

Spring 2001

Using satellite remote sensing to model and map the distribution of Bicknell's thrush (*Catharus bicknelli*) in the White Mountains of New Hampshire

Stephen Roy Hale

University of New Hampshire, Durham

Follow this and additional works at: <https://scholars.unh.edu/dissertation>

Recommended Citation

Hale, Stephen Roy, "Using satellite remote sensing to model and map the distribution of Bicknell's thrush (*Catharus bicknelli*) in the White Mountains of New Hampshire" (2001). *Doctoral Dissertations*. 17.
<https://scholars.unh.edu/dissertation/17>

This Dissertation is brought to you for free and open access by the Student Scholarship at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact nicole.hentz@unh.edu.

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI[®]

USING SATELLITE REMOTE SENSING TO MODEL AND MAP
THE DISTRIBUTION OF BICKNELL'S THRUSH (*Catharus bicknelli*)
IN THE WHITE MOUNTAINS OF NEW HAMPSHIRE

BY

STEPHEN R. HALE

Bachelor of Science, University of Miami, 1989
Master of Science, Louisiana State University, 1994

DISSERTATION

Submitted to the University of New Hampshire
in Partial Fulfillment of
the Requirements for the Degree of

Doctor of Philosophy

In

Natural Resources

May, 2001

UMI Number: 3006138

UMI[®]

UMI Microform 3006138

Copyright 2001 by Bell & Howell Information and Learning Company.

*All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.*

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

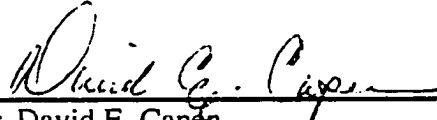
This dissertation has been examined and approved.



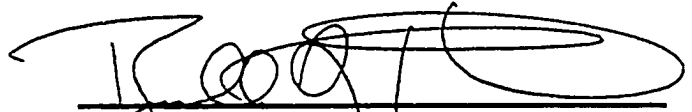
Dissertation Director, Dr. Barrett N. Rock,
Professor of Natural Resources and Earth,
Oceans and Space
University of New Hampshire



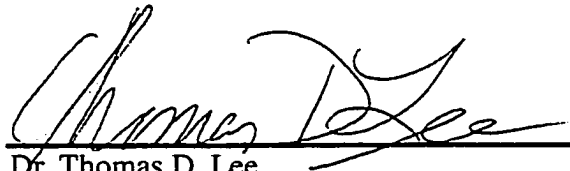
Dr. Richard W. Birnie,
Professor of Earth Sciences
Dartmouth College



Dr. David E. Capen,
Professor of Natural Resources
University of Vermont



Dr. Russell G. Congalton,
Professor of Remote Sensing & Geographic
Information Systems
University of New Hampshire



Dr. Thomas D. Lee,
Associate Professor of Plant Biology
(Ecology)
University of New Hampshire

January 5, 2001

DEDICATION

I dedicate this dissertation to those who have most partnered in its completion,

my wife and son.

ACKNOWLEDGEMENTS

Special thanks are extended to my graduate advisory committee Drs. Barrett N. Rock, Richard W. Birnie, David E. Capen, Russell G. Congalton, and Thomas D. Lee for their advise, guidance, and commitment, which fostered my own development, growth, and understanding of the scientific process. I especially thank my major advisor Dr. Barrett N. Rock, without whose initial enthusiasm and encouragement, this program of study would not have happened.

I offer special thanks to Dr. Mary E. Martin, for her invaluable hours of patience and assistance, which greatly improved this study. There were many individuals who provided invaluable assistance for field data collection under extremely adverse conditions to whom I extend thanks. Among these I especially thank Jay Raymond without whose energy, ability, and camaraderie this project would have fallen woefully short of its goals. I thank fellow students, Petya Entcheva, Shannon Spencer, and Ryan Huntley with whom I have worked closely.

Further thanks are extended to the White Mountain National Forest, New Hampshire Fish and Game Department, and the Audubon Society of New Hampshire for collecting and making available the High Elevation Bird Monitoring Program data. Additionally, I thank Mariko Yamasaki, Leslie Rowse, Chris Costello, and Mike Madeiros for logistical support and housing during the field season. Also, I thank Carl Demrow of the Appalachian Mountain Club for providing use of remote campsites.

Financial support making this study possible included funds from New Hampshire Space Grant, ForestWatch (NSF #ESI-9452792), the New England Regional Assessment (NSF #OCE-9903604), and The Garden Club of America's Francis M. Peacock Scholarship for Native Bird Habitat.

Finally, I extend my most gracious thanks to my wife, Eleanor, and son, Jackson, whose unwavering love, commitment, support, and companionship permitted me to fulfill a dream.

TABLE OF CONTENTS

DEDICATION.....	iii
ACKNOWLEDGEMENTS.....	iv
LIST OF TABLES.....	vii
LIST OF FIGURES.....	ix
ABSTRACT.....	xi
CHAPTER 1 INTRODUCTION.....	1
CHAPTER 2 COMPARISON OF TOPOGRAPHIC NORMALIZATION STRATEGIES IN LAND COVER CLASSIFICATION OF A MONTANE ENVIRONMENT IN NORTHERN NEW ENGLAND, USA.....	12
CHAPTER 3 IMPROVING LAND COVER CLASSIFICATION WITH ANCILLARY DATA.....	46
CHAPTER 4 MODELING BICKNELL'S THRUSH DISTRIBUTION USING ETM SATELLITE IMAGERY.....	57
CHAPTER 5 ESTIMATING BICKNELL'S THRUSH (<i>Catharus bicknelli</i>) DISTRIBUTION, DENSITY, AND ABUNDANCE FROM A HABITAT MODEL DERIVED FROM SATELLITE IMAGERY.....	90
CHAPTER 6 SUMMARY OF FINDINGS.....	106

LIST OF TABLES

2.1	Topographic normalization image treatments.....	35
2.2	Land cover class definitions and descriptions.....	38
2.3	Results of null hypothesis test that image treatment classification accuracies were not different from zero.....	39
2.4	Results of null hypothesis test that image treatment classification accuracies were not different from each other.....	39
2.5	Ranges of Producer's and User's Accuracies for each land cover class and all four image treatments.....	45
2.6	Image treatments yielding the highest class specific Producer's and User's Accuracies for each land cover class.....	45
4.1	Logistic modeling regressor variables.....	83
4.2	Results of univariate logistic regression.....	84
4.3	Double-odds scale for significant variables from univariate logistic regression analysis.....	84
4.4	Step-wise logistic multiple regression producing the best fit Bicknell's Thrush habitat model.....	85
4.5	Analysis of Variance for multiple regression of dominant vegetation height versus Enhanced Thematic Mapper Normalize Difference Vegetation Index and Band 1.....	86
4.6	Logistic multiple regression producing the applied Bicknell's Thrush habitat model.....	86
4.7	Estimated versus observed probabilities for the calibration data set using the applied Bicknell's Thrush habitat model.....	87

4.8	Observed versus predicted numbers of Bicknell's Thrush records at model validation stations.....	88
5.1	Estimates of area occupied by modeled probability deciles for Bicknell's Thrush presence.....	106
5.2	Estimates of Bicknell's Thrush relative abundance and density for modeled probability deciles.....	106

LIST OF FIGURES

1.1	The White Mountain National Forest of New Hampshire, USA.....	10
1.2	Diagram of data and analysis flow for the program of study.....	11
2.1	Aspect partitioning.....	36
2.2	Photograph from within a fir wave on Mt. Cabot (Kilkenny, NH).....	37
2.3	Scatterplots and regression equations for determination of Minnaert Constants.....	40
2.4	Discrete and fuzzy error matrices for image treatment I1.....	41
2.5	Discrete and fuzzy error matrices for image treatment I2.....	42
2.6	Discrete and fuzzy error matrices for image treatment I3.....	43
2.7	Discrete and fuzzy error matrices for image treatment I4.....	44
2.8	Land cover map from image treatment I4.....	46
3.1	Comparison of spectral signatures for the Softwoods, Krummholz, and Kampfzone land cover classes.....	50
3.2	Discrete and fuzzy error matrices for adjusted classification I4 after combining Kampfzone and High Mortality into the single Fir Sapling land cover class.....	56
3.3	Land cover map of adjusted classification I4.....	57
4.1	Scatterplot of elevation estimates from a digital elevation model versus estimates from differential GPS.....	85
4.2	Logit and probability solution for the applied Bicknell's Thrush habitat model.....	87
4.3	Scatterplot of the estimated probability versus observed proportion of Bicknell's Thrush at model calibration stations.....	88

4.4 Predicted versus observed numbers of Bicknell's Thrush presence records at model validation stations.....	89
4.5 Predictive maps of Bicknell's Thrush habitat distribution at three spatial scales.....	90

ABSTRACT

USING SATELLITE REMOTE SENSING TO MODEL AND MAP THE DISTRIBUTION OF BICKNELL'S THRUSH (*Catharus bicknelli*) IN THE WHITE MOUNTAINS OF NEW HAMPSHIRE

By

Stephen R. Hale

University of New Hampshire, May, 2001

Landsat-7 Enhanced Thematic Mapper satellite imagery was used to model Bicknell's Thrush (*Catharus bicknelli*) distribution in the White Mountains of New Hampshire. The proof-of-concept was established for using satellite imagery in species-habitat modeling, where for the first time imagery spectral features were used to estimate a species-habitat model variable. The model predicted rising probabilities of thrush presence with decreasing dominant vegetation height, increasing elevation, and decreasing distance to nearest Fir Sapling cover type. To solve the model at all locations required regressor estimates at every pixel, which were not available for the dominant vegetation height and elevation variables. Topographically normalized imagery features Normalized Difference Vegetation Index and Band 1 (blue) were used to estimate dominant vegetation height using multiple linear regression; and a Digital Elevation Model was used to estimate elevation. Distance to nearest Fir Sapling cover type was obtained for each pixel from a land cover map specifically constructed for this project.

