortunately, recent technical problems with the sensor's scan line corrector (SLC) resulted in high quality coverage of only about 75% of each scene (NASA, 2004). Instead, Landsat 5, which was launched in 1984 but continues to provide high quality data, was used for this research. Landsat 5's Thematic Mapper (TM) sensor provided the necessary spatial resolution and multispectral bands at a much lower cost than the slightly higher resolution SPOT sensor. The Landsat 5 data that were used included six multispectral bands (Table 5), each with a re-sampled spatial resolution of 25 m x 25 m. Unfortunately, unlike the Landsat 7 Enhanced Thematic Mapper (ETM+), a higher resolution (10 m pixel) panchromatic band was not available.

Two Landsat scenes from different orbital paths (Path 18 Row 30 and Path 19 Row 30), with a scene overlap of approximately 65 km, were used to cover the majority of southern Ontario. For leaf-on conditions, the only cloud-free data consisted of a scene from June 4th 2004 (Path 18 Row 30) and a scene from September 15th 2004 (Path 19 Row 30). Both scenes were projected to the UTM Zone 17T NAD83 datum.

Table 5. Landsat 5 TM bands used in this project. (Chander et al., 2004)

Band	Wavelength (μm)
Blue	0.45 – 0.52
Green	0.52 – 0.60
Red	0.63 – 0.69
Near Infrared (NIR)	0.76 – 0.90
Mid Infrared 1 (MIR 1)	1.55 – 1.75
Mid Infrared 2 (MIR 2)	2.08 – 2.35

These acquisitions were almost completely cloud free. However, being more than three months apart caused significant reflectance differences between the two scenes. Many fields that were bare in early June had grown into mature crops by mid-September.

The two scenes were co-registered by collecting tie points in the overlapping region (all tie points and ground control points are shown in Appendix 1). Before a mosaic could be produced, a relative calibration procedure was performed in order to minimize reflectance differences in forested areas between the two scenes. A simple Manual No-Change Regression (MNCR) approach was used (Over et al., 2003), which compares a set of dark and bright unchanged areas in both images. The mean DN values in each spectral band were calculated for dark water bodies (10 areas) and large impervious areas (urban and industrial) and quarries (10 areas). The scene that

corresponded to the Ikonos area (Path 18 Row 30) was used as the master, while the other scene was considered the slave. A least squares regression was performed in order to derive the calibration equation for each band (regression results and equations are shown in Appendix 3), and the transformation was applied to the slave's bands. To check the quality of the calibration in forested areas, the mean DN values in each original and calibrated spectral band were extracted from twenty forest sites in the overlapping region. Before attempting to relatively calibrate the western Landsat scene to the eastern one, it was found that the blue, green, and red bands had absolute differences of 1.44%, 10.03%, and 2.28%, respectively for forests in the overlap area. The first iteration of the calibration procedure caused these differences to increase to 6.46%, 15.74%, and 11.24%, so these bands were left in their original state. The NIR and MIR bands showed significant improvements following two iterations of the calibration (Table 6), so the resulting data from the second calibration were used in further analysis. A third iteration was not performed as it was decided that further significant improvements were unlikely.

Table 6. Average differences of forest validation areas as a percent of the average DN value of the same areas from the master scene. Bolded values show the minimum difference obtained, and therefore the level of calibration that was applied to each band

	Original	After 1st Calibration	After 2nd Calibration
Blue	**-1.44**	6.46	-
Green	**10.03**	15.74	-
Red	**-2.28**	11.24	-
NIR	38.86	7.09	**6.93**
MIR 1	33.92	5.61	**0.56**
MIR 2	37.67	7.57	**3.39**

A mosaic was produced from the two scenes using the previously collected tie points, and a first order nearest neighbour transformation in order to correct any geometric distortions between the images. A Root Mean Squared (RMS) error of (0.61, 0.51) pixels was achieved.

The Landsat mosaic was then georeferenced using a nearest neighbour second order transformation with the thirty-eight GCPs collected at road intersections spread throughout the imagery. This resulted in an RMS error of (0.61, 0.62) pixels. These GCPs were collected with same sub-meter Trimble GPS previously mentioned. Several of the same points that were used to georeference the Ikonos imagery were also used for this processing in order to ensure accurate alignment between the two data types. The United States was manually masked out of the imagery since the project was only concerned with forests in Canada (Figure 18).

As the study was focused on forested land, and there were still large differences in brightness between the scenes for other land cover types, forested land was extracted from the imagery using a supervised Maximum Likelihood classification. Originally a K-Means algorithm was tested, as was used for the Ikonos imagery. However, it did not provide the desired results. Training data for forests were therefore selected from across the entire image in order to ensure that the full variability of the forests was captured. A 3 x 3 Mode filter was then applied to smooth out the salt and pepper appearance of the map caused by unclassified forest pixels that were spectrally mixed (Yang and Lo, 2002). It was also necessary to visually scan the map and manually digitize some areas that had been missed by the classification. This forest map was obviously not perfect, nor was it

quantitatively validated, but it suited the needs of this research in terms of masking out non-forested areas in the image. Figure 19 shows an example of the forest classification.

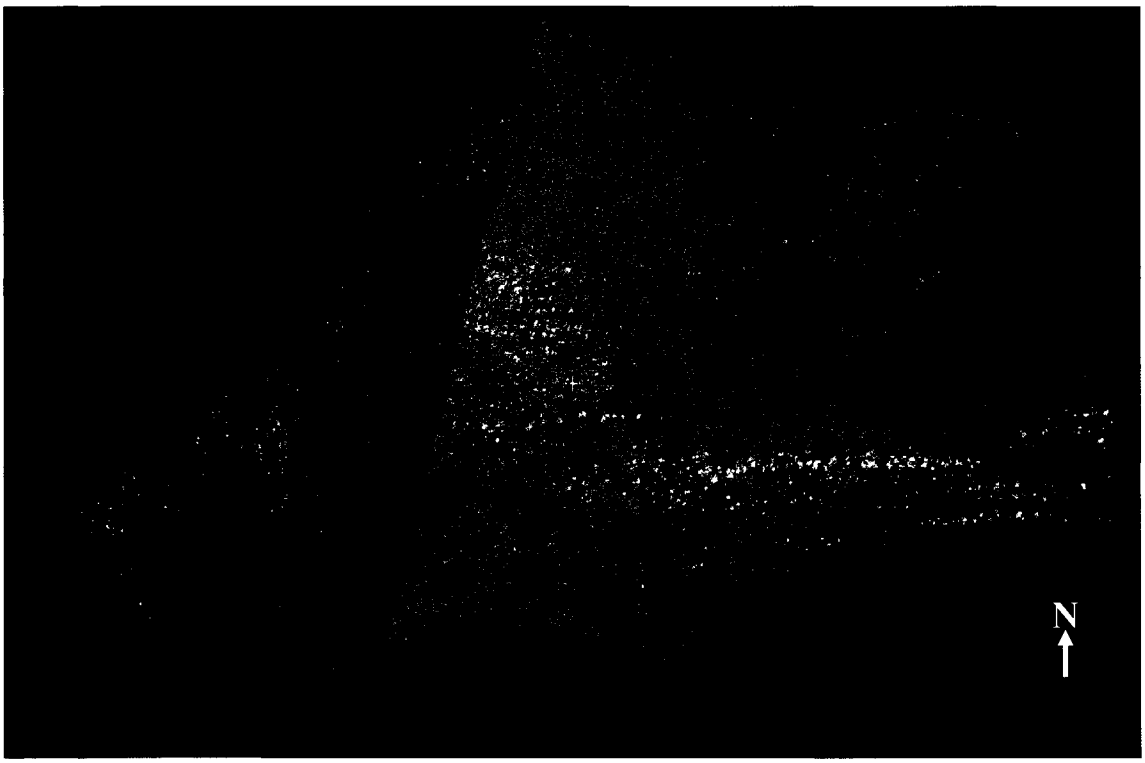


Figure 18. Final Landsat mosaic, shown as a colour infrared composite. Temporal differences are clearly visible in agricultural areas, with bare fields in the eastern half of the image and cropped fields in the western half.

The final forest map was then filtered to remove small forest patches that would never be used by the hooded warbler and therefore were simply cluttering the map. Flaxman's (2004) habitat mapping research showed that 96% of hooded warbler nests were in patches larger than 100 ha, and in fact two thirds of the nests were in patches larger than 1000 ha. However, this was based on forest areas measured from a land cover map produced from Landsat imagery. Such a map is subject to errors of omission and

commission that could cause a larger forest to be split in two or two smaller patches to be joined into one large forest (as shown in Figure 19). In addition, few forests in southwestern Ontario are of such large size. Consequently, a smaller threshold of 15 ha was selected, which Flaxman stated was the minimum patch size used by hooded warblers in Maryland, New York, and Ohio. Potentially, the interpatch distance could have been factored into this thresholding, however, this would have required a better understanding of hooded warbler movement between forest patches.

Using this map, a final analysis of the quality of the relative calibration procedure was tested. The forests in the eastern half of the mosaic (June imagery – master) were compared to the forests in the western half of the mosaic (September imagery – slave). Table 7 shows the mean and standard deviation DN values for each of the six spectral bands. The green, red, NIR, and MIR 2 bands were retained for further analysis.

Table 7. Mean and standard deviation of DN values for each Landsat spectral band used to determine which bands were well calibrated for forested areas.

	East Forests (June - Master)		West Forests (September - Slave)	
	Avg.	**Stdev.**	**Avg.**	**Stdev.**
Blue	60.99	2.81	64.74	2.47
Green	27.28	1.75	25.70	1.34
Red	20.04	2.25	20.92	1.41
NIR	117.91	16.51	117.46	14.05
MIR 1	74.98	8.84	81.43	8.41
MIR 2	22.80	3.36	24.14	3.60

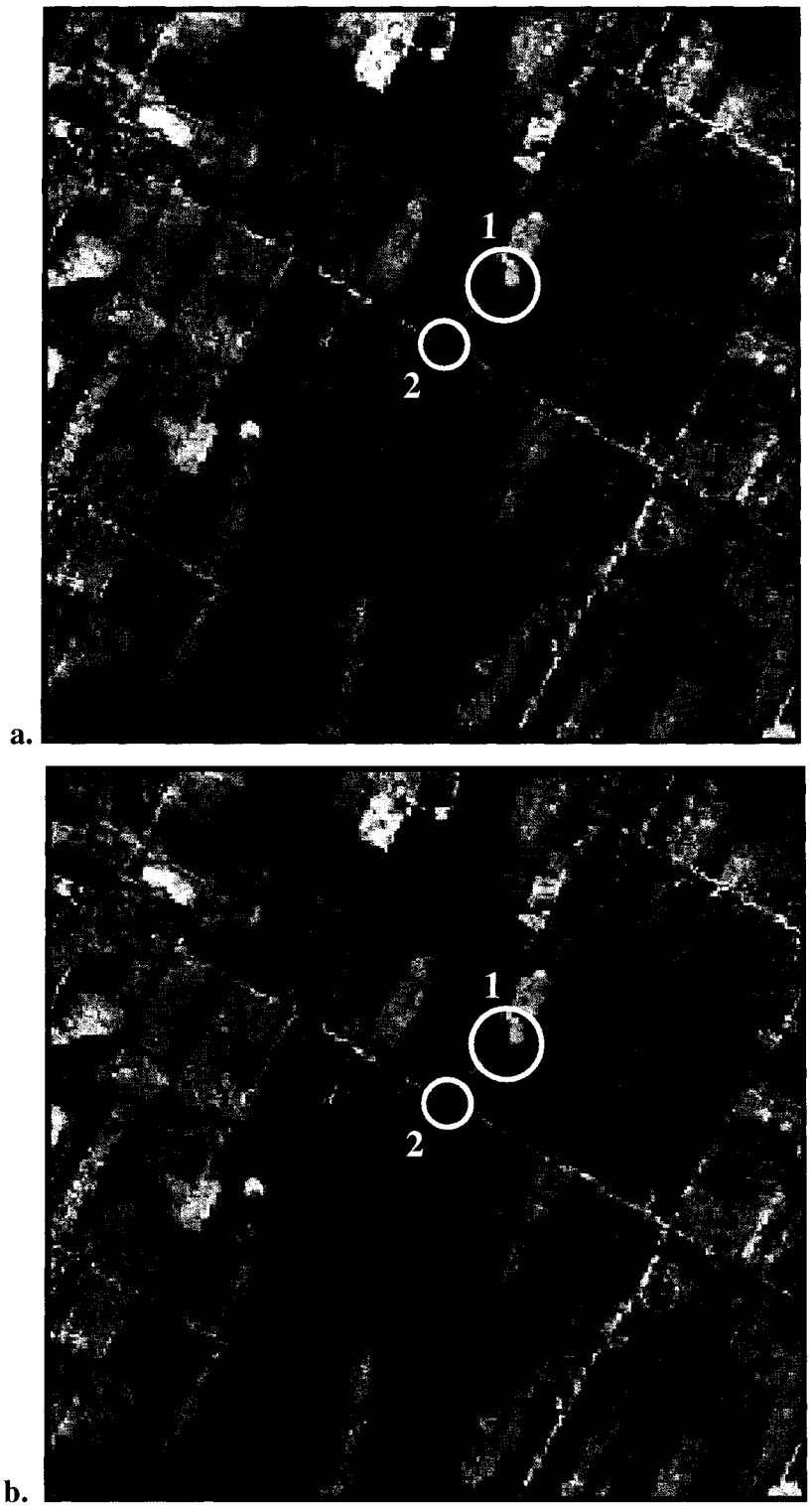


Figure 19. Examples of errors of omission (1) and commission (2) in forest classification, shown in a CIR composite (a) and the classified forest map (b).

9.3.4 Extraction of Spectral and Textural Information

The six GLCM textures previously discussed were calculated for the Landsat forest mask image using a 3 x 3 window to calculate local canopy texture. The spectral and textural information was extracted from the Landsat imagery at each study site using a 3 x 3 pixel window (75 m x 75 m area) average, which was larger than any of the expected GPS or image registration errors.

9.3.5 Linear Spectral Unmixing

Using the knowledge derived from the endmember selection and unmixing of the Ikonos imagery, endmembers were manually selected from the Landsat imagery, and the six bands were unmixed under the forest mask for the three endmembers to create fraction images of GPV, soil, and shadow. Sub-pixel values were averaged using a 3 x 3 window, and extracted for each study site.

9.3.6 Comparison of Nest and Non-Nest Sites

The spectral, textural, and image fraction variables extracted from the Landsat imagery were examined for differences between the 279 sample nests and 63 samples of non-nest sites across southern Ontario.

Principal components were created using all the variables that showed a significant difference ($p \leq 0.05$). Since this threshold cut-off a large number of potentially important variables, a second set of principal components was created including those variables that showed a difference using $p \leq 0.10$. The factors created by these analyses

were explored in terms of the percent variance each represented, as well as their factor loadings, and in each case a final set of PCs was selected for habitat mapping.

9.3.7 Maximum Likelihood Classification and Binary Logistic Regression

The supervised classification and logistic regression methods previously described for the Ikonos imagery were repeated using Landsat. Maximum Likelihood supervised classification was conducted as in Section 9.2.9. However, training data were extracted from 3 x 3 pixel windows centred on each site, and the final maps produced were filtered with a 3 x 3 mode filter. The algorithm was applied to each set of PCs generated above. Of the 279 confirmed breeding sites (nest sites), 140 of them were randomly selected for training, leaving 139 for validation. Of the 63 non-nest sites, 38 were randomly selected for training, leaving 25 sites for validation. The locations of these sites are shown in Figure 20. Error matrix analysis was conducted separately for the training and validation data to assess the quality of the two habitat maps.

As well, binary logistic regression was carried out as in Section 9.2.10 using the two sets of PCs of the Landsat imagery as independent variables.

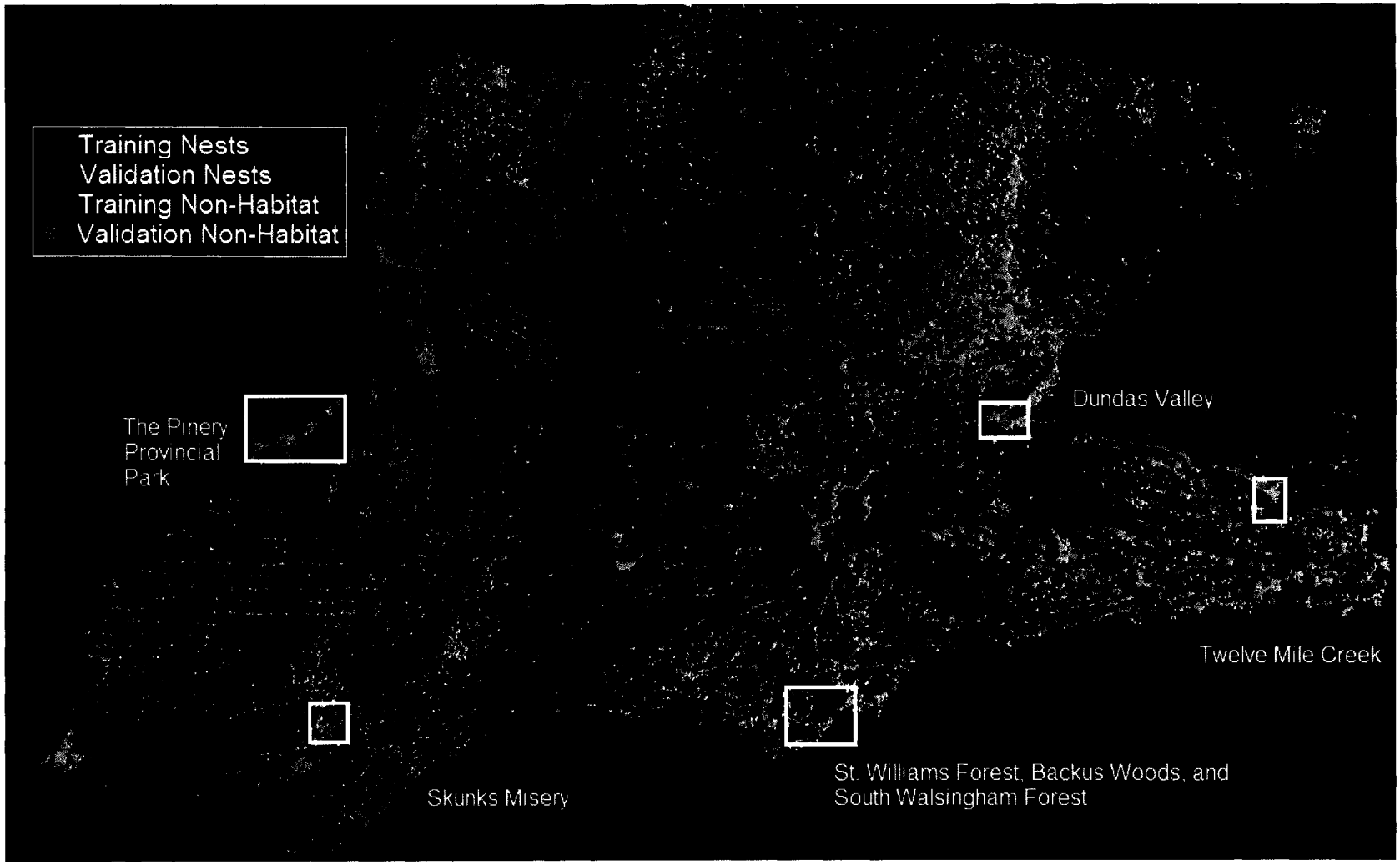


Figure 20. Forested areas (green) with locations of nest and non-nest training and validation sites used for Maximum Likelihood supervised classification habitat/non-habitat. Outlined areas show significant forests across Ontario that were looked at in more detail.

10.0 RESULTS

10.1 Mapping with Ikonos Imagery

10.1.1 Endmember Selection and Spectral Unmixing

Endmembers were created using the three methods discussed in Section 9.2.5.1, each represented by a DN value for each of the spectral bands. In the case of the Ikonos imagery, the values ranged from 0 to 2048 because the imagery was acquired as 11-bit (i.e. 2^{11} possible DN values).

The initial testing of endmember selection methods resulted in three known endmembers, representing green photosynthetic vegetation (GPV), soil, and shadow, for the first two methods tested. Manually delineating pure areas directly in the imagery was easily done, while selecting vertex pixels in the n-dimensional scatterplot was more challenging. However, both methods resulted in very similar endmembers (Table 8). The automated method produced four endmembers, however only the first two could be identified as shadow and green photosynthetic vegetation.

Table 8. Ikonos endmembers derived from three endmember selection methods. Cell entries are the mean brightness of each endmember in each spectral band.

EM Selection Method	Endmember	Blue	Green	Red	NIR
Image Delineation	**GPV**	278.14	310.39	195.12	1306.68
	Soil	462.35	620.16	622.71	700.56
	Shadow	262.77	234.71	114.17	141.70
Scatterplot	**GPV**	282.70	329.83	209.17	1617.24
	Soil	521.94	743.94	766.72	853.67
	Shadow	261.31	233.00	115.62	63.15
Automatic Generation	**Shadow**	85.20	122.20	114.80	155.70
	GPV	281.60	317.60	203.20	1669.90

All three selection methods resulted in very similar unmixed fractions. Sub-pixel fractions extracted at fifty sites were significantly correlated (Table 9) between and within the three methods. The correlations within a given selection method were moderate, but still artificially inflated because they were constrained to sum to one. High correlations of each endmember between the manual delineation and scatterplot selection methods showed that they represented essentially the same information.

Table 9. Pearson correlation matrix for sub-pixel fractions calculated using three endmember selection methods (n = 50). Significance values shown in brackets.

		Image Delineation			**Scatterplot**			**Automatic Generation**	
		GPV	**Soil**	**Shadow**	**GPV**	**Soil**	**Shadow**	**Shadow**	**GPV**
Image Delineation	**GPV**	-	x	x	x	x	x	x	x
	Soil	0.295 (0.038)	-	x	x	x	x	x	x
	Shadow	-0.255 (0.074)	-0.408 (0.003)	-	x	x	x	x	x
Scatterplot	**GPV**	0.977 (0.000)	0.270 (0.058)	-0.203 (0.157)	-	x	x	x	x
	Soil	0.405 (0.003)	0.974 (0.000)	-0.415 (0.003)	0.403 (0.004)	-	x	x	x
	Shadow	-0.405 (0.004)	-0.380 (0.006)	0.905 (0.000)	-0.279 (0.050)	-0.377 (0.007)	-	x	x
Automatic Generation	**Shadow**	-0.941 (0.000)	-0.412 (0.003)	0.563 (0.000)	-0.905 (0.000)	-0.514 (0.000)	0.660 (0.000)	-	x
	GPV	0.944 (0.000)	0.369 (0.008)	-0.557 (0.000)	0.906 (0.000)	0.473 (0.000)	-0.659 (0.000)	-0.999 (0.000)	-

The second of the two automatically generated endmembers ended up being labelled as GPV due to its high correlation with the GPV endmembers created by the other two methods. The first endmember appeared to be all other spectral information, seen from the extremely high negative correlation with the GPV endmember.

Given that manually selecting endmembers was by far the most straightforward and easiest method, and given that it showed the same results as using n-dimensional scatterplots, it was selected for spectral unmixing of the Ikonos imagery.

The fraction images derived from the Ikonos imagery (Figure 21) appeared visually correct, although there was no way to validate the fraction values calculated for each pixel. Figure 21 (a), shows a small area of St. Williams Forest with a neighbouring pine plantation and some agricultural fields. The corresponding GPV fraction image (Figure 21, b) shows, at point A in the pine plantation, the small sub-pixel fractions of GPV, except along the paths through the plantation, which were covered by dense ground vegetation. As previously mentioned, the conifers themselves were mixtures of GPV and shadow, whereas homogeneous deciduous areas (e.g. point B) with less shadow had higher GPV fractions. This deciduous area had close to 0% soil fraction (Figure 21, c), while some soil was detected in the pine plantation, due to the large linear gaps in the canopy caused by the rows of trees. The soil fraction image shows the sand/dirt paths in St. Williams (e.g. point C). These areas also had high shadow fractions as a result of shading from the trees alongside the paths. The shadow fraction image (Figure 21, d) showed very high sub-pixel fractions within the pine plantation, and as well through the more mixed forested areas.

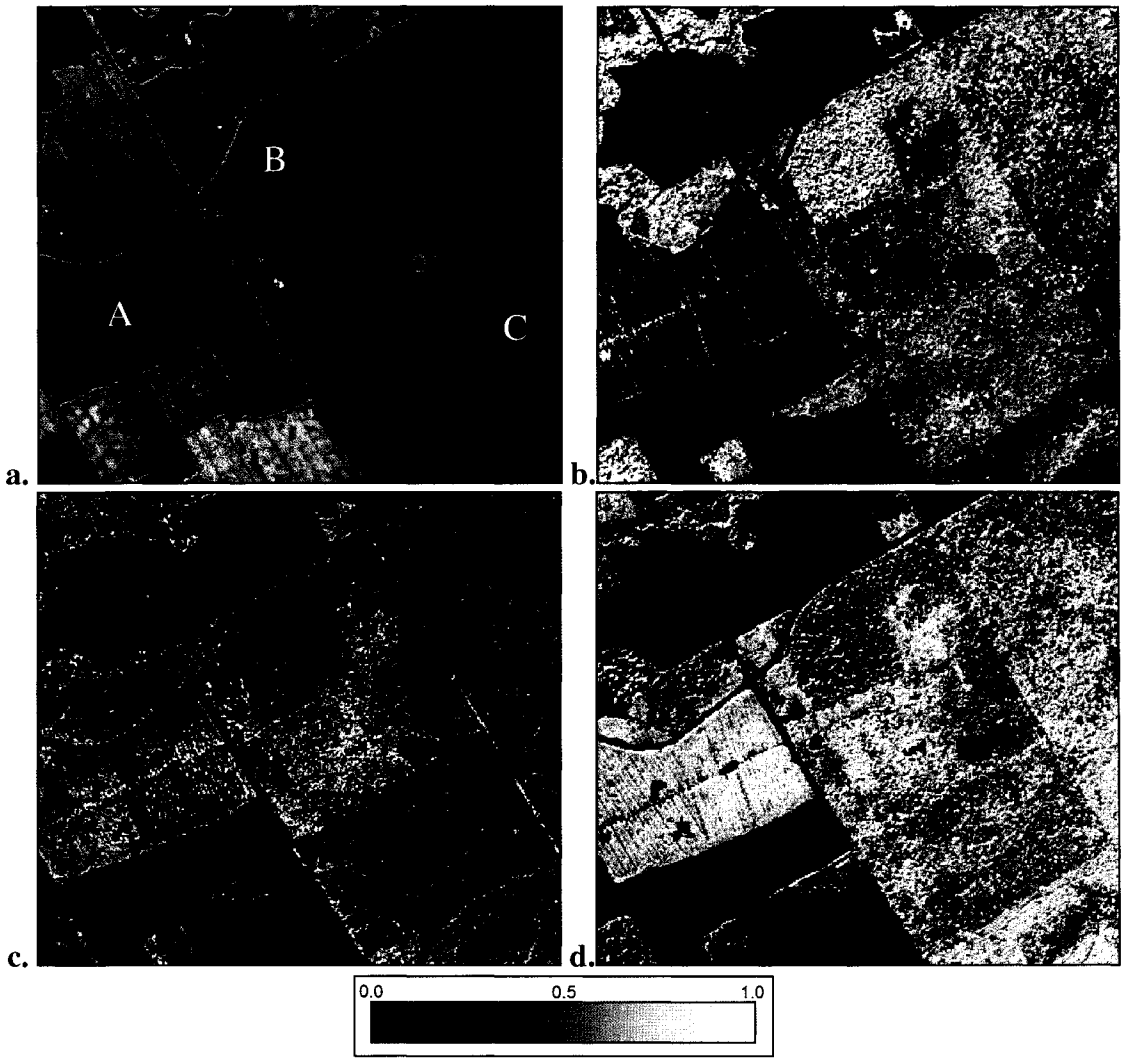


Figure 21. Ikonos imagery (a) fraction images for GPV (b), soil (c), and shadow (d), created through constrained least squares linear spectral unmixing using endmembers selected manually from the imagery.

10.1.2 Comparison of Canopy Gaps in Hemispherical Photographs to Shadows in Ikonos Imagery

Comparing the panchromatic Ikonos imagery to the hemispherical photos revealed that the pattern of gaps in the canopy was only partially similar to the corresponding shading in the imagery (e.g. Figure 22a and b), and in many cases completely dissimilar. For example, Figure 22c shows a very large canopy gap that

resulted in very little shading in the Ikonos imagery, and potentially the understory vegetation was directly illuminated. Figure 22d showed the opposite problem, with significant shading in the lower left quadrant of the satellite imagery but almost no gaps visible in the hemispherical photograph. This is an example of mutual shading in the overstory where one tree is smaller than the adjacent tree in an irregular canopy. Overall, the correlation between gaps in hemispherical photographs and the shadows in high resolution panchromatic satellite imagery was poor.

10.1.3 Modelling Canopy Cover Using Ikonos Imagery

The shadow and GPV fractions were not highly correlated to percent cover calculated from the hemispherical photographs (using the 22.5° zenith angle). Percent cover and the shadow fraction showed a positive correlation ($r = 0.283$, $p = 0.046$), while cover and GPV fraction were marginally correlated ($r = -0.234$, $p = 0.102$).

Finally, modelling hemispherical photo forest cover (22.5°) against the various spectral, textural, and sub-pixel fraction variables extracted from the Ikonos imagery resulted in a reasonably strong model. The multiple linear regression included red and, and GPV fraction ($R^2 = 0.45$; SE = 14%, Cover range = 14% - 98%), based on the model shown in Equation 2. This simple model included two independent variables, which varied as expected, with red reflectance decreasing and GPV fraction increasing, as the percent cover increased.

$$\%Cover = 3.16 - 0.018(\mathrm{Re}\,d) + 0.657(GPV) \qquad [2]$$

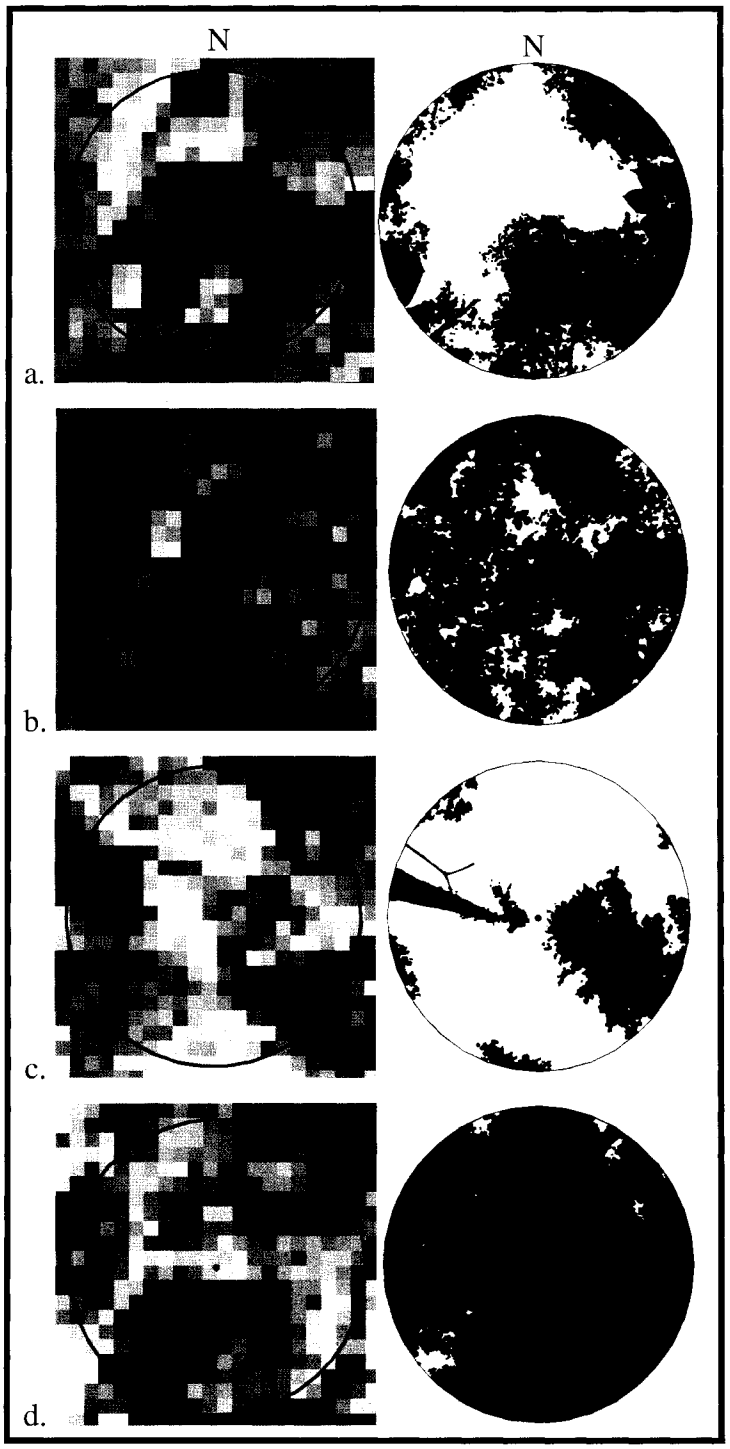


Figure 22. Examples of hemispherical photographs and the corresponding Ikonos 1 m x 1 m panchromatic imagery. In the Ikonos images, bright areas are vegetation and dark areas are shadows, while in the hemispherical photographs the white areas are gaps and the dark areas are canopy.

10.1.4 Comparison of Nest and Non-Nest Sites

The blue, green, and red spectral bands were found to be significantly brighter for nest sites than non-nest sites (Table 10), which was contrary to expectations. Textures (Table 11) were found to be greater at nest sites, as expected since nest sites were thought to have more heterogeneous canopies. Five of the six green and red textures and one of the NIR textures were significantly different. Of the sub-pixel fractions (Table 12), the soil fraction showed a significant difference between nest and non-nest sites. However, with an average proportion of 1% at nest sites and 0% at non-nest sites, this significant difference may simply have been a result of unmixing error. As well, the shadow fraction showed a greater proportion of shadow at non-nest sites compared to nest sites. With the high resolution imagery, it was thought that potentially a more local analysis using single pixel or a 3 x 3 window for extracting information may have been more appropriate than the 5 x 5 pixel window that was used. However, tests were done extracting single pixel spectral values, as well as 3 x 3 windows for spectral and textural information, and in all cases it was found that the significance of the differences between nest and non-nest sites increased when the 5 x 5 window was used.