The Bicknell's Thrush habitat model was derived using logistic regression, which produced the probability of detecting a singing male based on the pattern of model

covariates. Model validation using Bicknell's Thrush data not used in model calibration, revealed that the model accurately estimated thrush presence at probabilities ranging from 0 to < 0.40 and from 0.50 to < 0.60 . Probabilities from 0.40 to < 0.50 and greater than 0.60 significantly underestimated and overestimated presence, respectively.

Applying the model to the study area illuminated an important implication for Bicknell's Thrush conservation. The model predicted increasing numbers of presences and increasing relative density with rising elevation, with which exists a concomitant decrease in land area. Greater land area of lower density habitats may account for more total individuals and reproductive output than higher density less abundant land area. Efforts to conserve areas of highest individual density under the assumption that density reflects habitat quality could target the smallest fraction of the total population.

CHAPTER 1

INTRODUCTION

Ornithology has a long tradition of quantitative species-habitat modeling. In the late 1960's, several researchers noted the compatibility of hypergeometric niche theory (Hutchinson, 1944) with multivariate statistics, and began to model bird habitat in quantitative terms (Shugart, 1981). Discriminant Function Analysis was first used to model habitats of grassland birds (Cody, 1968) and later for forest birds (Hespenheide, 1971; James, 1971). Principal Components Analysis was also used to model the habitat of forest birds (James, 1971; Anderson and Shugart, 1974). In the early 1980's, species-specific multivariate habitat models became focused toward guiding wise land use activities for species conservation (Schamberger and O'Neil, 1986). Most recently, species habitat modeling and distribution mapping has employed advances in geospatial technologies, including the Global Positioning System (GPS), geographic information systems (GIS), and a variety of remote sensing methods including aerial photography, airborne and satellite spectral sensors, and synthetic aperture radar (Green et al., 1987; Schwaller et al., 1989; Palmeirim, 1988; Breininger et al., 1991; Franklin and Steadman, 1991; Aspinall and Veitch, 1993; Herr and Queen, 1993; Breininger et al., 1994; Green and Griffiths, 1994; Jørgensen and Nøhr, 1996; Hepinstall and Sader, 1997; Imhoff et al., 1997; Dettmers and Bart, 1999).

Mosher et al. (1986) suggested the following objectives for species habitat models: 1) greater than 80% accuracy of prediction; 2) be affordable; 3) include variables that are readily measured; 4) include variables that have typically been used in the past; 5) include variables that can be demonstrated as important; and 6) be applicable of most or all of a species' range. These authors noted further that these objectives were difficult to obtain singly, and increasingly difficult to satisfy multiple objectives with a single model. Contributions to habitat modeling from geospatial technologies offer the greatest potential to maximize the likelihood of satisfying multiple model objectives.

Satellite remote sensing imagery provides extremely cost effective information over vast geographic extents at fine spatial scales (on the order of 10's of meters). Thus, objectives 2 and 6 above can be maximized with satellite remote sensing. Objectives 3 and 4 can be satisfied because satellite imagery have been used by remote sensing analysts to estimate the same types of forest stand parameters used by ornithologists in habitat modeling, including basal area, stem density, canopy closure, species composition, vegetation height, seral stage, forest damage, biomass and leaf-area-index (Vogelmann and Rock, 1988; Sader et al., 1990; Spanner et al., 1990; Cohen and Spies, 1992; Franklin, 1994; Hall et al., 1995; Jakubauskas, 1996; Jakubauskas and Price, 1997).

The remaining objectives, 1 and 5, require a detailed program of study to determine if satellite remote sensing can accurately predict species-habitat using variables demonstrated as important. Of published bird habitat modeling studies that use satellite imagery, none has utilized its potential for estimating model habitat variables. Instead, most used Landsat-5 Thematic Mapper (TM) imagery to generate thematic

classifications or maps of vegetative land cover as input into a geographical information system (GIS) (Green et al., 1987; Palmeirim, 1988; Franklin and Steadman, 1991; Herr and Queen, 1993). Analyses were performed with accompanying data layers, to predict localities of species occurrence. Other studies have applied Bayes' Theorem to allow pixel Brightness Values (BVs) to be probabilistically recoded based on pixel BVs for sites with the species present versus those with the species absent (Aspinall and Veitch, 1993; Hepinstall and Sader, 1997). The products of these Bayesian type methods are best described as species distribution maps rather than habitat maps, because no habitat model was used and no formal link between spectral and habitat variables was made.

The hypothesis tested was that there would be no difference between estimates of the number of stations surveyed for the presence of Bicknell's Thrush (BITH; *Catharus bicknelli*) and the number of BITH presence stations as predicted from a habitat model applied using satellite imagery. Moreover, a primary objective was to determine if a species-habitat model could be applied using satellite imagery and satisfy the six objectives of Mosher et al. (1986). Bicknell's Thrush, which is endemic to the northeastern US and southeastern Canada, and has been identified as a conservation priority (Rosenberg and Wells, 1995). Bicknell's Thrush is Neotropical migrant, and habitat specialist during its breeding season that utilizes northeastern US montane coniferous forests above approximately 900 m (Wallace, 1939; Noon, 1981; Atwood et al., 1996; Rimmer et al., 1996). Because BITH is a habitat specialist, it was selected for study to determine if satellite imagery could be used to effectively identify areas of use. The study was conducted on the White Mountain National Forest (WMNF) of New Hampshire (Fig. 1.1A).

The WMNF, consists of approximately 300,000 hectares and lies within the northern portion of New Hampshire (Coos, Grafton, and Carroll Counties), and has a small presence in eastern Maine (there was no sampling in Maine). All sampling localities lie within mixed coniferous-deciduous forest, spruce-fir coniferous forest, and spruce-fir krummholz habitats within an elevation range of 600-1500 m above-sea-level (Fig. 1.1B). There are approximately 110,000 acres of these cover types at this elevation range in the WMNF (Miller-Weeks and Smoronk, 1994). Nearly one-fourth of the sampling points fall within or adjacent to Wilderness Area boundaries.

Dominant tree species within the study area are balsam fir, red spruce, yellow birch (*Betula allegheniensis*), mountain paper birch (*Betula papyrifera* var. *cordifolia*), and showy mountain-ash (*Sorbus decora*). Within disturbed sites, such as fir-waves, pin-cherry (*Prunus pensylvanica*) may represent a significant fraction of deciduous woody cover. Herbaceous understory vegetation varies with each site, but the most common taxa found include clintonia (*Clintonia borealis*), goldthread (*Coptis trifolia*), wood-sorrel (*Oxalis* sp.), and bunchberry (*Cornus canadensis*). Also conspicuous are various other taxa including ferns, mosses, grasses, and lichen (both ground and epiphytic forms).

The overall research study approach and flow consisted of a number of steps including data collection, land cover mapping, habitat modeling, variable estimation, and model validation (Fig. 1.2). An existing data set of BITH point count data from 1993-1997 in the White Mountains was used to identify survey stations as BITH presence or absence. The same point count stations were visited during the summer of 1998 for vegetation sampling and to acquire GPS location and elevation estimates. A

land cover map of the forest was constructed using 1999 Landsat-7 Enhanced Thematic Mapper (ETM) satellite imagery.

Areas characterized by strong topographic relief pose challenges to land cover mapping because of differences in incident radiation and non-Lambertian reflectance behavior (Teillet et al., 1982; Colby, 1991; Ekstrand, 1996). Chapter 2 of this dissertation compares the resulting map accuracies from applying various topographic normalization treatments. The treatment that delivered the most accurate map (using several measures of map accuracy) was selected for use in deriving landscape level variables for habitat analysis. Chapter 3 describes adjustments made to the spectra-based land cover classification from ancillary data that were believed to improve classification accuracy.

Chapter 4 details the habitat modeling, model application, and model validation procedures and results. Using the BITH binary data as the outcome variable and regressor variables derived from field visits and the land cover map, a logistic multiple regression habitat model was constructed. Three variables: 1) dominant vegetation height, 2) elevation, and 3) distance to nearest Fir Sapling cover type were found to be effective predictors of BITH presence/absence. Because data collection only occurred at point count stations, the variables had to be estimated from other sources. A digital elevation model was used to estimate elevation, spectral features from the ETM imagery were used to estimate dominant vegetation height, and the land cover map was analyzed in a GIS to find the distance to nearest Fir Sapling cover type at every pixel. The model was solved at every pixel using the variable estimates and validated with BITH

presence/ absence data that were not used in model calibration. The applied habitat model validated well at predicting the probability of BITH presence.