Table 10. Mann-Whitney U tests comparing nest (n = 24) and non-nest (n = 21) sites in terms of spectral variables extracted from Ikonos imagery using a 5 x 5 window.

	Nest Sites		Non-Nest Sites			
	Avg.	Stdev.	Avg.	Stdev.	U	P
Blue 5x5 Average	264.79	11.77	253.81	4.81	78.00	**0.000**
Green 5x5 Average	264.83	4.28	259.28	2.19	52.00	**0.000**
Red 5x5 Average	160.79	10.61	151.57	3.48	104.50	**0.001**
NIR 5x5 Average	800.62	152.74	776.09	44.76	213.50	0.381
Pan. 5x5 Average	411.08	85.02	417.00	46.19	239.50	0.776

Table 11. Mann-Whitney U tests comparing nest (n = 24) and non-nest (n = 21) sites in terms oftextural variables extracted from Ikonos imagery using a 5 x 5 window.

	Nest Sites		Non-Nest Sites		U	p
	Avg.	Stdev.	Avg.	Stdev.		
Green Contrast	14.72	9.98	5.93	4.51	102.50	**0.001**
Green Dissimilarity	2.88	1.11	1.79	0.71	101.00	**0.001**
Green Entropy	3.01	0.23	2.67	0.31	89.50	**0.000**
Green Homogeneity	0.31	0.11	0.43	0.12	114.00	**0.002**
Green ASM	0.06	0.02	0.08	0.03	101.00	**0.001**
Green Correlation	-0.09	0.25	-0.12	0.25	233.50	0.674
Red Contrast	13.46	8.50	4.94	4.24	97.50	**0.000**
Red Dissimilarity	2.75	1.07	1.61	0.75	102.00	**0.001**
Red Entropy	2.91	0.32	2.53	0.42	119.00	**0.002**
Red Homogeneity	0.32	0.14	0.47	0.15	108.00	**0.001**
Red ASM	0.06	0.02	0.10	0.05	120.00	**0.003**
Red Correlation	-0.03	0.25	-0.05	0.32	224.00	0.524
NIR Contrast	89.42	54.85	91.15	45.08	231.00	0.633
NIR Dissimilarity	6.99	2.46	7.64	2.23	207.00	0.306
NIR Entropy	3.29	0.14	3.32	0.12	223.00	0.505
NIR Homogeneity	0.17	0.06	0.12	0.07	141.00	**0.012**
NIR ASM	0.04	0.01	0.04	0.01	216.50	0.414
NIR Correlation	-0.02	0.28	-0.14	0.29	195.00	0.195
Pan. Contrast	67.38	53.61	63.57	47.00	251.50	0.991
Pan. Dissimilarity	5.69	2.94	5.93	2.45	237.00	0.733
Pan. Entropy	3.08	0.45	3.25	0.22	201.00	0.246
Pan. Homogeneity	0.24	0.16	0.19	0.11	217.00	0.426
Pan. ASM	0.06	0.06	0.04	0.02	204.00	0.273
Pan. Correlation	0.13	0.22	0.11	0.22	166.00	0.275

Table 12. Mann-Whitney U tests comparing nest (n = 24) and non-nest (n = 21) sites in terms of sub-pixel fractions extracted from Ikonos imagery.

	Nest Sites		Non-Nest Sites		U	p
	Avg.	Stdev.	Avg.	Stdev.		
GPV 3x3 Average	0.56	0.13	0.54	0.04	213.5	0.380
Soil 3x3 Average	0.01	0.01	0.00	0.00	168.0	**0.023**
Shadow 3x3 Average	0.29	0.17	0.41	0.13	124.5	**0.004**

Principal components were created from the 3 visible bands, 11 texture bands, and 2 image fractions that showed significant differences between nest and non-nest sites. The first four were deemed to account for a significant percent variance (Table 13). The factor loadings (Table 14) for these components showed PC1 to be positively correlated

with the visible bands and negatively correlated with shadow. PC2 had strong negative correlations with green and red texture. PC3 and PC4 were not significantly correlated with the input variables and were therefore dropped from further analysis.

Table 13. Principal components created from the Ikonos imagery that showed significant differences between nest and non-nest sites

Factor	% Variance
PC1	**67.61**
PC2	**28.32**
PC3	2.17
PC4	1.04
PC5	0.82
PC6	0.04

Table 14. Factor loadings for Ikonos principal components that showed significant differences ($p \leq 0.05$) between nest and non-nest sites.

	PC1	PC2	PC3	PC4
Blue	**0.90**	0.26	0.15	0.05
Green	**0.95**	0.27	-0.18	-0.05
Red	**0.96**	0.21	0.16	0.04
Green Contrast	0.47	**-0.84**	-0.10	0.16
Green Entropy	0.30	-0.35	-0.13	0.08
Green Homogeneity	-0.25	0.53	0.15	-0.14
Green ASM	-0.24	0.28	0.10	-0.07
Green Dissimilarity	0.41	**-0.81**	-0.17	0.18
Red Contrast	0.48	**-0.83**	0.08	-0.16
Red Entropy	0.26	-0.38	-0.01	-0.08
Red Homogeneity	-0.22	0.55	0.02	0.11
Red ASM	-0.21	0.30	0.01	0.06
Red Dissimilarity	0.40	**-0.82**	0.01	-0.16
NIR Homogeneity	0.08	0.39	-0.09	-0.07
Soil	0.55	0.19	-0.17	-0.09
Shadow	**-0.62**	-0.24	-0.18	-0.06

10.1.5 Maximum Likelihood Classification

The two principal components were used as input variables for the supervised classification. Training data for nest and non-nest sites in this area had a Transformed Divergence separability of 1.64, which was considered moderate (TD ranges from 0 to 2.0, with a value of 2.0 representing complete statistical separation of the training data distributions (Jensen, 2005)). The output map was filtered with a 5 x 5 window in order to reduce noise. Maximum likelihood classification produced a map composed of 92.2% nesting habitat and only 7.8% non-nesting habitat. Figure 23 shows this habitat/non-habitat map, the training sites, and the nest sites used for validation.

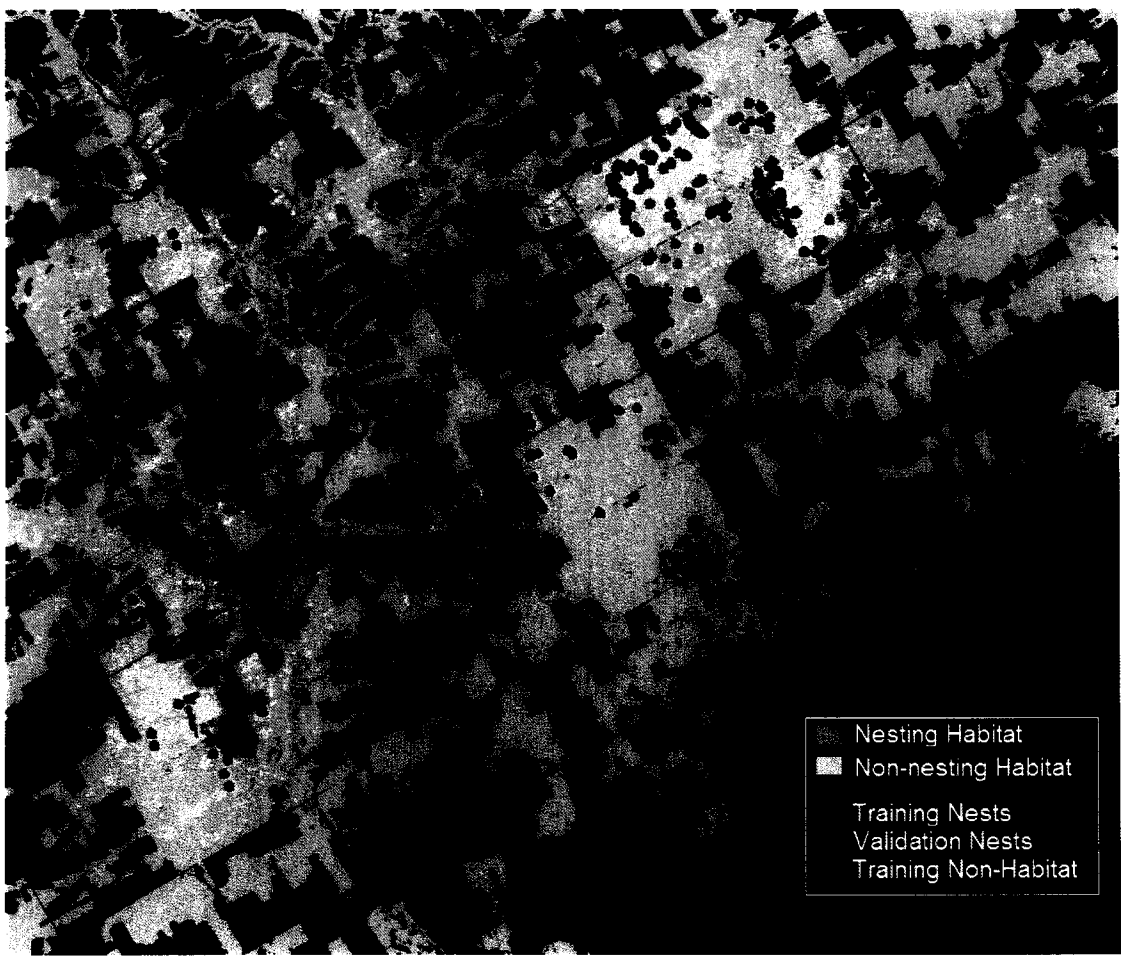


Figure 23. Supervised classification of Ikonos imagery, showing hooded warbler nesting and non-nesting habitat, along with training and validation sites.

Of the 157 validation nest sites in the area, 109 (70%) of them were located in pixels classified as nesting habitat. Figure 24 shows a close-up view of St. Williams Forest, where most of the sites were located. Several large areas where nests had not been found in the past few years were well classified as non-nesting habitat, including the transect used for training. The rectangle outlined in Figure 24 delineates an area where there was a high degree of confusion with many nest sites classified as non-nesting habitat. Further analysis showed that in most cases the nest sites were right at the edge of habitat/non-habitat patches, as shown in Figure 25 (a portion of the rectangular area in greater detail). To further investigate this, a map was produced measuring the distance to the nearest habitat patch, and the distance was extracted for each of the 157 validation nest sites (Table 15). The distribution of distances showed that of the 48 nest sites classified as non nesting habitat, 28 (18%) of them were within 10 m (2.5 pixels) of nesting habitat.

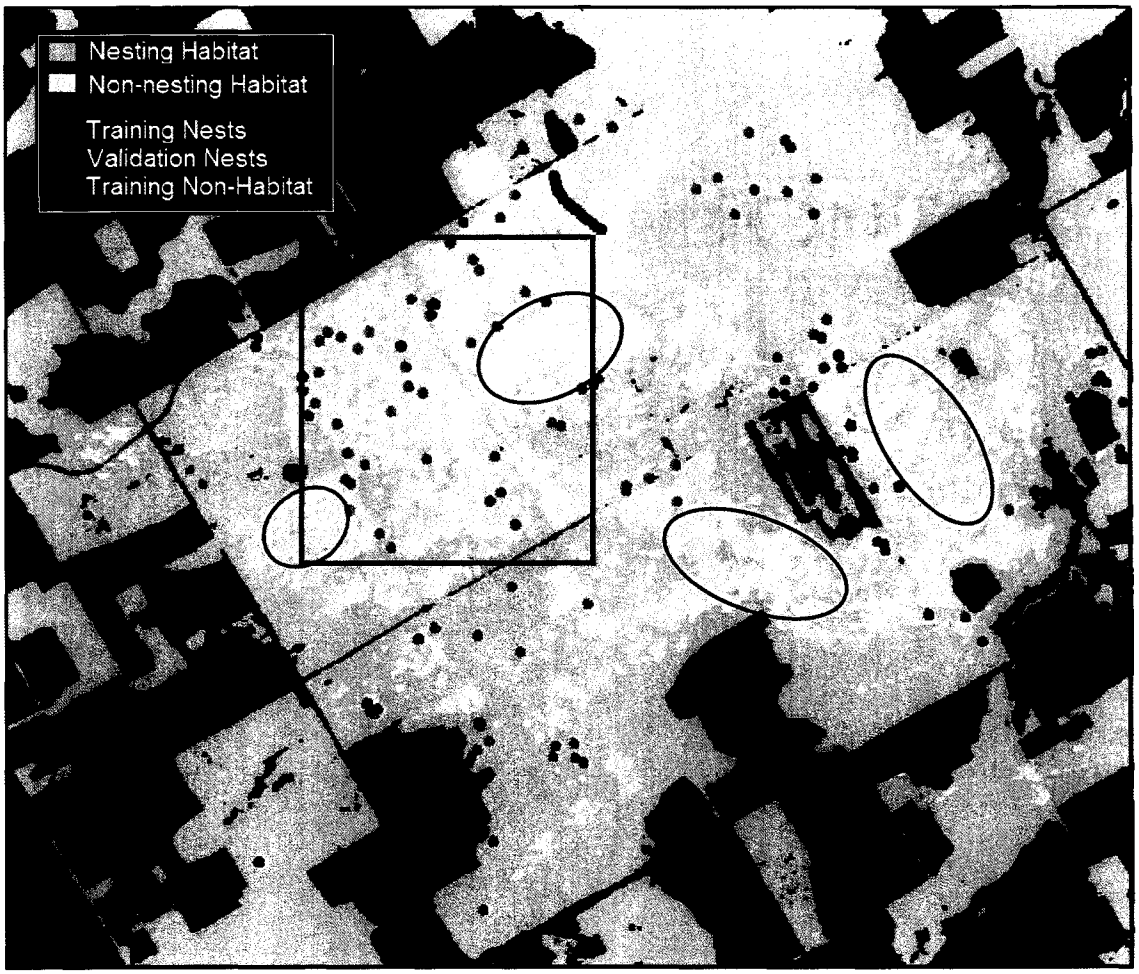


Figure 24. Close-up of St. Williams Forest showing areas classified as habitat and non-habitat. Areas outlined in red are classified as non-nesting habitat and that were previously unused for nesting. The black box indicates the location of the subset area shown in Figure 26.

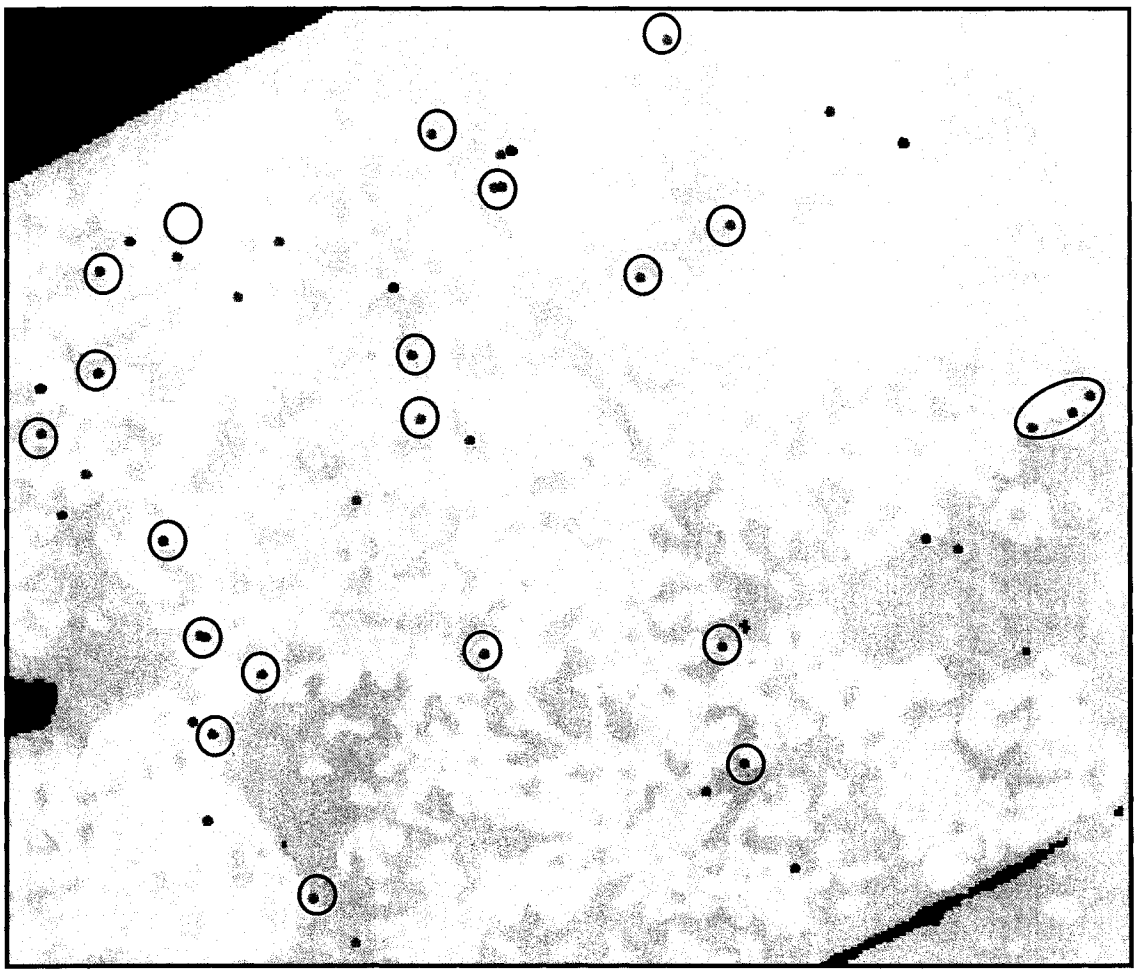


Figure 25. Subset area within St. Williams Forest, showing examples (circled) of validation nest sites classified as non nesting habitat that were located at the boundary of habitat/non-habitat.