Finally, Chapter 5 examines the output of the habitat model in terms of the predicted distribution, relative abundance, and relative density of BITH in the White Mountains. A relationship was observed where the relative density of BITH rises with increasing elevation. With increasing elevation, however, there was a concomitant decrease in available land area. Thus, habitats that support the greatest densities of BITH were also most limited in terms of their abundance. In the absence of reproductive fitness data relating to habitat quality, it cannot be assumed that habitats with the greatest densities of individuals represent the most significant source of BITH added to the population with each breeding season. This dissertation effectively demonstrated the proof-of-concept for using satellite imagery in detailed species-habitat modeling studies.

Within the following chapters, there is mention of several specific brand name software packages and field equipment used in the study. Inclusion of specific manufacturer's products was for the purpose of methodological reconstruction as these may contain unique proprietary components that effect data results. Mention of specific products or manufacturers was not intended as an endorsement.

Literature Cited

Anderson, S. H. and H. H. Shugart Jr. 1974. Habitat selection of breeding birds in an east Tennessee deciduous forest. *Ecology* 55:828-837.

Aspinall, R. and N. Veitch. 1993. Habitat mapping from satellite imagery and wildlife survey data using a Bayesian modeling procedure in a GIS. *Photogrammetric Engineering & Remote Sensing* 59(4): 537-543.

- Atwood, J. L., C. C. Rimmer, K. P. McFarland, S. H. Tsai, and L. R. Nagy. 1996. Distribution of Bicknell's Thrush in New England and New York. *Wilson Bull.* 108(4): 650-661.
- Breining, D. R., M. J. Provancha, and R. B. Smith. 1991. Mapping Florida Scrub Jay habitat for purposes of land-use management. *Photogrammetric Engineering and Remote Sensing* 51:1467-1474.
- Breining, D. R., V. L. Larson, B. W. Duncan, R. B. Smith, D. M. Oddy, and M. F. Goodchild. 1994. Landscape patterns of Florida Scrub Jay habitat use and demographic success. *Conservation Biology* 9(6):1442-1453.
- Cody, M. L. 1968. On methods of resource division in grassland bird communities. *Am. Nat.* 102:107-147.
- Cohen, W. B. and T. A. Spies. 1992. Estimating structural attributes of Douglas Fir/western hemlock forest stands from Landsat and SPOT imagery. *Remote Sensing of Environment* 41: 1-17.
- Colby, J. D. 1991. Topographic normalization in rugged terrain. *Photogrammetric Engineering & Remote Sensing* 57(5): 531-537.
- Dettmers, R. and J. Bart. 1999. A GIS modeling method applied to predicting forest songbird habitat. *Ecological Applications* 9(1): 152-163.
- Ekstrand, S. 1994. Assessment of forest damage with Landsat TM: Correction for varying forest compartment characteristic's assessment of damage. *Remote Sens. of Environ.* 47:291-302.
- Franklin, J. and D. W. Steadman. 1991. The potential for conservation of Polynesian birds through habitat mapping and species translocation. *Conservation Biology* 5:506-521.
- Franklin, S. E. 1994. Discrimination of subalpine forest species and canopy density using digital CASI, SPOT PLA, and Landsat TM data. *Photogrammetric Engineering & Remote Sensing* 60(10): 1233-1241.
- Green, K. M., J. F. Lynch, J. Sircar, and L. S. Z. Greenberg. 1987. Landsat remote sensing to assess habitat for migratory birds in the Yucatan Peninsula, Mexico. *Vida Silvestre Neotropical* 1(2): 457-471.
- Green, R. E. and G. H. Griffiths. 1994. Use of preferred nesting habitat by Stone Curlews *Burhinus oedicephalus* in relation to vegetation structure. *J. Zool., Lond.* 233:457-471.

- Hall, F. G., Y. E. Shimabukuro, and K. F. Huemmrich. 1995. Remote sensing of forest biophysical structure using mixture decomposition and geometric reflectance models. *Ecological Applications* 5(4): 993-1013.
- Hepinstall, J.A., & S.A. Sader. 1997. Using Bayesian statistics, Thematic Mapper satellite imagery, and Breeding Bird Survey data to model bird species probability of occurrence in Maine. *Photogrammetric Engineering and Remote Sensing* 63(10): 1231-1237.
- Herr, A. M. and L.P. Queen. 1993. Crane Habitat Evaluation Using GIS and Remote Sensing. *Photogrammetric Engineering & Remote Sensing* 59(10): 1531-1538.
- Hespenheide, H. A. 1971. Flycatcher habitat selection in the eastern deciduous forest. *Auk* 88:61-74.
- Hutchinson, G. E. 1944. Limnological studies in Connecticut. VII. A critical examination of the supposed relationship between phytoplankton periodicity and chemical changes in lake waters. *Ecology* 25:3-26.
- Imhoff, M. L., T. D. Sisk, A. Milne, G. Morgan, and T. Orr. 1997. Remotely sensed indicators of habitat heterogeneity: Use of Synthetic Aperture Radar in mapping vegetation structure and bird habitat. *Remote Sens. Environ.* 60:217-227.
- Jakubauskas, M. E. 1996. Canonical correlation analysis of coniferous forest spectral and biotic relations. *Int. J. Remote Sensing* 17(12): 2323-2332.
- Jakubauskas, M. E. and K. P. Price. 1997. Empirical relationships between structural and spectral factors of Yellowstone Lodgepole pine forests. *Photogrammetric Engineering & Remote Sensing* 63(12): 1375-1381.
- James, F. C. 1971. Ordinations of habitat relationships among breeding birds. *Wilson Bulletin* 83:215-236.
- Jørgensen, A. F. and H. Nøhr. 1996. The use of satellite images for mapping of landscape and biological diversity in the Sahel. *Int. J. Remote Sensing* 17:91-109.
- Miller-Weeks, M. and D. Smoronk. 1994. Aerial assessment of red spruce and balsam fir conditions in the Adirondack region of New York, the Green Mountains of Vermont, the White Mountains of New Hampshire and the mountains of western Maine. USDA, Forest Service Northeastern Area, NA-TP-16-93.
- Mosher, J. A., K. Titus, and M. R. Fuller. 1986. Developing a practical model to predict nesting habitat of woodland hawks. Pp. 31-35. in *Wildlife 2000: Modeling Habitat Relationships of Terrestrial Vertebrates* (Verner, J., M. L. Morrison, and C. J. Ralph, eds.). University of Wisconsin Press. Madison, WI.

- Noon, B. R. 1981. The distribution of an avian guild along a temperate elevation gradient: The importance and expression of competition. *Ecological Monographs* 51(1): 105-124.
- Palmeirim, J. M. 1988. Automatic mapping of avian species habitat using satellite imagery. *Oikos* 52:59-68.
- Rimmer, C. C., J. L. Atwood, K. P. McFarland, and L. R. Nagy. 1996. Population density, vocal behavior, and recommended survey methods for Bicknell's Thrush. *Wilson Bulletin* 108(4): 639-649.
- Rosenberg, K. V. and J. V. Wells. 1995. Importance of Geographic Areas to Neotropical Migrant Birds in the Northeast. Report submitted to US Fish and Wildlife Service, Region-5, Hadley, MA.
- Sader, S. A., R. B. Waide, W. T. Lawrence, and A. T. Joyce. 1990. Tropical forest biomass and successional age class relationships to a vegetation index derived from Landsat TM data. *Remote Sensing of Environment* 28: 143-156.
- Schamberger, M. L. and L. J. O'Neil. 1986. Concepts and constraints of habit-model testing. Pp. 5-10 in *Wildlife 2000: Modeling Habitat Relationships of Terrestrial Vertebrates* (Verner, J., M. L. Morrison, and C. J. Ralph, eds.). University of Wisconsin Press. Madison, WI.
- Schwaller, M. R., C. E. Olson Jr., Z. Ma, Z. Zhu, and P. Dahmer. 1989. Remote sensing analysis of Adelie Penguin rookeries. *Remote Sensing of Environment* 28:199-206.
- Shugart, H. H. Jr. 1981. An overview of multivariate methods and their application to studies of wildlife habitat. Pp.4-10 in *The use of multivariate statistics in studies of wildlife habitat*. (Capen, D. E., ed.). USDA Forest Service General Technical Report RM-87.
- Spanner, M. A., L. L. Pierce, D. L. Peterson, and S. W. Running. 1990. Remote sensing of temperate coniferous leaf area index: The influence of canopy closure, understory vegetation, and background reflectance. *International Journal of Remote Sensing* 11(1): 95-111.
- Teillet, P. M., B. Guindon, and D. G. Goodenough. 1982. On the slope-aspect correction of Multispectral Scanner Data. *Canadian Journal of Remote Sensing* 8(2): 84-106.
- Vogelmann, J. E. and B. N. Rock. 1988. Assessing forest damage in high-elevation coniferous forests in Vermont and New Hampshire using Thematic Mapper data. *Remote Sensing of Environment* 24: 227-246.

Wallace, G. J. 1939. Bicknell's Thrush, its taxonomy, distribution, and life history. Proc. Boston Soc. Nat. Hist. 41(6): 211-402.

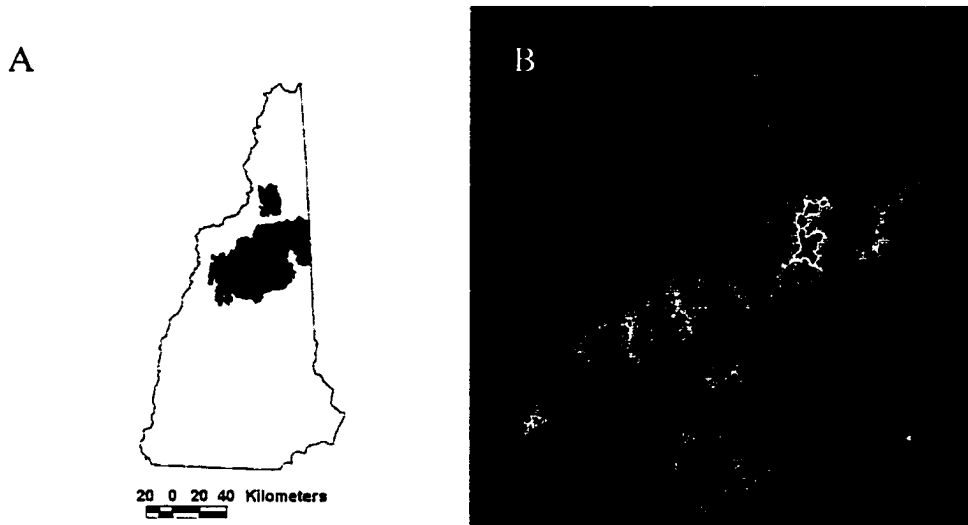


Figure 1.1. The White Mountain National Forest A) within the state boundaries of New Hampshire B) study area consisting of elevations from 615-1525 meters. Black areas were not included in the study, darker gray tones represent lower elevations and lighter tones represent higher elevations.

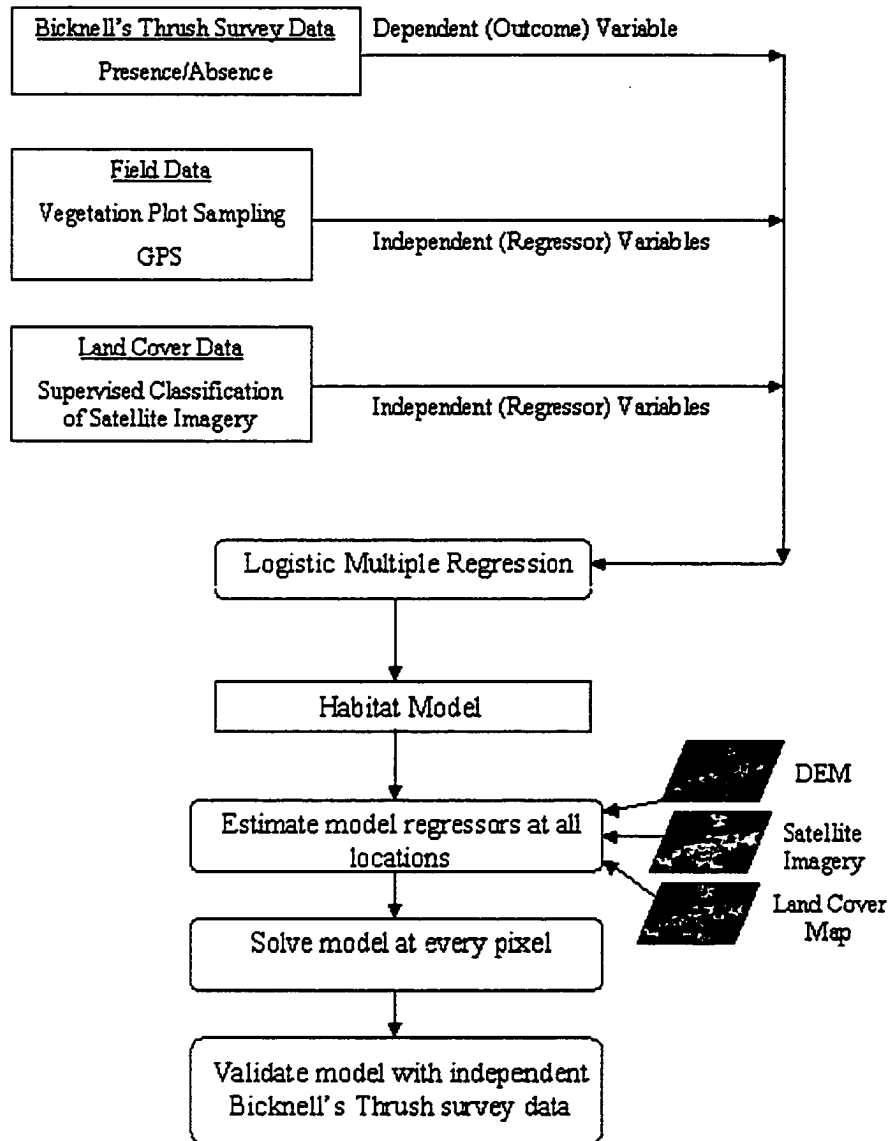


Figure 1.2. Diagram of research data and analysis flow for the construction, application, and validation of the Bicknell's Thrush habitat model.

CHAPTER 2

COMPARISON OF TOPOGRAPHIC NORMALIZATION STRATEGIES IN LAND COVER CLASSIFICATION OF A MONTANE ENVIRONMENT IN NORTHERN NEW ENGLAND, USA

Introduction

Remote sensing of mountainous areas, characterized by high topographic relief, presents challenges not encountered on level or constant terrain (Teillet et al., 1982; Colby, 1991; Ekstrand, 1996). Problems arise when spatial variations in ground target illumination and radiance, owing to differences in incident radiation and non-Lambertian reflectance behavior, cause identical surface features (e.g., land cover types) to reflect differently, or cause different surface features to reflect similarly. For pixel classifying algorithms to perform effectively, effects due to topographic relief must be minimized or removed. Several topographic normalization techniques are available, but few studies have compared their effectiveness for land cover classification. The purpose of this part of the study was to investigate the efficacy of various topographic normalization strategies toward construction of a land cover classification of montane forest in northern New England, USA.

Normalization strategies investigated were 1) band-ratios, 2) Minnaert Correction, 3) aspect partitioning, and 4) combinations of these strategies. Aspect partitioning is the geographical stratification of an image (Jensen, 1996), where the image study area is divided into subsets with pixel aspect values of one image subset

("sunlit" image) within or equal to 180 degrees of the solar azimuth, and another image subset ("sunshade" image) greater than 180 degrees from the solar azimuth. The two image subsets are classified separately using only training sites from within the respective subset, and resulting classifications subsequently merged back into the full study area. Aspect partitioning is not, *sensu stricto* a normalizing method because there is no spectral adjustment of pixel values. However, its inclusion here is based on its potential for classification improvement. Band-ratios have been employed to minimize the effects of topography on reflectance (Rock et al., 1986; Vogelmann and Rock, 1988; Vogelmann 1990; Colby, 1991; Ekstrand, 1994; Ekstrand, 1996), as has the Minnaert Correction (Smith et al., 1980; Teillet et al., 1982; Colby 1991; Ekstrand, 1996).

Another commonly applied method is the cosine correction, which accounts for differing solar incidence angles when the sun is not at zenith and assumes Lambertian reflectance properties of the target (Teillet et al., 1982; Jensen, 1996). Most studies employing the cosine correction have produced unsatisfactory results, especially at middle and low solar elevation angles, where over-correction of northern slope pixels occurs (Ekstrand, 1996). The cosine correction factor increases greatly as incidence angles approach 90 degrees and can cause radiance values for 8-bit radiometric data to reach saturation (Teillet et al., 1982). Areas in complete shadow ($\cos i \leq 0$) are set to zero radiance as there is by definition no incident solar irradiance, and diffuse sky irradiance is ignored. Thus, instances occur where brightly over-corrected pixels occur adjacent to pixels with zero radiance (Teillet et al., 1982). The cosine correction was not treated in this study.