Table 15. Distance to habitat measured for 157 validation nest sites in St. Williams Forest

Distance to Classified Habitat	Number of Nests
0 m	109
5 m	11
10 m	17
15 m	5
20 m	6
25 m	5
30 m	3
35 m	0
40 m	1

The distribution of *a posteriori* probabilities from the maximum likelihood classification algorithm (Table 16) showed 62.5% of the nest sites being in pixels classified as habitat with high probability (> 0.70 (O^2 Planning and Design, 2003)). Interestingly, there were not very many nests in areas of moderate habitat probability but 23.6% of the nests were in pixels classified as very low probability of being habitat (probability < 0.20).

Table 16. Distribution of maximum likelihood *a posteriori* probabilities for validation nest sites.

Probability	Number of Nests	% of Nests
0-10%	24	15.3
10-20%	13	8.3
20-30%	3	1.9
30-40%	5	3.2
40-50%	5	3.2
50-60%	5	3.2
60-70%	4	2.6
70-80%	4	2.6
80-90%	6	3.8
90-100%	88	56.1

10.1.6 Binary Logistic Regression

Logistic regression using the two principal components as independent variables produced a model with the first principal component as a predictor for nest presence. Equation 3 shows the model for deriving the probability of a pixel being hooded warbler nesting habitat. This model had a Nagelkerke r^2 of 0.495, and correctly predicted 67% of the training nest sites and 76% of the training non-nest sites. The prediction curve (Figure 26) created using this equation, showed that pixels with values of the principal component less than approximately 70 were predicted to have zero probability of being nesting habitat. From this value to approximately 125, the probability of occurrence of

nesting habitat increased following an s-curve, with values greater than 125 showing 100% probability of being appropriate nesting habitat.

$$\text{Probability (nest)} = \frac{e^{-14.267+0.152[PC1]}}{1+e^{-14.267+0.152[PC1]}} \qquad [3]$$

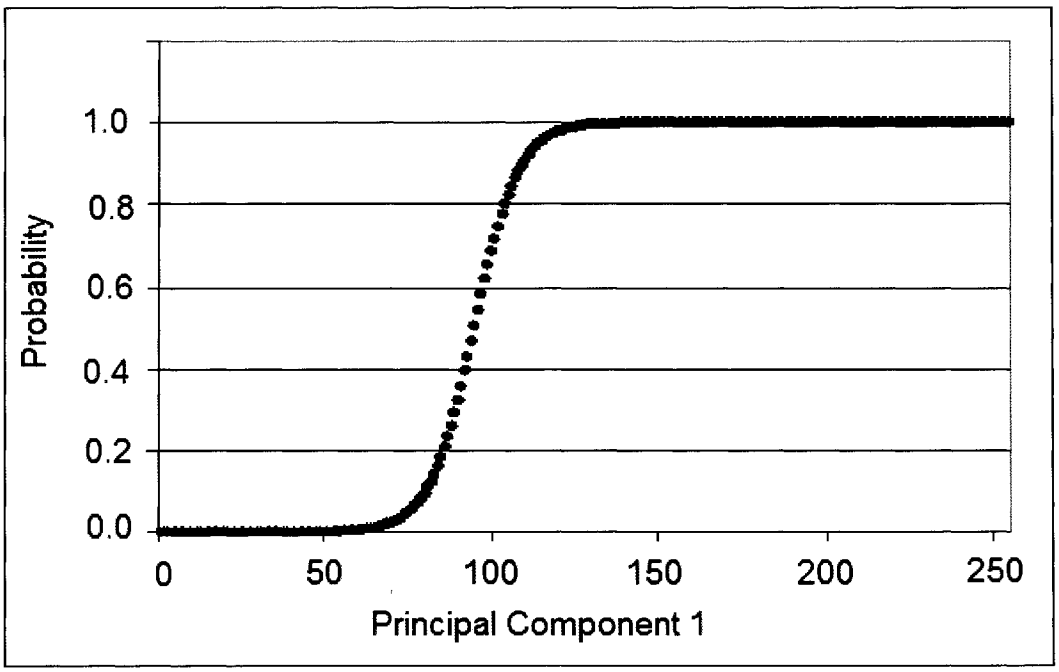


Figure 26. Probability curve corresponding to the binary logistic regression model (Equation 3) predicting the presence of hooded warbler nesting habitat.

The distribution of the probabilities (Table 17), similar to the maximum likelihood *a posteriori* probabilities, showed 61.2% of the nest sites with high probability ($p > 0.70$), but the remaining nests were spread throughout the other probabilities without a cluster in the lowest ranges. When the logistic regression probabilities were compared to the maximum likelihood probabilities they showed quite good agreement with a Pearson correlation of 0.82 ($p = 0.000$).

Table 17. Distribution of logistic regression predicted probabilities for validation nest sites.

	Numbers of Nests	% of Nests
0-10%	1	0.6
10-20%	7	4.5
20-30%	14	8.9
30-40%	13	8.3
40-50%	5	3.2
50-60%	16	10.2
60-70%	5	3.2
70-80%	8	5.1
80-90%	26	16.6
90-100%	62	39.5

10.2 Mapping with Landsat Imagery

10.2.1 Spectral Unmixing

The three endmembers (Table 18) were more difficult to manually delineate in the Landsat imagery due to the coarser spatial resolution. The unmixed Landsat fraction images (Figure 27), corresponded well to features visually identified in the imagery. Figure 27 (a) shows the same area as in Figure 21 with the three areas (A-C). Although clearly coarser and harder to discern, these features are definitely visible within the three sub-pixel fraction images.

Table 18. Average DN of the three endmembers used in spectral unmixing of the Landsat imagery.

Endmember	Blue	Green	Red	NIR	MIR 1	MIR 2
GPV	61.40	27.43	19.27	150.58	86.08	23.85
Soil	115.59	66.08	91.75	100.11	186.25	120.16
Shadow	61.90	21.68	17.17	10.14	9.36	5.65

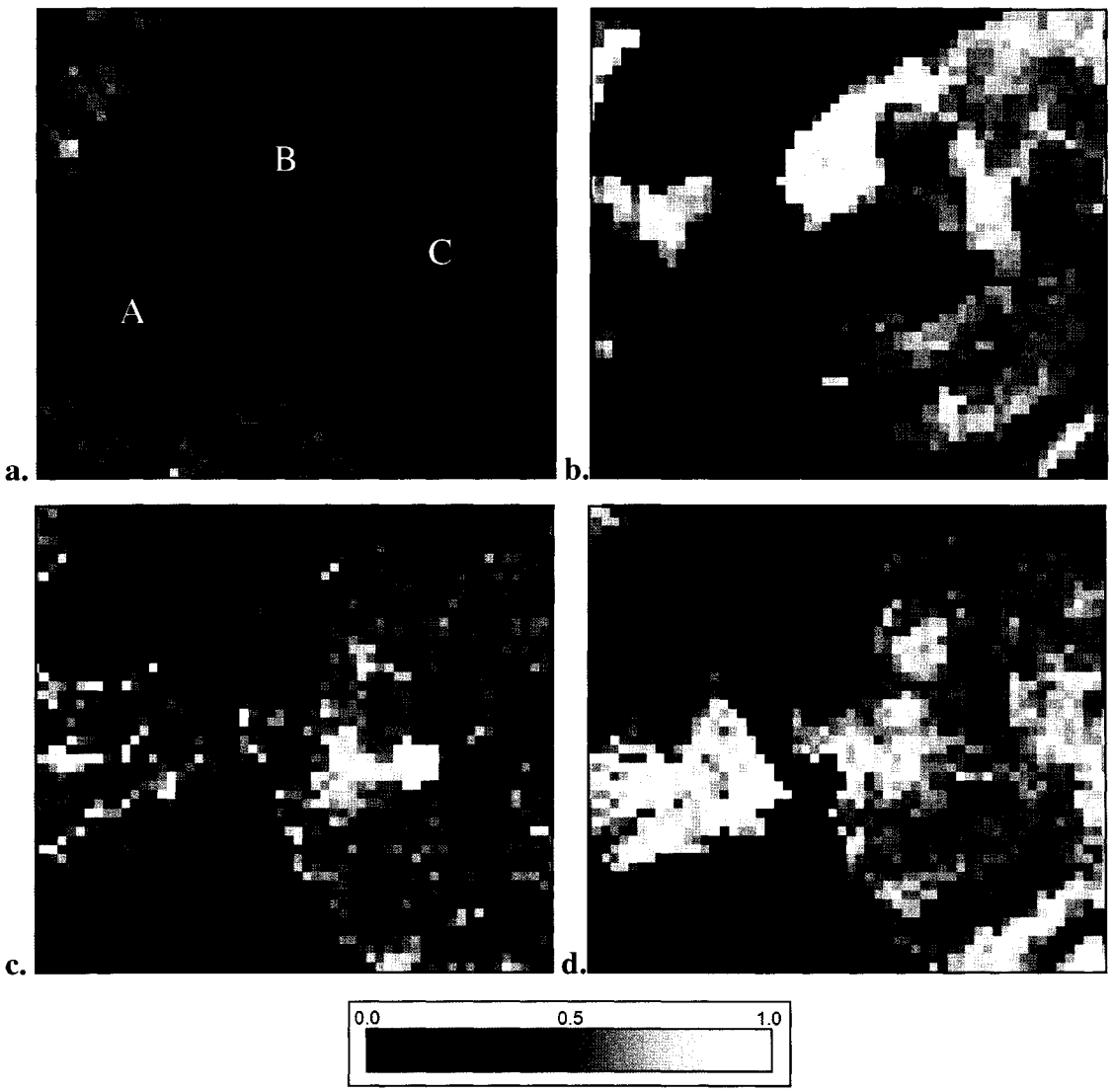


Figure 27. Landsat imagery (a) fraction images for GPV (b), soil (c), and shadow (d), created through constrained least squares linear spectral unmixing using endmembers selected manually from the imagery.

10.2.2 Comparison of Nest and Non-Nest Sites

The results of comparing the four useable Landsat spectral variables, based on a 3 x 3 pixel window, between nest and non-nest sites showed that three were significantly different using $p \leq 0.05$ and the fourth was significantly different using $p \leq 0.10$ (Table 19). All four variables were higher at non-nest sites, which was as expected.

Table 19. Mann-Whitney U tests comparing nest (n = 279) and non-nest (n = 63) sites in terms of spectral variables extracted from Landsat imagery.

	Nest Sites		Non-Nest Sites		U	p
	Avg.	Stdev.	Avg.	Stdev.		
Green 3x3 Average	25.75	1.26	26.90	2.40	6139.0	**0.000**
Red 3x3 Average	18.98	1.42	20.90	2.40	4788.0	**0.000**
NIR 3x3 Average	110.24	17.96	116.03	14.43	7455.5	**0.060**
MIR 2 3x3 Average	20.94	2.527	23.40	4.229	5784.0	**0.000**

When comparing the textures, it was found that five of the six red, two NIR, and two MIR textures showed significant differences between nest and non-nest sites (Table 20). In all cases, the non-nest sites showed greater canopy heterogeneity compared to the nest sites. This was opposite to expected and to what was found previously with the Ikonos analysis.

With regards to the fraction images, the soil and shadow fractions showed significant differences between nest and non-nest sites (Table 21). The soil fractions for both groups were extremely low, with slightly more bare soil found at non-nest sites, perhaps suggesting less ground vegetation. The shadow fraction was significantly greater at nest sites, as expected, and while the differences for shadow and soil were small they were significant as a result of the small variance, whereas for GPV, higher variance resulted in no significant differences.

Principal components were created using the nine image variables ($p \leq 0.05$) and then fifteen image variables ($p \leq 0.10$), which were found to have significant differences between nest and non-nest sites. Table 22 shows the PCs created along with the corresponding percent variance attributed to each. In the first PCA, the first four PCs were considered for further analysis, based on the percent variance and as well visual

analysis. For the second analysis, five components were saved for further analysis. However, in both cases, only the first three components were thought to hold useful information for differentiating nest from non-nest sites, based on the percent variance explained by each component.

Table 20. Mann-Whitney U tests comparing nest (n = 279) and non-nest (n = 63) sites in terms of textural variables extracted from Landsat imagery using a 3 x 3 window.

	Nest Sites		Non-Nest Sites		U	p
	Avg.	Stdev.	Avg.	Stdev.		
Green Contrast	5.85	11.20	13.18	22.33	7946.0	0.234
Green Dissimilarity	1.47	1.20	2.16	2.24	7990.5	0.258
Green Entropy	1.36	0.43	1.40	0.46	8356.5	0.539
Green Homogeneity	0.56	0.26	0.52	0.28	8164.0	0.378
Green ASM	0.32	0.16	0.30	0.17	8408.0	0.589
Green Correlation	-0.20	0.41	-0.17	0.37	8524.5	0.709
Red Contrast	7.26	18.80	19.26	36.25	7565.5	**0.084**
Red Dissimilarity	1.55	1.45	2.49	2.60	7497.0	**0.067**
Red Entropy	1.26	0.45	1.43	0.41	6971.5	**0.010**
Red Homogeneity	0.54	0.26	0.47	0.24	7416.0	**0.057**
Red ASM	0.34	0.19	0.28	0.13	7007.5	**0.013**
Red Correlation	-0.24	0.41	-0.24	0.41	8637.5	0.865
NIR Contrast	5.30	8.17	5.87	5.93	7655.0	0.109
NIR Dissimilarity	1.54	1.13	1.71	1.01	7567.0	**0.083**
NIR Entropy	1.57	0.44	1.61	0.38	8639.5	0.832
NIR Homogeneity	0.50	0.22	0.46	0.19	7714.5	0.140
NIR ASM	0.25	0.17	0.24	0.14	8683.0	0.916
NIR Correlation	-0.24	0.36	-0.33	0.36	7433.0	**0.061**
MIR 2 Contrast	67.07	87.81	95.11	116.88	7669.0	0.114
MIR 2 Dissimilarity	5.52	4.08	6.33	4.56	7940.5	0.231
MIR 2 Entropy	1.61	0.37	1.63	0.41	8170.0	0.377
MIR 2 Homogeneity	0.71	0.26	0.61	0.25	6800.5	**0.005**
MIR 2 ASM	0.52	0.35	0.36	0.25	6898.0	**0.008**
MIR 2 Correlation	-0.17	0.31	-0.18	0.38	7779.5	0.159

Table 21. Mann-Whitney U tests comparing nest (n = 279) and non-nest (n = 63) sites in terms of sub-pixel fractions extracted from Landsat imagery.

	Nest Sites		Non-Nest Sites		U	p
	Avg.	Stdev.	Avg.	Stdev.		
GPV 3x3 Average	0.68	0.14	0.71	0.13	8145.0	0.364
Soil 3x3 Average	0.03	0.02	0.04	0.03	6648.0	**0.002**
Shadow 3x3 Average	0.21	0.12	0.17	0.10	7291.5	**0.035**

Table 22. Factors created through a PCA of spectral, textural, and sub-pixel fraction variables that showed significant differences between nest and non-nest sites.

$p \leq 0.05$			$p \leq 0.10$	
Factor	**% Variance**		**Factor**	**% Variance**
PC1	**49.11**		**PC1**	**42.98**
PC2	**39.24**		**PC2**	**30.39**
PC3	**5.96**		**PC3**	**11.32**
PC4	**2.43**		**PC4**	**5.66**
PC5	1.79		**PC5**	**4.08**
PC6	0.94		**PC6**	2.27
			PC7	1.24
			PC8	0.84

The factor loadings for the components from the first PCA (Table 23) showed that the first three PCs represented various red and mid-infrared textures, while the fourth strongly represented shadows. In the second PCA (Table 24), the first component represented red textures, the second represented MIR textures, and the third strongly represented NIR textures. The fourth and fifth components showed weak relations with a red texture and a couple of MIR textures respectively. Spectral brightness did not account for any significant variance in the data. This analysis used unstandardized PCs as opposed to standardized PCs, which may have caused variables with greater variance, such as texture, to dominate the analysis. However, for data compression purposes unstandardized components are most common (Eastman, 2001).

Table 23. Factor loadings for independent variables, extracted from Landsat imagery, using only those showing significant differences ($p \leq 0.05$) between nest and non-nest sites using.

	PC1	PC2	PC3	PC4
Green	-0.22	-0.08	0.02	-0.19
Red	-0.30	-0.04	-0.10	0.04
MIR 2	0.01	0.04	-0.04	0.13
Red Entropy	**-0.60**	**-0.76**	0.00	0.00
Red ASM	**0.57**	**0.75**	-0.02	0.00
MIR 2 Homogeneity	**0.82**	**-0.46**	**0.39**	0.00
MIR 2 ASM	**0.76**	**-0.52**	**-0.29**	-0.04
Soil	-0.15	0.09	-0.09	0.30
Shadow	0.34	-0.22	-0.13	**0.89**

Table 24. Factor loadings for independent variables, extracted from Landsat imagery, using only those showing significant differences ($p \leq 0.10$) between nest and non-nest sites.