The Minnaert Correction has been employed by many authors to account for both incidence angle and non-Lambertian reflectance properties of forested landscapes with mountainous topography. It was first used to describe the surface roughness of the moon (Minnaert, 1941) and has been adapted for terrestrial remote sensing. Its derivation can be found in Smith et al. (1980), Teillet et al. (1982), and Colby (1991), and is reproduced here for efficiency. The relationship between observed and normalized BVs under the non-Lambertian assumption is given as

$$BV_{o(\lambda)} = BV_{n(\lambda)} \cos^{k(\lambda)} i \cos^{k(\lambda)-1} \theta_n \quad (\text{eq. 1})$$

where

- $k(\lambda)$ = Minnaert Constant for band λ
- $BV_{n(\lambda)}$ = normalized brightness value of band λ
- $BV_{o(\lambda)}$ = observed brightness value of band λ
- θ_n = surface slope
- i = incident angle of solar beam

$BV_{n(\lambda)}$ is the radiance if the surface were horizontal given $BV_{o(\lambda)}$ over inclined terrain (Ekstrand, 1996). The $\cos i$ term can be calculated by

$$\cos i = \cos(90 - \theta_s) \cos \theta_n + \sin(90 - \theta_s) \sin \theta_n \cos(\Phi_s - \Phi_n) \quad (\text{eq. 2})$$

where

- θ_s = solar elevation
- Φ_s = solar azimuth
- θ_n = surface slope
- Φ_n = surface aspect

Linearizing equation 1 using a log transformation yields

$$\log(BV_{o(\lambda)} \cos \theta_n) = \log BV_{n(\lambda)} + k(\lambda) (\log \cos i \cos \theta_n) \quad (\text{eq. 3})$$

Equation 3 is in the slope-intercept form for a line $y = b_0 + b_1x$, and regression of $(\log \cos i \cos \theta_n)$ as the independent variable and $\log(BV_{o(\lambda)} \cos \theta_n)$ as the response variable, and can be used to generate the slope estimate b_1 , which equals the Minnaert Constant $k(\lambda)$. The empirically derived $k(\lambda)$ is then used in a Backwards Radiance Correction Transformation (Smith et al., 1980; Colby, 1991) given by

$$BV_{n(\lambda)} = (BV_{o(\lambda)} \cos \theta_n) / (\cos^{k(\lambda)} i \cos^{k(\lambda)} \theta_n) \quad (\text{eq. 4})$$

The primary objectives in the present study were to determine which normalization method or combination of methods generates BVs, which when analyzed with a pixel classifying algorithm, produces the most accurate classification over a large geographic extent, using general and few cover types. From a raw Landsat-7 Enhanced Thematic Mapper (ETM) image, the following images were generated and compared: I1- raw BV's with the addition of the ETM5/ETM4 ratio (5/4 ratio), Normalized Difference Vegetation Index (NDVI), and the ETM4/ETM3 ratio (4/3 ratio); I2- same values as I1 and aspect partitioned into two image subsets; I3- Minnaert Corrected image with new 5/4 ratio, NDVI, and 4/3 ratio calculated with Minnaert Corrected BVs; and I4- same values as I3 and aspect partitioned into two image subsets. Based on the error matrix for each classification an estimated Kappa statistic (Khat) was calculated and used to test the formal null hypotheses that no difference in accuracies exist between the various normalization procedures (i.e., $H_0: \text{Khat1} = \text{Khat2} = \text{Khat3} = \text{Khat4}$) (Congalton, 1991; Congalton and Green, 1999). Examination of Producer's and User's Accuracies was also done for class level effects (Story and Congalton, 1986; Congalton, 1991; Congalton and Green, 1999).

Methods

Image Preprocessing

Image preprocessing includes steps that were performed on the base imagery prior to the generation of any image derivatives. A single ETM scene (Path 13/Row 29; August 31, 1999) was reduced to include areas both within the boundaries of the White Mountain National Forest (excluding a small portion in Maine) and lying within an

elevation range from 610 to 1525 m. The image subset rectification accuracy was checked using 46 ground control points (GCPs). All GCPs were geocoded with the average of 10 static positions acquired using Forced Overdetermined Differential GPS (DGPS; i.e., satellites ≥ 5 , PDOP < 6.0 , Signal/Noise ratio > 6.0) with a Trimble Pathfinder Pro XR receiver. GCPs were collected to represent both the spatial extent and elevation range within the image. The GCPs at low elevations include road-road intersections, powerline corridor-road intersections, parking lots, gravel pits, and water body edges. At higher elevations, GCPs include water body edges, bald mountain summits, rocky outcrops, and man-made huts/cabins. The root mean square error (RMSE) using the full set of 46 GCPs was 1.3 pixels. The 46 GCPs were subsequently divided into a control set and a validation set. The control set was used to develop a first-degree geometric transformation model resulting in a root mean square error (RMSE) of 0.66 pixels. Application of the transformation model (nearest neighbor resampling) and comparison with the validation set of GCPs produced an RMSE of 0.97 pixels. Rectification to within a single pixel assured that plot sampling areas on the ground would closely correspond with representative training site spectra. A minimum Brightness Value (BV) per band subtraction was performed on each of the 6 reflective bands of imagery (thermal band omitted).

Topographic Normalization

Application of the Minnaert Correction requires a DEM of spatial resolution equal to or greater than the imagery to which the correction is applied (Civco, 1989). A DEM for the study area was assembled from USGS Level 1 DEM quadrangles converted into ARC/INFO GRID coverages, merged, and nearest neighbor resampled

from 31.4 to 28.5 meter grid spacing, to match the spacing of the base imagery. Small seams of NODATA were filled using the NIBBLE command, which replaces NODATA pixel values with the value of the nearest neighbor containing data. Two iterations of a 3x3 grid cell low pass filter were applied to generate the final DEM product, which was converted into an ERDAS Imagine image file. Low pass filters were used to reduce systematic errors present in the DEM (Brown and Bara, 1994). Slope (in degrees) and aspect images were derived from the DEM image in Imagine.

Topographic normalization of the imagery was performed using a non-Lambertian reflectance model with Minnaert Constants calculated for each ETM band. Minnaert Constants for each band were calculated by systematically sampling coregistered pixels from the ETM, slope, and aspect images, and inserting the values into equations 2 and 3. Sampled pixels were gridded and spaced 300 m in the x -, and y -directions resulting in 17455 sampling locations. From this sample, 1592 were dropped because they did not represent forest pixels, leaving a reduced set of 15863 pixels for derivation of a Minnaert Constant for each band. Pixel BVs were normalized scene-wide in ERDAS Imagine using equation 4 to solve for BV_n .

Aspect Partitioning

This treatment involved dividing a single image into two separate images based on pixel orientation. Pixels of aspect oriented away from the incident solar ray were segregated to one image subset, and pixels of aspect oriented toward the incident solar ray were segregated to another image subset. Using the solar azimuth at the time and date of image acquisition as 147 degrees, the image was partitioned into a subset

including “sunshade” pixels with aspects from 238° (through north) to 56°, and into a subset including “sunlit” pixels with aspects from 57° to 237° (Fig. 2.1).

Image Treatments

Derivative images were generated that included combinations of uncorrected band-ratios, Minnaert Corrected band-ratios and aspect partitioning (Table 2.1). Band-ratios employed were the 5/4 ratio, NDVI, and 4/3 ratio, and these were stacked with the single band features of each image. Band-ratio calculations for I1 and I2 were from raw BVs, and calculations for I3 and I4 were from Minnaert Corrected BVs. Contributions to target irradiance from atmospheric scattering (i.e., diffuse sky irradiance) were ignored. This scattering most greatly effects smaller wavelengths, and bands ETM1 and ETM2 were not used as classification features. Band ETM3 was included computation of NDVI and the 4/3 ratio.

Land cover Descriptions and Definitions

The thematic classification was designed to capture the major naturally occurring land cover types encountered within the WMNF study area. Because the use for the final classification product is directed to wildlife habitat modeling efforts, these cover types were selected based on both vegetation type and structure, and not toward discrimination of vegetation at the species level. The land cover types Hardwoods, Mixedwoods, Softwoods, High Mortality, Kampfzone, Krummholz, and Open Rock were classified thematically using the various topographic normalization treatments. For this study, a tree was arbitrarily defined as any woody stem with DBH greater than 7.6 cm, and a sapling defined as any woody stem with DBH less than 7.6 cm. Canopy cover

refers to vegetative cover contributed by trees only. Land cover classes are defined and described below, and are summarized in Table 2.2.

Hardwoods- This class is defined as containing a hardwood canopy contribution greater than 67% of the live canopy. It occurs largely unbroken throughout the lowland valleys up to an elevation of approximately 915 m. From 610-915 m elevation, the canopy consists of hardwood tree species including mountain paper birch (*Betula papyrifera* var. *cordifolia*), yellow birch (*Betula alleghaniensis*), mountain ash (*Sorbus americana*), sugar maple (*Acer saccharum*), and American beech (*Fagus grandifolia*). The understory may consist of saplings of the species listed above, but also encountered are hobblebush (*Viburnum alnifolium*), striped maple (*Acer pensylvanicum*), red maple (*Acer rubrum*), mountain maple (*Acer spicatum*), shadbush (*Amelanchier* sp.), red spruce (*Picea rubens*), and balsam fir (*Abies balsamea*).

Mixedwoods- Rising in elevation from 915-1160 m a continuous, albeit rapid transition into Mixedwoods occurs, possessing more than 33% or more of both hardwood and softwood canopy species. Hardwood overstory species continuing into the mixedwood zone include paper birch, yellow birch, and showy mountain ash, and the hardwood understory is largely comprised of these species as well. The softwood species include balsam fir and red spruce. The elevation where mixedwood forest begins is highly variable, and its elevation range narrow owing to steeper slopes, which incur a rapid climatic gradient. Larger spruce and fir reach into the overstory canopy and the understory fraction of these softwood species rises rapidly with elevation. Mixedwood forest grades into softwoods at approximately 1160 m.

Softwoods- From approximately 1160 m elevation to tree line (ca. 1370 m), the trees most encountered include balsam fir, red spruce, mountain paper birch, and showy mountain ash with balsam fir predominating. This class was defined as containing a softwood canopy contribution greater than 67% of the live canopy. The understory is mostly softwoods, but mountain paper birch and showy mountain ash saplings are also present. Changes into the vegetation types at higher elevations can occur gradually or rapidly depending on slope and wind exposure conditions.

High Mortality- Within the softwoods zone there are naturally occurring mortality areas known as fir waves that interrupt the continuity of the softwood zone (Fig. 2.2). Fir waves occur as alternating bands of live and dead balsam fir stands (Sprugel, 1974). Dead stands are often less than 50 m in width and therefore vegetation sampling (see below) usually included both live and dead portions of a wave. Within the dead portion of a fir wave very high stem density of regenerating balsam fir occurs along with some mountain paper birch and sometimes pin cherry (*Prunus pensylvanica*) amid standing dead fir stems of any size class. Where fir waves occur they can represent a major feature of the landscape and therefore were included in the classification. Other sources of mortality within the softwood zone were not distinguished and were collectively classed as High Mortality and defined as having a live to dead basal area ratio less than 2.0.

Kampfzone- Blending continuously with increasing elevation, Softwoods transition into Kampfzone (Hugentobler, 1994), characterized by nearly pure balsam fir with stems less than 7.6 cm DBH and very high stem density. This vegetation type occurs between the upper elevation limit (greater than 1200 m) of the Softwoods and the

lower elevation limit of the Krummholz class (below). There is an observable increase in live lateral branching compared to individuals of similar size occurring as forest understory, and there are many remnant standing or fallen dead snags. As there are no appreciable numbers of trees (i.e., stems with DBH > 3 in.), there is less than 10% canopy cover. A challenge to this classification effort is distinguishing between the kampfzone vegetation, the high mortality portions of fir waves, and Krummholz as these are all dominated by balsam fir and are structurally similar.

Krummholz- At elevations greater than 1220 m and onto windward slopes, kampfzone vegetation grades into Krummholz, characterized by greatly stunted and twisted stems of balsam fir with lesser amounts black spruce (*Picea mariana*). Dwarf forms of mountain paper birch persist along with alpine woody and herbaceous species. Soils are very shallow and much of the land surface is occupied by lichen-covered rocky exposures, boulders and coarsely weathered gravels. Krummholz most often occurs as a patchwork of islands rather than as continuous uninterrupted tracts, and is defined here as consisting of softwood vegetation less than breast height in stature.

Open Rock- Generally, from 1525 m elevation to the highest summits, krummholz vegetation fails to persist and grades into rocky areas sparsely vegetated with sedges, mosses, and alpine-tundra species. No woody trees or saplings are present. Additionally, this class occurs throughout the elevation range of the study as slides and rocky exposures maintained by rain, wind, and ice erosion.

Training Sites and Image Classification

Training sites for image classification were acquired from vegetation plot sampling for the Hardwoods, Mixedwoods, and Softwoods classes; and from photo-

interpretation for the High Mortality, Kampfzone, Krummholz, and Open Rock classes. As part of a bird-habitat modeling project during the summer of 1998, 192 stations were located on 34 transects with 250 m interval spacing along the established hiking trails on selected mountains. Each station, represented by a 50-m radius circle within the forested portion of the study area, was sampled for vegetation characteristics and location coordinates found with DGPS. Four 5-m circular sample plots were placed with stratified random locations within each station. Each station quadrant was constrained to hold only one plot, but randomly placed therein. Circular boundaries were determined using a Haglof Forestor DME 201 sonic rangefinder. Within each plot, trees were tallied using Biltmore sticks into size classes 7.6-15.2 cm, 15.2-22.9 cm, 22.9-30.5 cm, and >30.5 cm. All stems greater than 30.5 cm were additionally measured with a DBH tape. Diameter for each size class was taken as the midpoint of the diameter range (e.g., 11.2 cm for the 7.6-15.2 cm class) and used to estimate basal area for each size class. Basal area was calculated directly for the >30.5 cm size class using measured diameters. All stems were characterized by life status (live or dead) and species. Sapling density was estimated by traversing perpendicular transects through the plot center and tallying stem hits to a meter-stick held at breast height (Noon, 1981). Saplings were further characterized by life status and species. Estimates from the four plots were averaged to produce a station estimate. Canopy cover estimates were derived from foliage presence/absence counts using sighting tube densiometers (The GLOBE Program, 1997). Presence of canopy was defined as a tree's leaf or needle seen at the cross hairs of the densiometer. The canopy was sampled from 10 locations at 4 meter intervals

along a randomly placed transect extending outward from the station center. Two transects were run at each station for a total of 20 canopy sampling locations.

Training sites for the Hardwoods, Mixedwoods, and Softwoods classes were selected from among the 192 sample station estimates. Of these, the best 3-5 stations representing each of these class definitions were selected (Table 2.2.). Selection was first based on class definition (Table 2.2), but additionally on the percentage mixtures of total basal area and sapling densities. After digitizing and displaying the geocoded station centers on the imagery, a 3x3 pixel window centered on each station selected was chosen as a training site. A 3x3 pixel window approximately corresponds to the planimetric configuration of a 50-m radius circle, but it was assumed that the station center coincided with the pixel center within which it occurred. Given this assumption, training spectra correspond to the areas sampled on the ground.

Training sites for the High Mortality, Kampfzone, Krummholz, and Open Rock classes were selected based on photo-interpretation. Color-infrared (CIR) 1:15840 scale aerial photography from August 1995 was used to identify areas representing each class. Training sites for each class were identified and digitized directly onto the image as irregular polygons.

Training sites were first applied to the aspect partitioned subsets of I4 so that the number of “sunshade” training sites (18) balanced the number of “sunlit” training sites (17). The respective training site subsets were stored as Imagine AOI layers, so that the same sites could be retrieved and applied to all image treatments. For I1 and I3, which were not aspect partitioned, the two training site sets were combined producing 35 training sites applied to the entire image. Each image treatment was submitted for

supervised classification in ERDAS Imagine software using the Feature Space option and minimum distance decision rule for ambiguous pixels. Many of the training sites had non-invertible signature matrices precluding the use of a maximum likelihood classifying algorithm.

Accuracy Assessment

The performance of the image treatments was evaluated with error matrices (Story and Congalton, 1986; Congalton, 1991). Reference (“ground-truth”) data were collected using the CIR aerial photography examined with a 3x stereoscope. Fifty reference polygons for each class were selected throughout the extent of the study area (Congalton and Green, 1999). Polygons were traced onto an acetate overlay of the photographs and later digitized and labeled directly onto the raw imagery. Polygons that could not be located on the imagery were discarded and new polygons for the appropriate class selected. All reference polygons were stored as an Imagine AOI layer so that the same accuracy assessment sites were applied to all image classifications. In addition to each reference polygon having a single “correct” class label, additional labels were assigned where appropriate for the generation of a fuzzy error matrix (Congalton and Green, 1999).

The discrete error matrix presents the accuracy in the strictest sense, where a mismatch between the classified and reference labels is considered of equal magnitude no matter the degree of confusion. A fuzzy error matrix accepts additional reference labels for inclusion in accuracy computation, based on the analyst’s willingness to apply less weight to errors occurring within areas of transition between cover types. Additional labels were assigned in instances where cover types graded together so that some

classification errors were acceptable based on the use of the product. Adherence to strict measures of accuracy can result in the perception of poor classification when commission errors are not severe (Gopal and Woodcock, 1994). Both discrete and fuzzy error matrices were used in interpreting the results.

To determine if a classified polygon was labeled correctly, reference polygons were overlain with each classification. A classified polygon was considered correctly labeled if the majority of its pixels within the reference polygon matched the reference label. In instances of a majority tie, correct labeling was attributed if one of the tied classes matched the reference label. In cases where the labeling was incorrect, it was checked if the classification label matched the fuzzy condition of accuracy.

It has been argued that there is no best measure of classification accuracy, and that presentation of multiple measures is advised (Congalton, 1991; Stehman, 1997; Congalton and Green, 1999). Accuracy measures evaluated were the Overall, Producer's, and User's Accuracies for both the discrete and fuzzy assessments. Producer's and User's Accuracy permit assessment of individual class-level accuracies. Producer's Accuracy represents the percentage of correctly labeled polygons within a class from the total number of within class reference polygons, and reflects how well the analysis reproduced conditions on the ground. User's Accuracy represents the percentage of correctly labeled polygons within a class from the total number of polygons committed, including labeling errors, to the class. The Kappa statistic (khat; Congalton, 1991) from the discrete error matrices was used to formally test the null hypothesis of no difference in classification accuracies at the $\alpha = 0.05$ significance level.

Results

The relationships for empirical derivation of Minnaert Constants are shown in Figure 2.3. The slope $b_{1(\lambda)}$ of the least squares fitted line equals the Minnaert Constant ($k_{(\lambda)}$), so that $b_{1(\lambda)} = k_{(\lambda)}$ ($\lambda = 1$ to 6). Each $k_{(\lambda)}$ was then inserted into the topographical normalization model (eq. 4) and BVs adjusted for treatments I3 and I4. Bands 1 and 3 showed the greatest deviation from the Lambertian assumption ($k_{(\lambda)} = 1$), and bands 4 and 5 showed the least deviation.