	PC1	PC2	PC3	PC4	PC5
Green	-0.24	-0.03	0.04	0.04	0.04
Red	-0.30	-0.12	0.03	0.10	-0.10
NIR	-0.01	-0.29	0.05	-0.08	0.21
MIR 2	0.03	-0.03	0.00	-0.01	-0.05
Red Contrast	**-0.44**	0.03	-0.04	0.26	-0.06
Red Dissimilarity	**-0.70**	0.13	-0.05	0.38	-0.02
Red Entropy	**-0.86**	0.33	0.03	-0.25	-0.03
Red Homogeneity	**0.84**	-0.28	0.05	**-0.45**	-0.07
Red ASM	**0.81**	-0.34	-0.03	0.26	0.01
NIR Dissimilarity	-0.28	-0.14	**-0.34**	0.04	-0.15
NIR Correlation	0.05	0.11	**0.99**	0.07	-0.05
MIR 2 Homogeneity	0.45	**0.83**	0.02	-0.05	**0.38**
MIR 2 ASM	0.40	**0.81**	-0.08	0.05	**-0.28**
Soil	-0.08	-0.15	0.00	0.02	-0.11
Shadow	0.18	0.36	-0.07	0.05	-0.17

10.2.3 Maximum Likelihood Classification

Following training for the nest and non-nest sites, the Transformed Divergence separabilities (Lillesand and Kiefer, 2000) using the PCs from the above analyses were found to be 0.20 and 0.07, respectively. These values were extremely low, indicating that the training data for the two classes were extremely overlapping in the multivariate spectral-texture space. This suggested that classifications created using these training data would not be very accurate or useful.

The first classification classified 49.9% of the Landsat forest coverage as nesting habitat compared to 43.8% by the second classification. The error matrices for the two classifications showed mixed results for the training sites and the validation sites. In the first case (Table 25) classification of the training sites and the validation sites was not very good, with an overall classification accuracy of 69.7% and 69.5% respectively.

The second classification (Table 26) was worse, with an overall accuracy of 56.2% and 62.8% for the training and validation sites respectively.

These two habitat maps were extremely grainy in many places and did not appear to provide accurate representations of habitat and non-habitat in known areas. Figure 28 shows five examples of the first classification for areas where there are known hooded warbler populations.

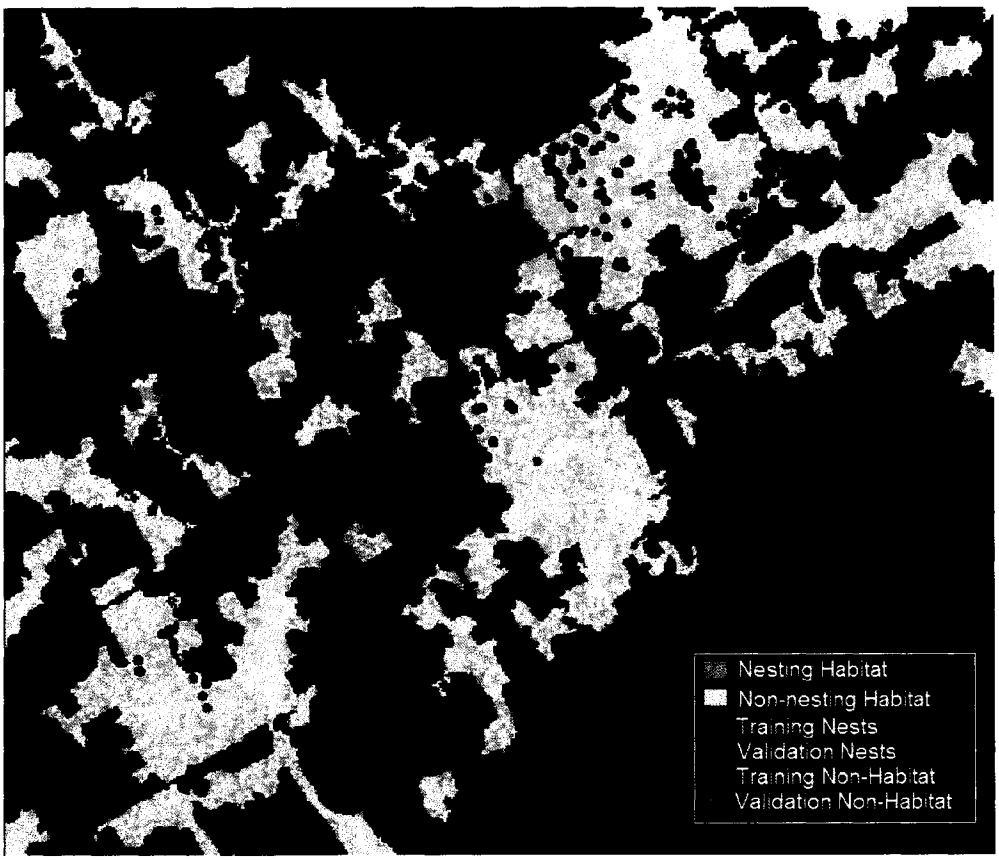
Table 25. Error matrix for training (a) and validation (b) sites for the habitat/ non-habitat map created using the first set of Landsat principal components.

		Image Class		
	(a)	**Nesting Habitat**	**Non- Habitat**	
Field Class	**Nests**	101 (72%)	39	
	Non-Habitat	15	23 (60%)	
				69.7%
	(b)	**Nesting Habitat**	**Non- Habitat**	
	Nests	101 (73%)	38	
	Non-Habitat	12	13 (52%)	
				69.5%

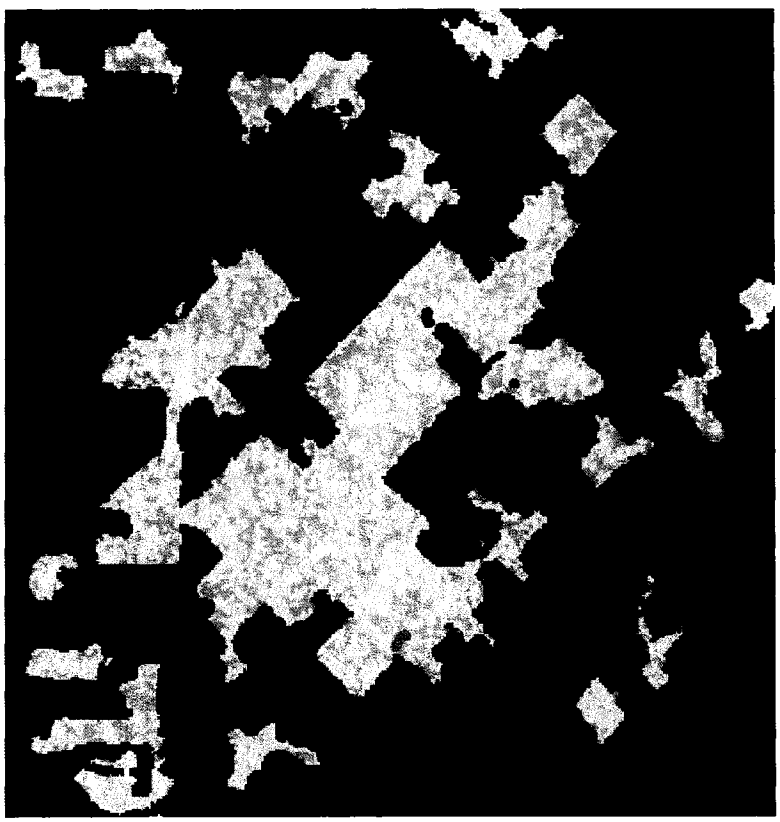
Table 26. Error matrix for training (a) and validation (b) sites for the habitat/ non-habitat map created using the second set of Landsat principal components.

		Image Class		
	(a)	**Nesting Habitat**	**Non- Habitat**	
Field Class	**Nests**	85 (61%)	55	
	Non-Habitat	23	15 (39%)	
				56.2%
	(b)	**Nesting Habitat**	**Non- Habitat**	
	Nests	85 (61%)	54	
	Non-Habitat	7	18 (72%)	
				62.8%

a.



b.



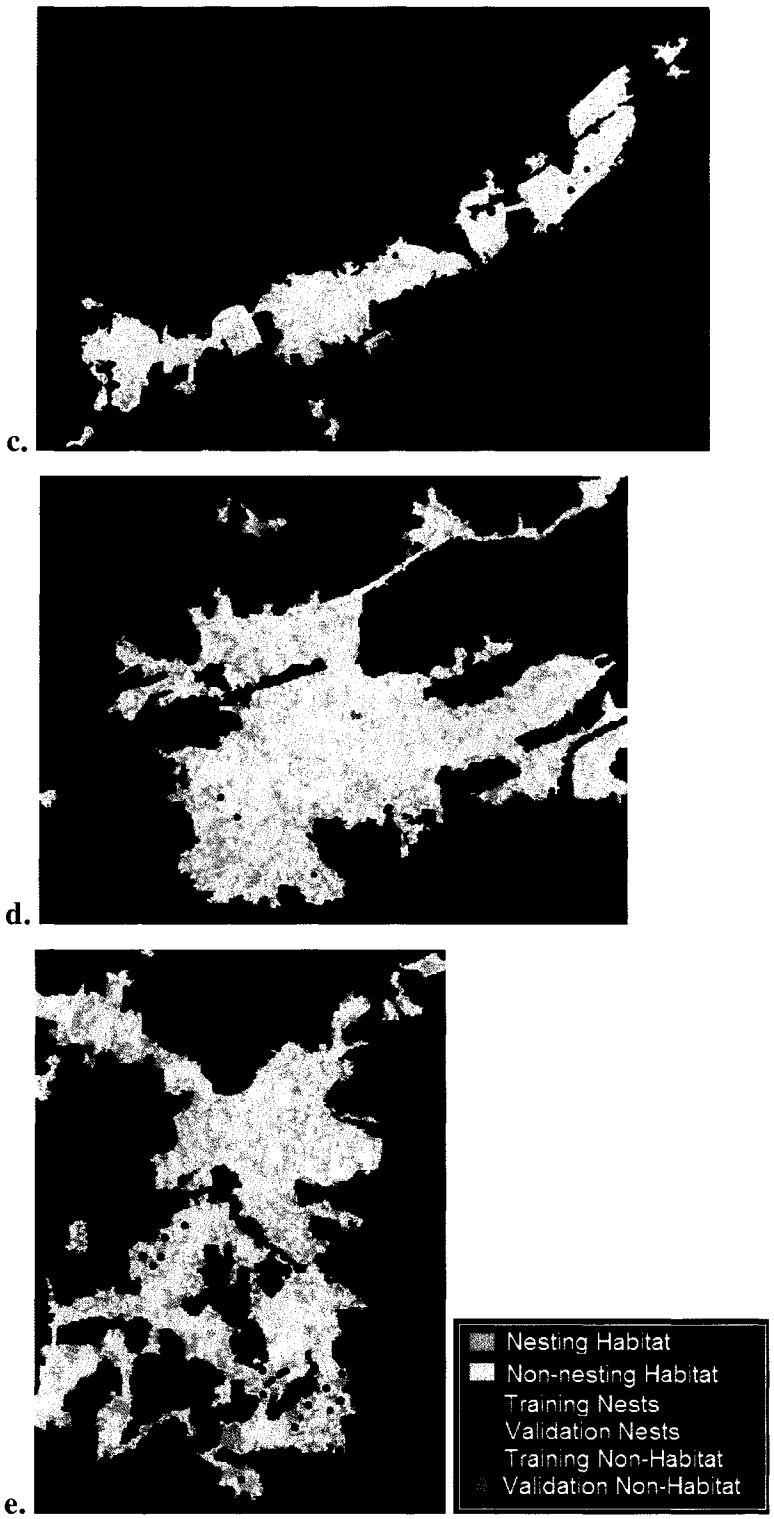


Figure 28. Examples of habitat mapping results for five well known forests across the study area: a. St Williams Forest, Backus Woods and South Walsingham Forest, b. Skunks Misery northeast of Chatham, c. The Pinery Provincial Park along the shore of Lake Huron, d. Dundas Valley near Hamilton, and e. Twelve Mile Creek near St. Catherines.

A close-up of St. Williams (Figure 29) showed that most of St. Williams Forest was classified as nesting habitat. Therefore, a second classification was conducted using training data only from nest sites that were visited in the field, along with the non-nesting transect. This was done to see how the classification would perform in a local target area using the Landsat imagery. Figure 30 shows the resulting map, which was assessed using one-hundred and forty-nine nest sites distributed throughout the image, and had 87% correctly classified. As was the case for the Ikonos mapping of the same area, quantitative assessment of non-nesting site accuracy was not possible with these data because no validation non-nesting sites were available. While the only definite non-nesting sites visited were located along the transect, the visibly clustered pattern of known nest sites in this forest suggested that there were indeed other areas that did not provide the necessary nesting conditions for the hooded warbler. This map shows the non-nesting transect (blue transect in Figure 30) as unsuitable nesting habitat, and areas unoccupied by nest sites (same red circles as in Figure 25) were also classified as non-nesting habitat.

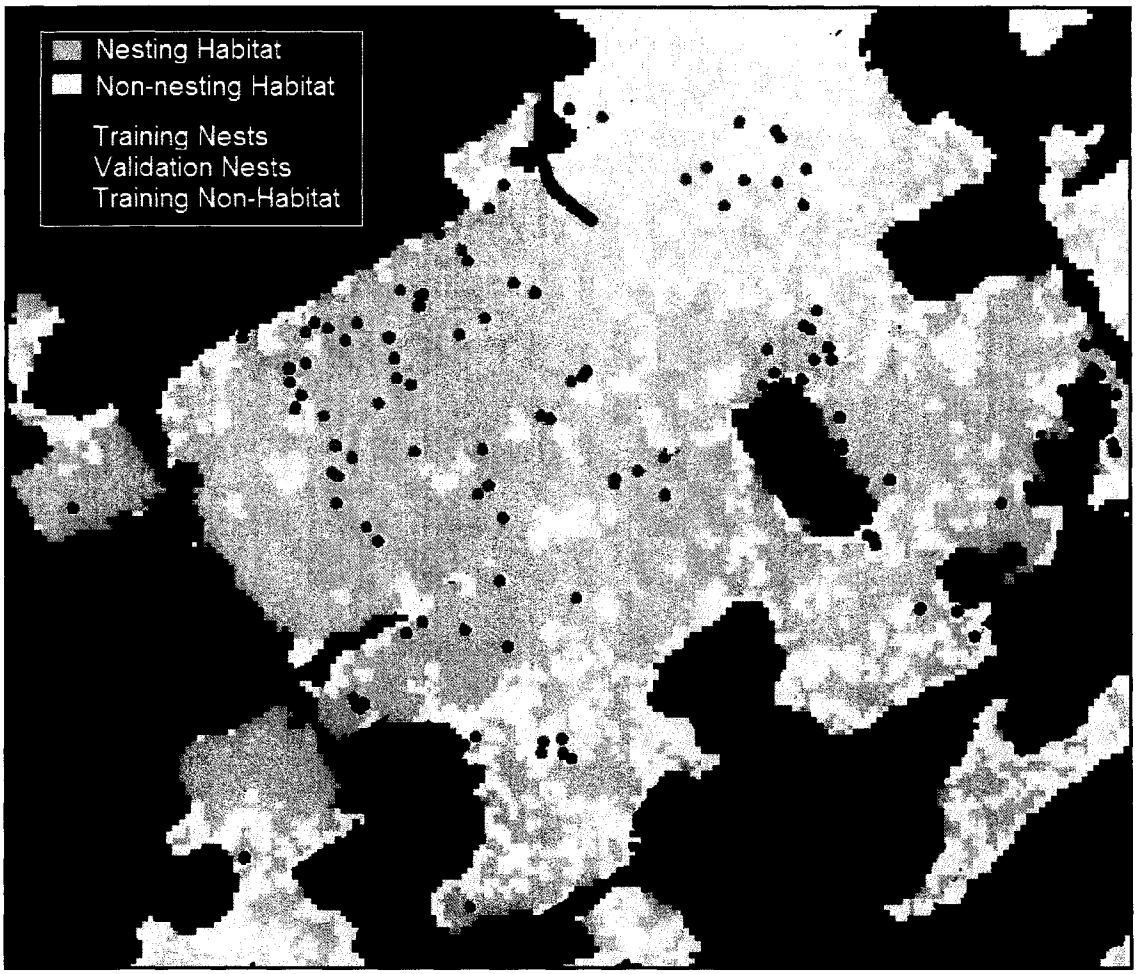


Figure 29. Close up of St Williams Forest showing classification results created using training data from across the entire study area.