Discrete and fuzzy error matrices, accuracies, and the discrete matrix k_{hat} for each image treatment are presented in Figures 2.4-2.7. Overall accuracies for all four treatments were similar and ranged from 59 to 63 percent. Likewise, k_{hat} varied little ranging from 52 to 57 percent. Kappa statistics for all image treatments were significantly different than zero, indicating that all performed better than random ($P < 0.0001$; Table 2.3.). Kappa statistics were used to test all pairwise comparisons of treatment classifications, and there were no significant differences between any of the classifications. Thus, rejection of the null hypothesis failed and suggested that no image treatment performed better than the others (Table 2.4).

Variability in class level accuracies delivered from the treatments was considerably greater than observed for Overall Accuracy (Table 2.5). The image treatments with the greatest frequency of possessing the highest Producer's and User's Accuracies were determined for each class (Table 2.6). Image treatments were considered tied for the highest accuracy if an image was less than 5% from the maximum. Image treatments without aspect partitioning (I1, I3) had the lowest frequency of having the highest class level accuracy for Producer's, User's, and

combined accuracies. I4 had the highest class level frequency of highest accuracy, with I2 next in rank. The relationships described above for the discrete error matrix apply to the fuzzy error matrix, except that I2 performed slightly better than I4 in User's and combined frequency of highest accuracy (not shown). Considering class level measures of accuracy, image treatment I4 performed best (Fig 2.8).

Discussion

Remote sensing applications in areas with high topographic relief have been problematic owing to differences in pixel orientation relative to the angle of incident solar radiation. The primary effects on target radiance arise result from 1) reduced target irradiance as the angle of incidence deviates from the surface normal, and 2) non-Lambertian reflectance properties of surface materials. (Smith et al., 1980; Teillet et al., 1982; Colby, 1991; Ekstrand, 1996; Richter, 1997). Richter (1997) divides topographic correction approaches into two types. The first type consists of scene-dependent empirical techniques including the use of band ratios (Holben and Justice, 1981); Vogelmann and Rock, 1988; Colby 1991; Ekstrand, 1996) or band specific regression constants, like the Minnaert Correction (Smith et al., 1980; Teillet et al., 1982; Colby, 1991; Ekstrand, 1996). The second correction type uses radiative transfer modeling to estimate radiance values at every pixel. The difficulty with the latter approach is that radiances, transmittances, and diffuse sky interactions must be estimated for every pixel, thus requiring the availability of atmospheric data, atmospheric models, and sensor calibration information (Richter, 1997). In this study, techniques of the first type were examined to determine if any competing topographic normalization techniques

performed best at producing a general land cover classification over an area exhibiting topographic relief.

Determination of the best image treatment was evaluated with an assessment of accuracy of resulting classifications. Because there is no single best measure of classification accuracy, multiple measures of accuracy were derived from error matrices (Congalton, 1991; Stehman, 1997; Congalton and Green, 1999). Discrete and fuzzy error matrices were used to examine Overall, Producer's and User's Accuracies. The kappa statistic was calculated for each discrete matrix and used to test the null hypothesis that all normalization procedure's classifications had equal Overall Accuracies (Congalton, 1991; Congalton and Green, 1999).

The differences in discrete overall accuracy for each image treatment were small and ranged from 59 to 63 percent. The khat values fell into the range described as "moderate agreement" between the classification and reference data (Landis and Koch, 1977), and there were no significant differences for any pair-wise comparison of image treatments. Thus, rejection of the null hypotheses that overall classification accuracies would differ between treatments was not possible. The range in overall accuracies from application of fuzzy rules (Congalton and Green, 1999) was also small (81 to 85 percent). Treatments I1 and I3 had equal overall accuracies in both the discrete and fuzzy assessments. These findings suggest that given the choice between normalization approaches used here, if one is searching for the greatest overall accuracy then one could chose the image treatment that is easiest or most cost-effective to apply.

Whenever data are represented by a single statistic like Overall Accuracy and khat, information can be lost. Examination of Producer's and User's Accuracies permits

further evaluation of classification performance at the class level (Story and Congalton, 1986; Congalton, 1991). A simple metric was constructed to determine which treatment most often produced the highest class specific accuracy (HCSA; Table 2.6). This metric ignores the magnitude of accuracy differences, except that a treatment was credited with HCSA status if its difference from the true maximum accuracy was within 5 percent. Among discrete matrices, the frequency at which each treatment produced the HCSA showed that the treatment using only band-ratios generated the lowest number of HCSA occurrences (I1, 5/14) followed by the treatment using Minnaert Correction and band ratios (I3, 6/14). Both of these treatments were performed without aspect partitioning, thus there were a greater number of HCSA from treatments with aspect partitioning (8/14 and 10/14 for I2 and I4, respectively). This finding suggests that using aspect partitioning can produce class specific improvement in classification accuracy. Moreover, treatment I2 had a greater HCSA than I3, which suggests that the aspect partitioning of imagery using only band ratios can provide a greater increase in HCSA than employing Minnaert Correction alone. The treatment with all combinations of techniques (I4) outperformed all other treatments by having the greatest number of HCSA occurrences (10/14). Among fuzzy error matrices, the pattern did not vary except that treatment I2 produced one more HCSA for User's Accuracy than I4.

The results from comparing overall and class level accuracies were mixed. The comparisons of overall accuracies were not significantly different, but they did include information about the magnitude of classification errors. Comparisons of class level accuracies using the HCSA metric did not include the magnitude of classification errors, but did suggest, albeit not through any objective statistical measure, that classifications

in mountainous terrain could be improved by Minnaert Correct in combination with aspect partitioning, or even by aspect partitioning alone.

Few previous studies have investigated the effects of Minnaert Correction on general land cover classification. The Minnaert Correction, which models the non-Lambertian properties of a target, has been found to be dependent on wavelength and cover type (Smith et al., 1980; Teillet et al., 1982; Ekstrand, 1996). Without generalizing assumptions (Teillet et al., 1982), cover type dependence of $k(\lambda)$ precludes its application in land cover classification, because it is spatially explicit cover type information that is sought. In this investigation, mapping cover type was the primary goal, and a generalizing assumption that groups all forest classes for computation of the Minnaert Constant had to be made. Land cover mapping with Landsat MSS data, Teillet et al. (1982) found that Minnaert Corrected spectral bands did not improve upon overall land cover classification accuracy versus uncorrected spectral bands for a coastal mountain area in British Columbia. Their models included parameterization from all cover types (i.e., generalized) and found no overall improvement of weighted mean classification accuracy. They concluded that no generalized correction model could improve cover type classification. The present study confirms these results, because there was no significant gain in overall accuracy using any of the applied treatments. Direct comparison with other studies employing the same treatments was not possible because the others did not 1) include multiple image treatments for comparison within the conditions of the study, or 2) they did not assess the effects on land cover classification. The TM 5/4 band-ratio has been used to successfully characterize vegetation health of softwood stands in mountainous terrain in the northeastern US

(Rock et al., 1986; Vogelmann and Rock, 1988, 1989; Vogelmann 1990). Colby (1991) compared the variability of raw TM bands with the variability in the 5/4 ratio at sites of similar vegetation composition, but different slopes and aspect. The 5/4 ratio was found to reduce variability between sites compared to the variability in the raw bands, indicating that the 5/4 ratio removed some of the topographic influence. Ekstrand (1996) found that the 5/4 ratio and Normalized Difference Vegetation Index (NDVI) performed well at removing the topographic effect in a defoliation study in Norway spruce (*Picea abies*) stands in Sweden. The 5/4 ratio failed to outperform against a single band near-infrared variable under conditions of moderate defoliation and no chlorosis (Ekstrand, 1994). Using Minnaert Correction in the Norway spruce defoliation study, Ekstrand (1996) found that the Minnaert Constants were not constant, but varied with the cosine of the solar incidence angle, and produced adequate topographic correction.

Fiorella and Ripple (1993) included incidence angle as a classification feature to map seral stages of coniferous land cover. They compared classification error matrices for treatments with and without an incidence angle feature, but found no difference in overall accuracy. They did find some class level accuracy improvement with the incidence angle feature treatment, but not for classes dominated by coniferous vegetation. They attributed greater error in labeling coniferous stands to non-Lambertian reflectance behavior. In this study, the conifer dominated classes Softwoods and Krummholz classified well, but Kampfzone and High Mortality performed poorly ranging from 20-22 percent. Further inspection reveals that Kampfzone was mislabeled as High Mortality 38-44 percent of the time, and from the fuzzy assessment this error

was acceptably correct 22-24 percent of the time. Confusion of these classes by the pixel-classifying algorithm is understandable because both 1) have open canopies and 2) are dominated by high stem densities of 1-5 m. tall balsam fir. The primary difference between the classes is large amounts of standing dead trees in the High Mortality class. Balsam fir within the *kampfzone*, while thick with needle-bearing branches on the lower two-thirds of the stem, often is devoid of needles on the apical one-third of the stem. Thus, the overall reflectance from these two classes were very similar, and based on the structural similarities of the attendant balsam fir, these two classes could be combined into a new class indicative of balsam fir sapling thickets.