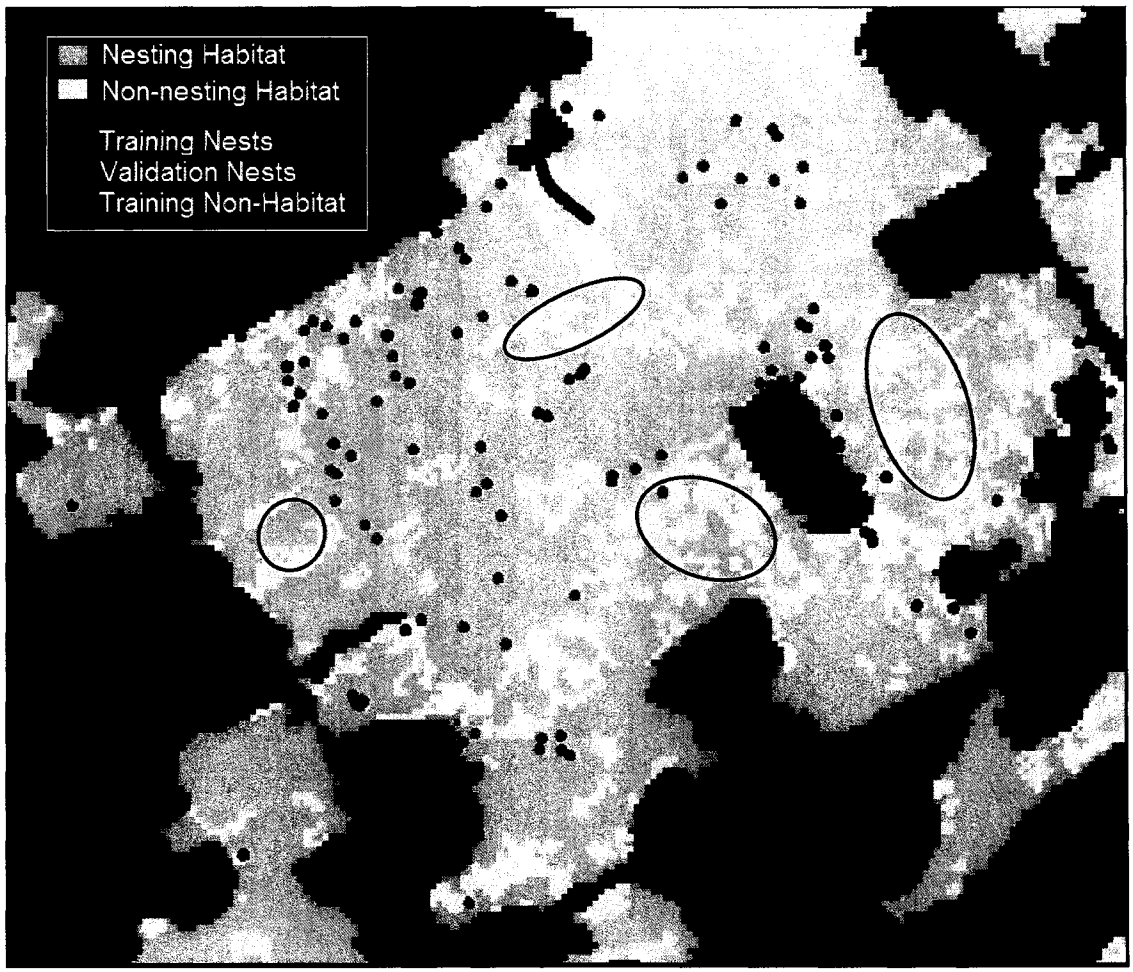


Figure 30. St Williams Forest results created using training data only from the area corresponding to the high resolution Ikonos imagery. Areas outlined in red are classified as non-nesting habitat and were previously unused for nesting.

10.2.4 Binary Logistic Regression

The results produced from regression modelling were disappointing, but perhaps simply showed similar results to the classifications. Using all available nest and non-nesting sites across southern Ontario, no significant models were created using either of the two sets of principal components. It was hypothesized that the imbalance of presence/absence information was causing this so 63 nest sites were randomly selected from the database to match the 63 non-nesting sites. With these input data, using the first set of PCs, a very weak model was produced that included the first three PCs. This model had a Nagelkerke r^2 of 0.178 and correctly predicted 74.6% of the non-nesting sites, but only 65% of the nest sites. As a test, a second random sample of 63 nest sites were selected and the resulting model showed an r^2 of 0.106, with 58.7% and 61.9% of non nest and nest sites correctly predicted. The second set of PCs produced a weaker model with lower accuracies.

11.0 DISCUSSION

11.1 Significant Findings

Overall, this analysis showed that while there may be significant relations of image spectral and textural measures with canopy gaps measured with hemispherical photographs, the actual gap sizes and shading caused by the gaps are not seen in the same way from below the canopy as they are from above the canopy. When hemispherical photographs were compared to the high resolution panchromatic imagery, with a ground pixel size of 1 m x 1 m, little agreement was seen. Discrepancies between the two measures can be attributed to differential tree heights in the canopy as well as view-angle

/ sun-angle effects in the imagery, both which cause differential shading. Figure 31 demonstrates this situation. A photograph taken under nearly closed canopy (A) would show at the most very small gaps, while imagery taken from above would consist mostly of off nadir data and see the shadows caused by differential tree height (noted just below A).

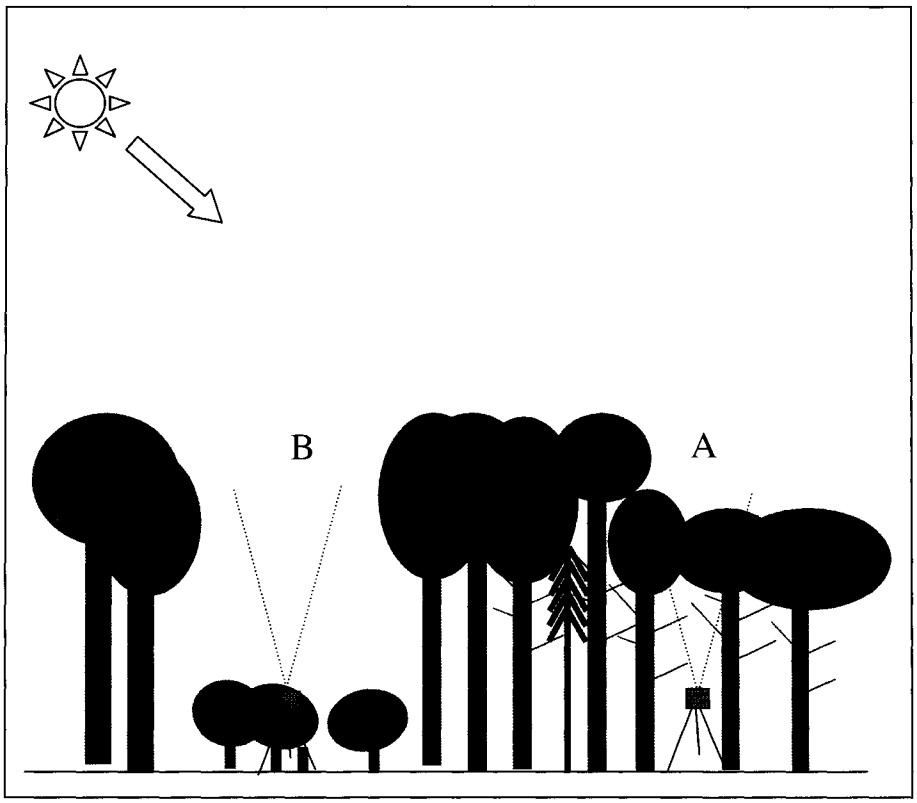


Figure 31. Demonstration of discrepancies between a vertical photograph and remotely sensed imagery. (A) Vertical photography images small gaps in closed canopy while overhead sensor would image shadow (S) due to differential tree height. (B) A large gap would dominate a vertical photograph while the sensor would detect illuminated ground vegetation

As well, the lack of agreement between the hemispherical photographs and the remotely sensed imagery was partially attributed to very dense understory within large gaps. When looking downwards from the sensor, while some deep shading can be seen,

some sites with extremely dense understory vegetation or directly illuminated understory most likely caused the canopy to appear more homogeneous than it really was (Figure 31, B).

A significant relation was found between the hemispherical photographs and variables extracted from the high resolution Ikonos imagery, and a simple two variable model for predicting percent cover with an R^2 of 0.48 and a standard error of 14% was produced. This showed similar results to a more complex model created with Landsat imagery of similar deciduous forest in eastern Ontario, which had an R^2 of 0.50 and a standard error of 12% for a much more limited range of cover (M. Lindsay, *unpublished*, 2005).

The positive correlation found between the sub-pixel shadow fraction and percent cover was the opposite of that expected, in the sense that gaps in the canopy, which reduce forest cover, were thought to have greater proportions of shadow compared to more homogeneous canopy. However, Seed and King (2003), using airborne imagery, showed that greater shadow fractions were associated with increased canopy closure as a result of more biomass (i.e. more branches, leaves) and a reduction in crown transparency.

The results of high resolution habitat mapping using Ikonos imagery showed 70% of the validation nest sites being situated in areas classified as nesting habitat. While this number was not as high as was desired, an additional 18% of the nest sites that had been classified as non nest habitat were found to be located within 10 m of classified nesting habitat. This 10 m error, to be conservative, can easily be attributed to the GPS measurements taken at these sites by members of Bird Studies Canada or volunteers with

the Ontario Breeding Bird Atlas surveying program. Unlike the sites surveyed in 2004 by the author, the geographic positions of these other nest sites were not differentially corrected (D. Badzinski, *pers. comm.*, 2004) and therefore could have easily been off by 10 m or more.

Mapping nesting habitat for hooded warblers across southern Ontario proved difficult. While significant differences in Landsat spectral and spatial information existed between nest and non-nest sites across the region, the classified maps predicted nesting and non-nesting habitat with less than 70% overall accuracy. Known nest sites were well classified, but approximately one-half of the non-nest validation sites were classified as suitable. These results may have been partially due to the disproportionate number of training sites within St. Williams Forest and the immediate surroundings, which may have dominated the training data distributions. The overall result was an overestimation of potential habitat, similar to Flaxman's (2004) hooded warbler habitat mapping research.

When the area from the Landsat imagery which corresponded to the Ikonos coverage was examined separately, fairly good results were obtained (correctly predicted 87% of validation nest sties). Whittam and McCraken (1999) did some hooded warbler habitat research in St. Williams Forest and the neighbouring South Walsingham Forest and they found significant differences between these two forests. For example, nest sites in South Walsingham were characterized by older gaps with lower shrub density surrounding nests than in St. Williams. As well, St. Williams generally consisted of a larger size-class of trees compared to South Walsingam. These various differences in forest structure could perhaps cause confusion in South Walsingham Forest, given that

the bulk of the data used for modelling was within St. Williams Forest. If such confusion could occur at this local scale, clearly forests across Ontario may have differed greatly from St. Williams Forest as well. While St. Williams Forest is currently the most populated forest in terms of hooded warblers in southern Ontario, it is definitely unique structurally (J. McCraken, *pers. comm.*, 2004), and any modelling done based solely on forest structure may be hampered by structural differences between forests.

St. Williams Forest actually looked different from the surrounding forests based on a visual analysis of the imagery. Perhaps visual interpretation of the largest forests in southern Ontario might be a better method for identifying potential habitat compared to automated methods.

Given the argument that a significant number of hooded warblers return to nesting and territory sites from one season to the next (Howlett and Stutchbury, 2003), it is difficult to say that the areas where nests have not been found in the past few years are definitely not suitable nesting habitat. These areas may either be suitable nesting habitat that simply have not been used. Hooded warblers tend to return to sites used previously, territorial size requires that some distance be maintained between nests, and the low numbers of birds in the area makes it impossible for all suitable areas to be used. Field validation is definitely required in order to properly check these areas.

In southern Ontario, there are only a few large continuous forests remaining. The database of hooded warbler nest sites corresponds quite well with these forests (St. Williams and surrounding area, Skunks Misery, The Pinery Provincial Park, Dundas Valley, and Twelve Mile Creek), and while the underlying goal of this research was to identify other forests across Ontario that show potential as hooded warbler nesting

habitat, it may be affordable and more beneficial to acquire high resolution Ikonos imagery for each of these sites in order to examine and map their habitat in more detail. Although difficult to define an exact threshold, past research has shown that hooded warblers will not live in small forest patches, which eliminates most of the forests across southern Ontario. Due to effects of mapping errors discussed in Section 9.3.3, it is difficult to accurately measure forest patch sizes based solely on a classified satellite image, and therefore difficult to apply an automated size threshold to identify patch sizes greater than a given size, aside from simply eliminating the very small patches. Flaxman (2004) showed that 96% of evaluated hooded warbler occurrences were in forest patches greater than 100 ha. If a threshold of this magnitude is used, not many forests remain across southern Ontario, and therefore perhaps it would be more beneficial to carry out detailed local analyses using high resolution imagery instead of regional analysis.

11.2 Research Limitations and Recommendations

Mapping habitat for a species can be done using only presence data, however errors of commission, meaning areas that are classified as habitat but are not, tend to be extremely high. Flaxman's (2004) hooded warbler habitat mapping suffered from these problems, however, this research attempted to use non-habitat sites in order to identify unsuitable areas, rather than simply rely on habitat sites to show habitat suitability, with the extreme outliers ending up as unsuitable areas.

Identifying non-habitat sites is very difficult and usually not the focus of field surveys. More often than not, researchers simply keep track of where a particular species has been found, but do not actively seek out areas that are completely unsuitable. Of

course, for a forest dwelling species such as the hooded warbler, eliminating all non-forested areas from the analysis is a simple, but key step, however, unsuitable areas must be also be identified at a finer scale within the forest.

Many researchers simply use presence data and identify areas in the landscape with similar spatial characteristics as the known presence locations. These methods are often referred to as "enveloping", and can be done using crisp or fuzzy models (Robertson et al., 2004). Predicting suitable habitat areas can be done by thresholding a set of predictor variables (using minimum and maximum values) based on characteristics obtained from known locations and combining the binary suitability maps in order to locate suitable areas. This basic method uses crisp boundaries in the sense that areas are either suitable or not suitable, and since this is often a gross simplification of reality, fuzzy boundary methods have been created to provide more realistic results. Fuzzy classifications rely on graded membership functions for each variable that can be derived using minimum, maximum, and median values, which can then be combined to create a fuzzy habitat distribution map showing levels of membership (Robertson et al., 2004).

Others have attempted to make use of what is known as "pseudo-absence" data for habitat mapping. Basically, when species absence data is not available, random points are generated in areas where species presence data is missing (e.g. GARP algorithm, Stockwell and Peters, 1999). The problem with this approach is that simply because a species was not detected in a certain area does not mean that that area is not perfectly suitable habitat (Brotons et al., 2004). Brotons et al. (2004) compared an algorithm that used presence only data with one that used presence-absence data for mapping breeding

bird distributions in an area of Spain and suggested that absence data should be included whenever possible.

Given all of this, a major limitation to this research, which has been previously mentioned in Part A of this thesis, was the lack of sufficient non-nesting habitat sites for training and validating the habitat maps produced. Within the immediate vicinity of St. Williams Forest, no non-nesting habitat sites were found other than the transect that was used for training the models. Ideally positions of non-nest sites would have been collected throughout this area, and more sites collected across all of southern Ontario. Most of the sites that were visited across the region were in small pockets of forest, as few large forests remain and many forests were not accessible, being at the backs of private farms.

Other spatial variables could also potentially have been included in this analysis had sufficiently detailed data been available. Pither (1997) made use of soil types for hooded warbler habitat mapping, and found it to be a useful predictor. While general soil maps do exist, more detailed variations in soil types and soil moisture information could be investigated for localized prediction of nesting habitat. As well, as mentioned briefly in Part A, a very detailed digital elevation model could be used in order to account for micro-topographic effects on ground vegetation growth.

Future remote sensing research for hooded warbler habitat mapping should include radar and/or lidar data, which have been shown to provide excellent information with regards to forest structure and therefore canopy gap characteristics.

12.0 THESIS SUMMARY AND CONCLUSIONS

The results of this research show strong evidence that hooded warbler nest sites have significantly lower canopy cover compared to non-nest sites. As well, this research showed that the maximum gap size at nest sites was significantly larger, and that non-nest sites had approximately the same number of gaps as nest sites, but they did not have the large canopy gap that was typically found at nest sites.

These results were obtained using hemispherical photographs, which provided a sophisticated and repeatable method for measuring forest canopy cover. Information with regards to the approximate gaps size distribution above nest sites was investigated, providing information that could potentially be used for forest management, in order to ensure suitable hooded warbler habitat will exist. Hooded warblers have been shown to colonize logged and managed forests, as a direct result of the overhead canopy gaps created by selective logging. Removing trees opens up the canopy, allowing increased light to reach the forest floor, which in turn allows ground vegetation to grow freely providing the necessary habitat for hooded warbler nesting. From a management perspective, understanding how much cover could be removed in order to encourage hooded warblers to inhabit specific forests is potentially an important part of increasing populations in Canada to a sustainable level. However, it must be kept in mind that managing a forest for a specific species may affect the survival of other species.

With evidence shown that the canopy structure differed significantly between nest and non-nest sites, satellite imagery was used to attempt to map potential hooded warbler habitat across southern Ontario. Remotely sensed data have been previously used to map forest structure at different scales, and this research tested the use of high and medium

resolution imagery for mapping habitat. High resolution Ikonos imagery provided additional maps of a local area that successfully showed the spatial pattern of nest and non-nest habitat. However, the high costs and many scenes required for mapping large areas make it virtually impossible to use high resolution imagery for anything more than targeted analysis of areas of interest. As well, in order to truly test these methods, a significantly larger sample of non-nest sites are needed for training and validation.

Landsat imagery was tested in order to provide an affordable means of mapping potential habitat across the Canadian range of the hooded warbler. While image spectral, textural, and sub-pixel information showed significant differences between nest and non-nest sites, the resulting maps suffered from high errors of commission for nesting habitat. However, local mapping of St. Williams Forest using the Landsat imagery showed similar patterns to those seen in the Ikonos map. It is therefore concluded that mapping hooded warbler nesting habitat with Landsat would be best conducted in local areas as discrete regions, requiring more detailed ground information for developing and validating the maps.

To help ensure the successful survival of hooded warblers in Canada, more work must be done in order to identify and map forests that contain sufficient nesting habitat. Remote sensing provides a means to map forest canopies, however more work must be done on the ground in order to link canopy characteristics seen in the imagery with ground vegetation present as a result of the overhead canopy structure. Other remote sensors, such as lidar or radar, could potentially be used in order to provide more detail with regards to vertical forest structure. As well, other environmental variables of importance to hooded warblers could potentially be combined with forest structure and

canopy gap information as part of a more robust modelling effort for predicting hooded warbler nesting habitat.

REFERENCES

Allair, J., D. Badzinski, B. Whittam, and J. McCracken. 2002. Southern Ontario hooded warbler Research: 2001 Final Report. Bird Studies Canada.

Annand, E.M. and F.R. Thompson. 1997. Forest bird response to regeneration practices in central hardwood forests. Journal of Wildlife Management 61(1): 159-171.

Anonymous. 2003. Species at Risk Act given royal assent. The Forestry Chronicle 79(2):181.

Asner, G.P., J.A. Hicke, and D.B. Lobell. 2003. Per-pixel analysis of forest structure. In Remote Sensing of Forest Environments: Concepts and case studies. Edited by M.A. Wulder and S.E. Franklin. Kluwer Academic Publishers. Boston.

Asner, G.P. and A.S. Warner. 2003. Canopy shadow in Ikonos satellite observations of tropical forests and savannas. Remote Sensing of Environment 87(4): 521-533.

Badzinski, D.S. 2003. Hooded warbler research in St. Williams Forest, Ontario: An investigation of nest productivity, nest concealment, territory size and species associations. Bird Studies Canada / Canadian Wildlife Service.

Baker, M.D. and M.J. Lacki. 1997. Short-term changes in bird communities in response to silvicultural prescriptions. Forest Ecology and Management 96(1): 27-36.

Bastin, L. 1997. Comparison of fuzzy c-means classification, linear mixture modelling and MLC probabilities as tools for unmixing coarse pixels. International Journal of Remote Sensing 18(17): 3629-3648.

Bechtel, R., A. Sanchez-Azofeifa, B. Rivard, G. Hamilton, J. Martin, and E. Dzus. 2004. Associations between woodland caribou telemetry data and Landsat TM spectral reflectance. International Journal of Remote Sensing 25(21): 4813-4827.

Bezdek, J.C. and S.K. Pal, editors. 1992. Fuzzy models for pattern recognition: Methods that search for structures in data. IEEE Press. New York.

Bisson, I.A. and B.J.M. Stutchbury. 2000. Nesting success and nest-site selection by a neotropical migrant in a fragmented landscape. Canadian Journal of Zoology 78(5): 858-863.

Borel, C.C. and S.A.W. Gerstl. 1994. Are leaf chemistry signatures preserved at the canopy level? In Proc. IGARSS' 94 International Geoscience and Remote Sensing Symposium, eds. IEEE, 996-998. Pasadena, CA., 8-12 August.

Braun, M. and M. Herold. 2003. Mapping imperviousness using NDVI and linear spectral unmixing of ASTER data in the Cologne-Bonn region Germany, Proceedings of SPIE 2003, Barcelona, Spain.

Brotons, L., W. Thuiller, M.B. Araujo, and A.H. Hirzel. 2004. Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. Ecography 27(4): 437-448.

Brown, D.G. 2001. A spectral unmixing approach to leaf area index (LAI) estimation at the alpine treeline ecotone. In GIS and Remote Sensing Applications in Biogeography and Ecology. Edited by A.C. Millington, S.J. Walsh, and P.E. Osborne. Kluwer Academic Publishers, Boston.

BSC (Bird Studies Canada). 2005. http://www.bsc-eoc.org.

Butson, C. and D.J. King. 1999. Semivariance analysis of forest structure and remote sensing data to determine an optimal sample plot size. Proceedings 4th International Airborne Remote Sensing Conference and Exhibition (Env. Res. Inst. of Michigan) / 21st Canadian Symposium on Remote Sensing (Canadian Remote Sensing Society), Ottawa, Ontario, June 21-24. Vol. II, pp. 155-162.

Carolinian Canada. 2005. http://www.carolinian.org.

Chan, J.C.W., N. Laporte, and R.S. Defries. 2003. Texture classification of logged forests in tropical Africa using machine-learning algorithms. International Journal of Remote Sensing 24(6): 1401-1407.

Chander, G., D.L. Helder, B.L. Markham, J.D. Dewald, E. Kaita, K.J. Thome, E. Micijevic, and T.A. Ruggles. 2004. Landsat-5 TM reflective-band absolute radiometric calibration. IEEE Transactions on Geoscience and Remote Sensing 42(12): 2747-2760.

Chen, J.N., P.M. Rich, S.T. Gower, J.M. Norman, and S. Plummer. 1997. Leaf area index of boreal forests: theory, techniques and measurements. Journal of Geophysical Research 102: 29429-29443.

Coburn, C.A. and A.C. Roberts. 2004. A multiscale texture analysis procedure for improved forest stand classification. International Journal of Remote Sensing 25(20): 4287-4308.

COSEWIC (Committee on the Status of Endangered Wildlife in Canada). 2003. http://www.cosewic.gc.ca.

Cosmopolous, P. and D.J. King. 2004. Temporal analysis of forest structural condition at an acid mine site using multispectral digital camera imagery. International Journal of Remote Sensing 25(12): 2259-2275.

Darwin, A.T., D. Ladd, R. Galdins, T.A. Contreras, and L. Fahrig. 2004. Response of forest understory vegetation to a major ice storm. Journal of the Torrey Botanical Society 131(1): 45-52.

Delta-T Devices Ltd. 1999. Hemiview Canopy Analysis Software. Cambridge, UK.

Eastman, J.R. 2001. Idrisi32 Release 2. Clark Lab, Clark University. Worcester, MA.

Elliott, K.A., J. McCracken, and A. Couturier. 1999. A management strategy for South Walsingham Sand Ridges / Big Creak Floodplain Forest. Bird Studies Canada. http://www.bsc-eoc.org/download/swalsrep.pdf.

Environment Canada. 2003. Species At Risk Website. http://www.speciesatrisk.gc.ca.

Environment Canada. 2004. Environment Canada species at risk recovery program, Federal policy discussion paper: Critical habitat. http://www.sararegistry.gc.ca/virtual_sara/files/policies/Critical%20Habitat%20Discussion%20Paper_e.pdf.

Flaxman, M. 2004. Habitat identification and mapping for the Acadian flycatcher, hooded warbler and prothonotary warbler in southern Ontario. IRF Project #31. Contract report for NWRC, CWS, EC. K1A-0H3.

Foody, G.M., N.A. Campbell, N.M. Trodd, and T.F. Wood. 1992. Derivation and applications of probabilistic measures of class membership from the Maximum-Likelihood Classification. Photogrammetric Engineering and Remote Sensing 58(9): 1335-1341.

Foody, G.M. and D.P. Cox. 1994. Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. International Journal of Remote Sensing 15(3): 619-631.

Foody, G.M., R.M. Lucas, P.J. Curran, and M. Honzak. 1997. Non-linear mixture modelling without end-members using artificial neural network. International Journal of Remote Sensing 18(4): 937-953.

Frazer, G.W., R.A. Fournier, J.A. Trofymow, and R.J. Hall. 2001. A comparison of digital and film fisheye photography for analysis of forest canopy structure and gap light transmission. Agricultural and Forest Meteorology 109(4): 249-263.

Friesen, L., M. Cadman, P. Carson, K. Elliott, M. Gartshore, D. Martin, J. McCracken, J. Oliver, P. Prevett, B. Stutchbury, D. Sutherland, and A. Woodliffe. 2000. National Recovery Plan for acadian flycatcher (*Empidonax virescens)* and hooded warbler (*Wilsonia citrina*). National Recovery Plan No. 20. Recovery of Nationally Endangered Wildlife (RENEW). Ottawa, Ontario. 33pp.

Friesen, L. and M. Stabb. 2001. Conserve Ontario's Carolinian Forests: Preserve endangered songbirds. Bird Studies Canada.

Hall-Beyer, M. 2004. GLCM Texture: A tutorial. www.ucalgary.ca/~mhallbey.

Haralick, R.M., K. Shanmugan, and I. Dinstein. 1973. Textural features for image classification. IEEE Transactions on systems, man, and cybernetics. SMC-3: 610-621.

Hayes, F.E. 1995. Definitions for migrant birds: What is a neotropical migrant? The Auk 112(2): 521-523.

Holben, B.N. and Y.E. Shimabukuro. 1993. Linear mixing model applied to coarse spatial resolution data from multispectral satellite sensors. International Journal of Remote Sensing 14(11): 2231-2240.

Howlett, J.S. and B.J.M. Stutchbury. 2003. Determinants of between-season site, territory, and mate fidelity in hooded warblers (*Wilsonia citrina*). The Auk 120(2): 457-465.

Huffman, R.D., M.A. Fajvan, and P.B. Wood. 1999. Effects of residual overstory on aspen development in Minnesota. Canadian Journal of Forest Research 29: 284-289.

Ishida, M. 2004. Automatic thresholding for digital hemispherical photography. Canadian Journal of Forest Research 34(11): 2208-2216.

ITRES. 2003. 1548 Pixel CASI-3 (Compact Airborne Spectrographic Imager) makes its debut. www.itres.com

James, F.C. and C.E. McCulloch. 2002. Predicting species presence and abundance. In Predicting species occurrences: Issues of accuracy and scale. Edited by J.M. Scott, P.J. Heglund, M.L. Morrison, J.B. Haufler, M.G. Raphael, W.A. Wall, and F.B. Samson. Island Press, Washington. pp. 461-465.

Jennings, S.B., N.D. Brown, and D. Sheil. 1999. Assessing forest canopies and understorey illumination: canopy closure, canopy cover and other measures. Forestry 72(1): 59-73.

Jensen, J.R. 2005. Introductory digital image processing: A remote sensing perspective. 3rd Ed. Pearson Prentice Hall, Upper Saddle River, NJ.

Jonckheere, I., S. Fleck, K. Nackaerts, B. Muys, P. Coppin, M. Weiss, and F. Baret. 2004. Review of methods for in situ leaf area index determination Part I. Theories, sensors and hemispherical photography. Agricultural and Forest Meteorology 121(1): 19-35.

Klein Gebbinck, M.S. 1998. Decomposition of mixed pixels in remote sensing images to improve the area estimation of agricultural field. PhD Thesis. Faculty of Mathematics and Informatics, University of Nijmegen, Jijmegen, The Netherlands.

Levesque, J. 2001. Modelling forest structure and health using high-resolution airborne imagery: Investigation of spectral unmixing and spatial analysis of radiometric fractions. PhD Thesis. Department of Earth Sciences, Carleton University, Ottawa, Canada.

Levesque, J. and D.J. King. 2003. Spatial analysis of radiometric fractions from high-resolution multispectral imagery for modelling individual tree crown and forest canopy structure and health. Remote Sensing of Environment 84(4): 589-602.

LiCor Inc., 1991. LAI-2000 Plant canopy analyzer operating manual. Li-Cor Inc., Lincoln, NE, USA.

Lillesand, T.M. and R.W. Kiefer. 2000. Remote sensing and image interpretation. 4th Ed. John Wiley & Sons, Inc. New York, NY.

Linderman, M., J. Liu, J. Qi, L. An, Z. Ouyang, J. Yang, and Y. Tan. 2004. Using artificial neural networks to map the spatial distribution of understorey bamboo from remote sensing data. International Journal of Remote Sensing 25(9): 1685-1700.

Liu, W. and E.Y. Wu. 2005. Comparison of non-linear mixture models: sub-pixel classification. Remote Sensing of Environment 94(2): 145-154.

Martens, S.N., S.L. Ustin, and R.A. Rousseau. 1993. Estimation of tree canopy leaf area index by gap fraction analysis. Forest Ecology and Management 61(1): 91-108.

Mitchell, M.S., R.A. Lancia, and J.A. Gerwin. 2001. Using landscape-level data to predict the distribution of birds on a managed forest: Effects of scale. Ecological Applications 11(6): 1692-1708.

Mladenoff, D.J., T.A. Sickley, R.G. Haight, and A.P. Wydeven. 1995. A regional landscape analysis and prediction of favorable gray wolf habitat in the Northern Great Lakes region. Conservation Biology 9(2): 279-294.

Nadeau, S. 2003. The Species At Risk Act: How it works. Recovery: An Endangered Species Newsletter. Issue #25, p. 2-3. ISSN: 0847-0294.

NASA. 2004. Landsat 7 – Exaggerated rumors of its demise http://landsat.gsfc.nasa.gov/main/PDF/Landsat7_SLCoff.pdf

Nobis, M. and U. Hunziker. 2005. Automatic thresholding for hemispherical canopy-photographs based on edge detection. Agricultural and Forest Metereology 128(4): 243-250.

O^2 Planning and Design. 2003. Whooping crane potential habitat mapping project. Contract report for NWRC, CWS, EC. K1A-0H3.

Oki, K., H. Oguma, and M. Sugita. 2002. Sub-pixel classification of alder trees using multitemporal Landsat Thematic Mapper imagery. Photogrammetric Engineering and Remote Sensing 68(1): 77-82.

Olthof, I., D.J. King, and R.A. Lautenschlager. 2003. Overstory and understory leaf area index as indicators of forest response to ice storm damage. Ecological Indicators 3(1):49-64.

OMNR (Ontario Ministry of Natural Resources). 2003. St. Williams Crown Lands Technical Advisory Group: Recommendations for the long-term management of the St. Williams Crown Lands. Ontario Ministry of Natural Resources. http://www.mnr.gov.on.ca/mnr/ebr/stwilliams/recommedations.PDF.

OMNR (Ontario Ministry of Natural Resources). 2004. Natural Areas Report: Backus Woods. http://www.mnr.gov.on.ca/MNR/nhic/areas/areas_report.cfm?areaid=19.

OBBA (Ontario Breeding Bird Atlas). 2003. Ontario breeding bird atlas guide for participants. http://www.birdsontario.org/download/atlas_feb03.pdf.

OBBA (Ontario Breeding Bird Atlas). 2005. http://www.birdsontario.org/atlas/atlasmain.html.

Over, M., B. Schöttker, M. Braun, and G. Menz. 2003. Relative radiometric normalisation of multitemporal Landsat data – A comparison of different approaches. In: Proceedings of IGARSS, July 2003, Toulouse, France.

Page, A.M. and M.D. Cadman. 1994. Status report on the hooded warbler (*Wilsonia citrina*) in Canada. Committee on the Status of Endangered Wildlife in Canada (Ottawa, Ontario).

PCI Geomatics. 2004. www.pcigeomatics.com.

Peddle, D.R., S.P. Brunke., and F.G. Hall. 2001. A comparison of spectral mixture analysis and ten vegetation indices for estimating boreal forest biophysical information from airborne data. Canadian Journal of Remote Sensing 27(6): 627-635.

Peddle, D.R., F.G. Hall, and E.F. LeDrew. 1999. Spectral mixture analysis and geometric-optical reflectance modelling of boreal forest biophysical structure. Remote Sensing of Environment 67(3): 288-297.

Peddle, D.R. and R.L. Johnson. 2000. Spectral mixture analysis of airborne remote sensing imagery for improved prediction of leaf area index in mountainous terrain. Kananaskis Alberta. Canadian Journal of Remote Sensing 26(3): 176-187.

Pellikka, P.K.E., E.D. Seed, and D.J. King. 2000. Modelling deciduous forest ice storm damage using CIR aerial imagery and hemispheric photography. Canadian Journal of Remote Sensing 26(5): 394-405.

Petrone, R.M, J.S. Price, S.K. Carey, and J.M. Waddington. 2004. Statistical characterization of the spatial variability of soil moisture in a cutover peatland. Hydrological Processes 18(1): 41-52.

Pither, R. 1997. An investigation into the use of satellite data and geographical information systems for developing habitat models of rare bird species in southwestern Ontario. M.Sc. Thesis. Department of Geography, University of Western Ontario, London, Canada.

Pringle, R.M., J.K. Webb, and R. Shine. 2003. Canopy structure, microclimate, and habitat selection by a nocturnal snake, *Hoplocephalus bungaroides*. Ecology 84(10): 2668-2679.

Pu, R., B. Xu, and P. Gong. 2003. Oakwood crown closure estimation by unmixing Landsat TM data. International Journal of Remote Sensing 24(22): 4433-4445.

Rankin, W.T. and E.J. Tramer. 2002. Understory succession and the gap regeneration cycle in a *Tsuga Canadensis* forest. Canadian Journal of Forest Research 32(1): 16-23.

RENEW. 2004. Recovery of national endangered wildlife in Canada, Annual Report No. 14. Ottawa, Ontario. 36 pp. ISBN: 0-662-68338-2.

Robertson, M.P., M.H. Villet, and A.R. Palmer. 2004. A fuzzy classification technique for predicting species' distributions: Applications using invasive alien plants and indigenous insects. Diversity and Distributions 10(5): 461-474.

Robinson, W.D. and S.K. Robinson. 1999. Effects of selective logging on forest bird populations in a fragmented landscape. Conservation Biology 13(1): 58-66.

Rodewald, P.G. and K.G. Smith. 1998. Short-term effects of undestory and overstory management on breeding birds in Arkansas oak-hickory forests. Journal of Wildlife Management 62(4): 1411-1417.

Rogers, A.S. and M.S. Kearney. 2004. Reducing signature variability in unmixing coastal marsh Thematic Mapper scenes using spectral indices. International Journal of Remote Sensing 25(12): 2317-2335.

Runkle, J.R. 1992. Guidelines and sample protocol for sampling forest gaps. Gen.Tech.Rep.PNW-GTR-283. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 44 p.

SARA. 2005. Species at Risk Act Public Registry. http://www.sararegistry.gc.ca.

Seed, E.D. and D.J. King. 2003. Shadow brightness and shadow fraction relations with effective leaf area index: Importance of canopy closure and view angle in mixedwood boreal forest. Canadian Journal of Remote Sensing 29(3): 324-335.

Settle, J.J. and N.A. Drake. 1993. Linear mixing and the estimation of ground cover proportions. International Journal of Remote Sensing 14(6): 1159-1177.

Space Imaging. 2004. Ikonos Products Guide Version 1.4. http://www.spaceimaging.com/whitepapers_pdfs/IKONOS_Product_Guide.pdf

SPSS. 2004. SPSS Inc. www.spss.com

Staenz, K., T. Szeredi, and J. Schwarz. 1998. ISDAS – a system for processing/analysing hyperspectral data. Canadian Journal of Remote Sensing 24(2): 99-113.

Stockwell, D.R.B. and D.P. Peters. 1999. The GARP modelling system: Problems and solutions to automated spatial prediction. International Journal of Geographic Information Systems 13(2):143-158.

Store, R. and J. Kangas. 2001. Integrating spatial multi-criteria evaluation and expert knowledge for GIS-based habitat suitability modelling. Landscape and Urban Planning. 55(2): 79-93.

Taulman, J.F. and K.G. Smith. 2002. Habitat mapping for bird conservation in North America. Bird Conservation International 12(4): 281-309.

Trexler, J.C. and J. Travis. 1993. Nontraditional regression analyses. Ecology 74(6): 1629-1637.

Trimble Navigation Ltd. 2004. http://www.trimble.com.

Tso, B. and R. C. Olsen. 2004. Scene classification using combined spectral, textual, and contextual information, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery X, Proceedings of the SPIE, Vol. 5425, edited by Sylvia S. Shen and Paul E. Lewis,

Twele, A. and P. Barbosa. 2004. Monitoring vegetation regeneration after forest fires using satellite imagery. Proceedings of the 24th Symposium of the European Association of Remote Sensing Laboratories, Dubrovnik, Croatia, May 25-27, 2004.

Van der Meer, F. and S.M. de Jong. 2000. Improving the results of spectral unmixing of Landsat Thematic Mapper imagery by enhancing the orthogonality of endmembers. International Journal of Remote Sensing 21(15): 2781-2797.

Van Horne, B. 2002. Approaches to habitat modelling: The tensions between pattern and process and between specificity and generality. In Predicting species occurrences: Issues of accuracy and scale. Edited by J.M. Scott, P.J. Heglund, M.L. Morrison, J.B. Haufler, M.G. Raphael, W.A. Wall, and F.B. Samson. Island Press, Washington. pp. 63-72.

Warner, T.A., and M.C. Shank. 1997. Spatial autocorrelation analysis of hyperspectral imagery for feature selection. Remote Sensing of Environment 60(1): 58-70.

Whittam, B. and J. McCracken. 1999. Productivity and habitat selection of hooded warblers in southern Ontario. Bird Studies Canada, Port Rowan, Ontario.

Whittam, R.M, J.D. McCracken, C.M. Francis, and M.E. Gartshore. 2002. The effects of selective logging on nest-site selection and productivity of hooded warblers (*Wilsonia citrina*) in Canada. Canadian Journal of Zoology 80(4): 644-654.

Yang, X. and C.P. Lo. 2002. Using a time series of satellite imagery to detect land use and land cover changes in the Atlanta, Georgia metropolitan area. International Journal of Remote Sensing 23(9): 1775-1798.

APPENDIX 1: TIE POINTS AND GROUND CONTROL POINTS (GCPS) FOR REGISTERING IMAGERY.

Tie points for co-registering two Landsat 5 scenes.

	Pixel (Col)	Line (Row)	UTM X	UTM Y
1	6297.563	5062.313	487973.438	4762773.438
2	6164.938	5444.688	484707.813	4753251.563
3	8285.875	3193.125	536315.625	4808915.625
4	7808.875	4524.625	524621.875	4775778.125
5	6505.031	4005.844	493128.125	4789153.125
6	7482.594	1941.219	517074.219	4840450.781
7	8882.156	1753.281	550923.438	4844764.063
8	6393.313	6169.188	490128.125	4735046.875
9	5805.375	6158.875	475927.344	4735511.719
10	7598.313	3594.688	519654.688	4799076.563
11	7007.094	3184.031	505398.438	4809532.813
12	7605.875	5692.625	519529.688	4746632.813
13	7347.563	4929.938	513404.688	4765773.438
14	7813.906	6394.031	524451.953	4729024.609
15	8470.188	2341.063	540900.781	4830197.656
16	6776.125	1943.125	499953.125	4840628.125
17	6199.125	4091.125	485696.875	4787090.625
18	7414.875	4223.625	515126.563	4783429.688
19	8428.063	3830.563	539689.063	4792964.063

GCPs for georeferencing Landsat mosaic.

	Pixel (Col)	Line (Row)	UTM X	UTM Y
1	544024.184	4717969.508	8504.188	6915.938
2	529060.809	4743530.064	7912.750	5893.750
3	493189.054	4735891.160	6496.625	6191.875
4	468102.663	4726771.057	5511.813	6550.188
5	453835.692	4712994.800	4953.469	7096.969
6	441491.512	4726332.052	4476.969	6558.156
7	416892.594	4713407.198	3523.375	7063.375
8	403255.227	4717236.864	3000.125	6902.875
9	382812.770	4723536.769	2218.938	6638.188
10	384503.749	4749321.211	2292.531	5606.906
11	408211.228	4748247.143	3199.250	5663.750
12	428334.491	4747848.362	3974.125	5690.125
13	432017.987	4761620.566	4119.875	5141.625
14	443298.007	4776163.016	4558.250	4564.750
15	440225.098	4798336.695	4445.625	3675.375
16	451468.590	4832790.672	4888.125	2302.625
17	476997.674	4817057.506	5874.625	2940.625
18	496192.673	4805300.569	6622.531	3415.594
19	513212.106	4801848.900	7290.227	3557.695
20	558483.739	4737264.170	9082.078	6145.891
21	584681.890	4751785.631	10130.469	5563.406
22	617057.748	4753418.893	11434.375	5495.125
23	623238.112	4776767.411	11682.734	4561.328
24	603248.386	4781626.794	10874.547	4369.797
25	577858.176	4782558.034	9856.531	4334.906
26	551995.396	4774489.350	8824.750	4655.750
27	552492.916	4758309.793	8843.750	5303.250
28	465620.388	4760922.763	5422.625	5182.625
29	497535.461	4776237.109	6671.625	4578.375
30	549858.706	4792191.412	8740.125	3947.875
31	530769.420	4818676.414	7984.656	2887.031
32	554758.649	4824321.861	8935.250	2663.750
33	577854.044	4819561.262	9855.250	2855.250
34	596267.120	4823142.727	10591.875	2710.625
35	536764.153	4728635.852	8215.875	6489.375
36	541397.441	4727917.064	8400.625	6518.375
37	545927.347	4727456.559	8581.242	6537.773
38	540963.974	4721498.481	8383.250	6774.750

Tie points for mosaicking panchromatic sections of Ikonos imagery.

	Pixel (Col)	Line (Row)	UTM X	UTM Y
1	494.992	11100.023	543988.919	4719435.211
2	729.969	8709.031	544223.981	4721827.149
3	494.984	5273.016	543988.919	4725265.211
4	245.969	7863.031	543739.981	4722674.211
5	817.969	10080.969	544311.919	4720454.149
6	454.047	6456.984	543947.919	4724081.149
7	349.016	4554.984	543842.981	4725984.274
8	49.961	2674.977	543544.966	4727865.196
9	162.938	2530.938	543657.981	4728009.149
10	212.094	180.969	543706.981	4730360.149
11	532.031	3614.969	544025.981	4726924.149
12	35.063	442.938	543529.919	4730098.211
13	607.016	9549.953	544101.044	4720986.211
14	640.938	12297.063	544134.919	4718237.274
15	1199.008	1271.992	544692.919	4729268.211

Tie points for mosaicking multispectral sections of Ikonos imagery.

	Pixel (Col)	Line (Row)	UTM X	UTM Y
1	71.969	1569.969	543787.075	4724255.305
2	254.016	217.016	544511.325	4729671.305
3	157.906	2112.969	544127.075	4722083.305
4	120.969	981.031	543982.825	4726611.305
5	95.984	3152.953	543879.075	4717919.305
6	61.969	1966.906	543742.825	4722667.305
7	59.953	2281.016	543739.075	4721407.305
8	177.016	2587.984	544206.825	4720183.305
9	171.969	688.031	544186.700	4727786.930
10	136.984	1306.016	544046.825	4725311.055
11	121.953	2775.984	543983.075	4719427.055
12	236.969	165.969	544446.575	4729875.055
13	251.094	470.031	544503.013	4728659.243
14	206.992	2528.961	544326.888	4720419.118
15	146.016	1778.984	544083.013	4723419.368

GCPs for georeferencing Ikonos panchromatic image.

	Pixel (Col)	Line (Row)	UTM X	UTM Y
1	923.063	12712.813	534628.616	4717822.765
2	2631.688	8499.438	536338.281	4722038.609
3	1207.344	6162.219	534913.348	4724375.720
4	3056.641	1903.047	536764.137	4728635.834
5	4458.344	4319.031	538166.453	4726218.247
6	5832.125	6678.625	539539.941	4723858.218
7	7256.875	9037.875	540963.995	4721498.483
8	8612.328	11415.297	542320.490	4719118.818
9	10319.938	12569.063	544027.726	4717965.122
10	11667.938	11711.938	545376.956	4718820.122
11	11073.698	8344.498	544782.145	4722189.511
12	9713.719	5980.719	543422.222	4724554.844
13	9056.313	4863.188	542765.214	4725674.486
14	8323.541	3674.375	542031.742	4726862.765
15	7688.813	2620.438	541397.469	4727917.053
16	12216.875	3080.375	545927.324	4727456.582
17	14231.813	6500.813	547943.640	4724035.408

Tie points for georeferencing multispectral Ikonos image to panchromatic image.

	Pixel (Col)	Line (Row)	UTM X	UTM Y
1	58.844	1549.156	533940.075	4724343.805
2	657.938	2125.063	536338.263	4722037.368
3	501.188	1138.938	535708.419	4725983.711
4	764.875	475.375	536764.153	4728635.852
5	1115.063	1079.688	538166.435	4726218.251
6	1458.438	1670.063	539539.961	4723858.214
7	1814.563	2259.688	540963.973	4721498.481
8	2153.688	2853.938	542320.468	4719118.796
9	2579.813	3141.563	544027.719	4717965.111
10	2917.813	2926.938	545376.958	4718820.134
11	2769.281	2085.406	544782.171	4722189.506
12	814.063	2546.813	536963.388	4720349.368
13	2260.156	1213.531	542750.325	4725681.055
14	2080.469	915.656	542029.575	4726875.055
15	1922.969	654.906	541397.441	4727917.064
16	3055.188	769.188	545927.347	4727456.559
17	3558.938	1622.688	547943.653	4724035.436
18	174.219	403.906	534404.638	4728925.118
19	1140.219	163.719	538265.825	4729887.055
20	3070.031	309.969	545985.013	4729290.868
21	3343.938	390.063	547085.888	4728967.993
22	3200.031	1020.469	546509.638	4726446.243
23	2863.719	1605.219	545160.044	4724111.024
24	2411.688	60.438	543356.294	4730294.899
25	35.313	2890.188	533842.200	4718977.430
26	754.688	2990.563	536721.075	4718572.305
27	1617.063	2719.563	540172.575	4719657.555
28	1361.313	606.438	539149.075	4728113.055
29	884.875	1601.625	537244.200	4724132.430
30	2465.938	1985.938	543568.763	4722591.743
31	1910.063	1370.938	541346.325	4725054.055
32	2049.656	115.969	541905.669	4730075.711
33	2554.875	645.375	543925.075	4727954.555

APPENDIX 2: TEXTURES CALCULATED FROM THE GREY LEVEL CO-OCCURRENCE MATRIX (Haralick et al., 1973; Hall-Beyer, 2004; Jensen, 2005).

In these equations $P_{i,j}$ is the probability of occurrence of grey level *i* next to grey level *j*, where *n* is the number of possible grey levels in the image. For calculating the Correlation texture, μ is the mean GLCM value for the i^{th} or j^{th} grey level, while σ is the variance.

$$HOMOGENEITY = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}\frac{P_{i,j}}{1+(i-j)^2}$$

$$CONTRAST = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}P_{i,j}(i-j)^2$$

$$DISSIMILARITY = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}P_{i,j}|i-j|$$

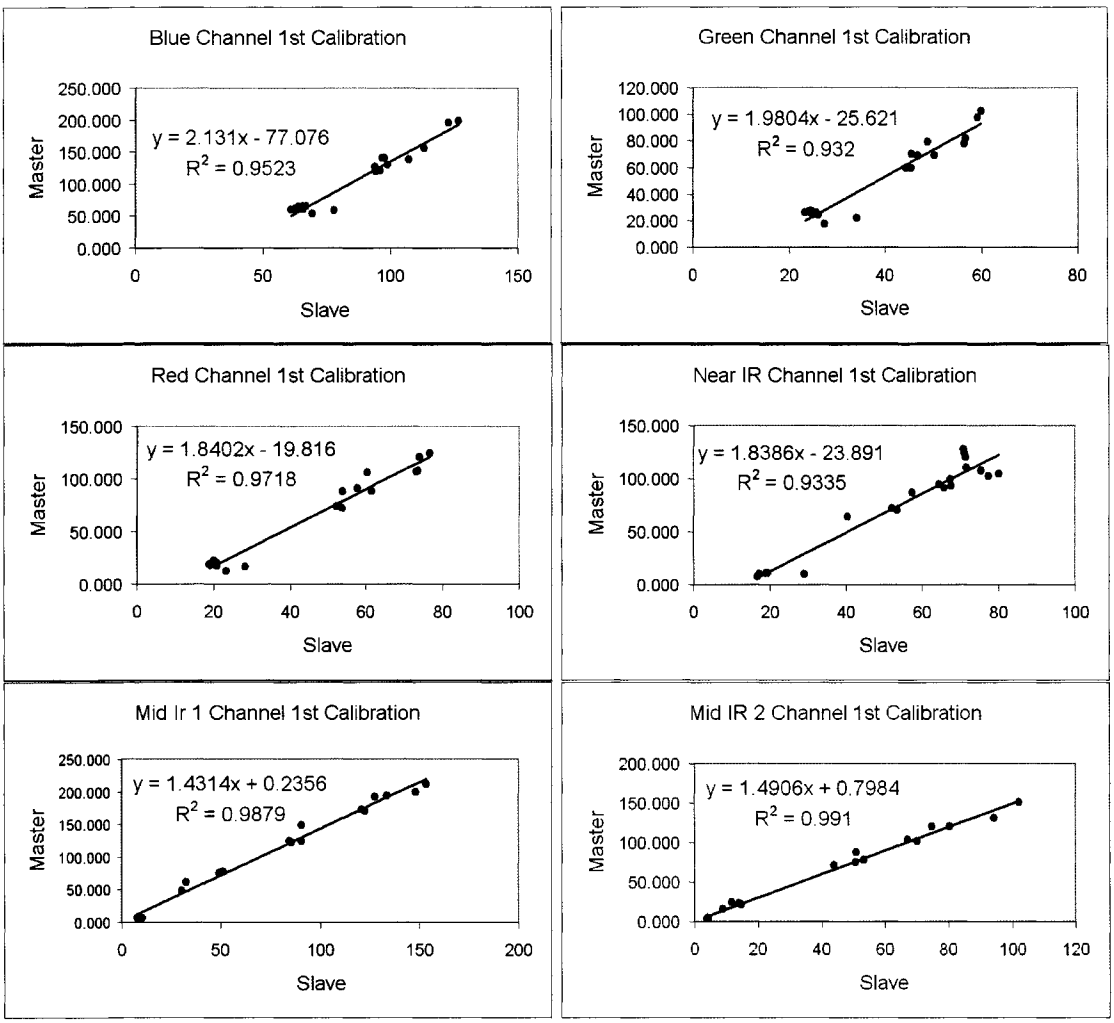
$$ENTROPY = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}P_{i,j}(-\ln P_{i,j})$$

$$ANGULAR\ SECOND\ MOMENT = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}(P_{i,j})^2$$

$$CORRELATION = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}P_{i,j}\left[\frac{(i-\mu_i)(i-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}}\right]$$

APPENDIX 3: CALIBRATION TRANSFORMATION APPLIED TO SPECTRAL BANDS (DN VALUES) OF THE SLAVE IMAGE

First Calibration



Second Calibration