Colby (1991) compared the effect of band-ratios and Minnaert Correction BV variability of lodgepole pine (*Pinus contorta*) stands in Colorado, USA. He concluded that both procedures effectively reduced the variability of BVs for stands of similar vegetation composition, but at different slopes and aspects. Furthermore, Minnaert Correction produced greater reduction in BV variability than band-ratios.

Conclusions

Comparisons of image treatments failed to reject the null hypothesis, because no image treatment outperformed the others in producing a greater Overall Accuracy. However, the highest class level Producer's and User's Accuracies were achieved most often when aspect partitioning was used. Aspect partitioning requires independent sets of training sites for each subset and may represent considerable cost if field data collection is used. Increases in accuracy with further stratification of aspect are unknown, but should be tested. Selecting training sites for multiple aspect strata from

large scale aerial photography could be an economical way to improve land cover classifications in montane environments.

Calculations for Minnaert Correction are not difficult, but do rely on slope and aspect data layers usually derived from a DEM. These DEM derivatives are of questionable accuracy and their effects on the Minnaert Constants is unclear. Band-ratios are the easiest to calculate and in combination with simple aspect partitioning may represent the most cost-effective normalization strategy.

Literature Cited

- Brown, D. G. and T. J. Bara. 1994. Recognition and reduction of systematic error in elevation and derivative surfaces from 7.5-Minute DEMs. *Photogrammetric Engineering and Remote Sensing* 60(2): 189-194.
- Civco, D. L. 1989. Topographic normalization of Landsat Thematic Mapper digital imagery. *Photogrammetric Engineering & Remote Sensing* 55(9): 1303-1309.
- Colby, J. D. 1991. Topographic normalization in rugged terrain. *Photogrammetric Engineering & Remote Sensing* 57(5): 531-537.
- Congalton, R. G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. of Environ.* 37: 35-46.
- Congalton, R. G. and K. Green. 1999. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. CRC Press. pp. 137.
- Ekstrand, S. 1994. Assessment of forest damage with Landsat TM: Correction for varying forest compartment characteristic's assessment of damage. *Remote Sens. of Environ.* 47:291-302.
- Ekstrand, S. 1996. Landsat TM-based forest damage assessment: Correction for topographic effects. *Photogrammetric Engineering & Remote Sensing* 62(2): 151-161.
- Fiorella, M. and W. J. Ripple. 1993. Determining successional stage of temperate coniferous forests with Landsat satellite data. *Photogrammetric Engineering and Remote Sensing* 59(2): 239-246.
- Globe Program. 1997. *The GLOBE Teacher's Guide*. Washington D.C.

- Gopal, S. and C. Woodcock. 1994. Theory and methods for accuracy assessment of thematic maps using fuzzy sets. *Photogrammetric Engineering and Remote Sensing* 60(2): 181-188.
- Holben, B. N. and C. O. Justice. 1980. The topographic effect on spectral response from nadir-pointing sensors. *Photogrammetric Engineering and Remote Sensing* 46(9): 1191-1200.
- Hugentobler, O. 1994. Waldgrenze und Kampfzone. *Bündnerwald* 47(1): 43-52.
- Jensen, J. R. 1996. *Introductory Digital Image Processing*. 2nd ed. Prentice Hall. Upper Saddle River, NJ. pp.316.
- Landis, J. and G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33: 159-174.
- Minnaert, M. 1941. The reciprocity principle in lunar photometry. *Astrophysical Journal* 93: 403-410.
- Noon, B. R. 1981. Techniques for sampling avian habitats. Pp. 42-50. in *The Use of Multivariate Statistics in Studies of Wildlife Habitat* (Capen, D. E., ed.). United States Forest Service General Technical Report RM-87.
- Richter, R. 1997. Correction of atmospheric and topographic effects for high spatial resolution satellite imagery. *Int. J. Remote Sensing* 18(5): 1099-1111.
- Rock, B. N., J. E. Vogelmann, D. L. Williams, A. F. Vogelmann, and T. Hoshizaki. 1986. Remote detection of forest damage. *Bioscience* 36(7):439-445.
- Smith, J. A., T. L. Lin, and K. J. Ranson. 1980. The Lambertian assumption and Landsat data. *Photogrammetric Engineering and Remote Sensing* 46(9): 1183-1189.
- Sprugel, D. G. 1974. Dynamic structure of wave-regenerated *Abies balsamea* forests in the North-eastern United States. *Journal of Ecology* 64: 889-911.
- Stehman, S. V. 1997. Selecting and interpreting measures of thematic classification accuracy. *Remote Sens. Environ.* 62:77-89.
- Story, M. and R. Congalton. 1986. Accuracy assessment: a user's perspective. *Photogrammetric Engineering and Remote Sensing* 52(3): 397-399.
- Teillet, P. M., B. Guindon, and D. G. Goodenough. 1982. On the slope-aspect correction of Multispectral Scanner Data. *Canadian Journal of Remote Sensing* 8(2): 84-106.

Vogelmann, J. E. 1990. Comparison between two vegetation indices for measuring different types of forest damage in the northeastern United States. *International Journal of Remote Sensing* 11(12): 2281-2297.

Vogelmann, J. E. and B. N. Rock. 1988. Assessing forest damage in high-elevation coniferous forests in Vermont and New Hampshire using Thematic Mapper data. *Remote Sens. Environ.* 24: 227-246.

Vogelmann, J. E. and B. N. Rock. 1989. Use of Thematic Mapper data for the detection of forest damage caused by pear thrips. *Remote Sens. of Environ.* 30: 217-225.

Image ID	Band Ratios	Minnaert Correction	Aspect Partition
I1	x		
I2	x		x
I3	x	x	
I4	x	x	x

Table 2.1. Table of Image Treatments. Summary of image treatments prior to supervised classification. All images include the band derivatives 5/4 ratio, NDVI, and 4/3 ratio. For images that have undergone Minnaert correction, the band derivatives were generated from source pixels that were Minnaert corrected.

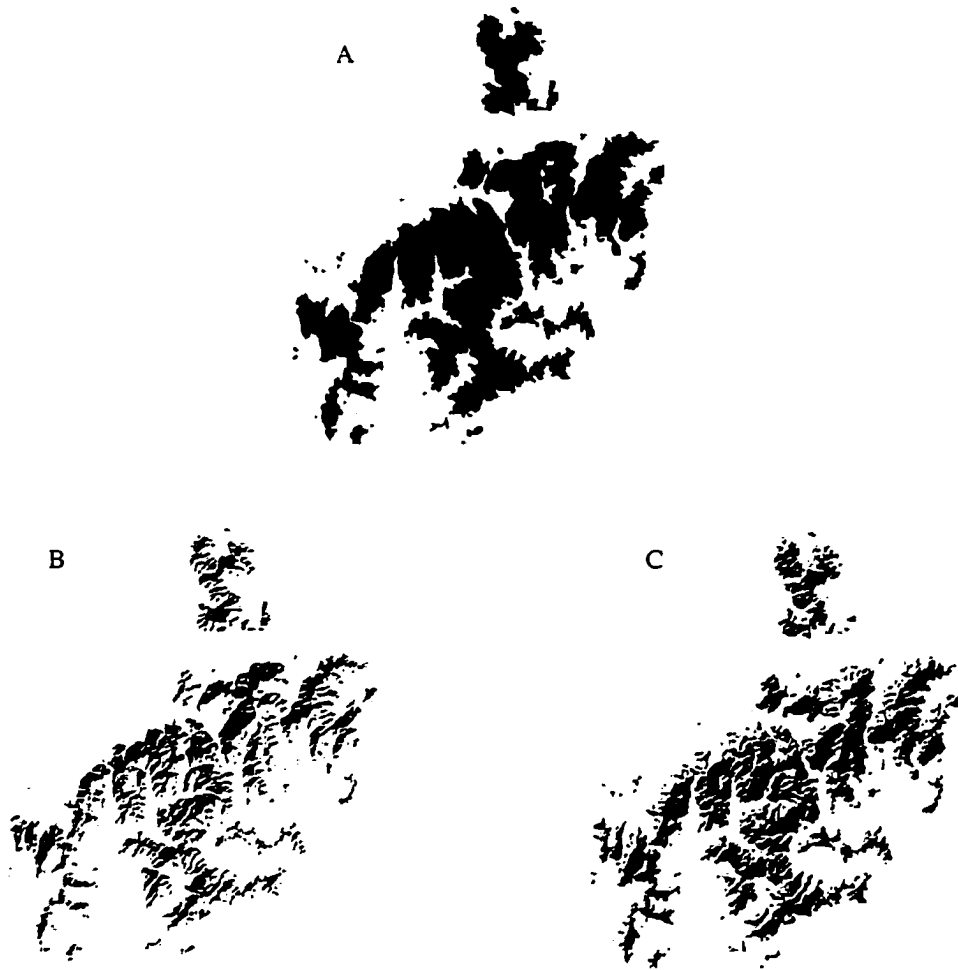


Figure 2.1. Aspect partitioning. The complete data set A has been partitioned into two image subsets. The solar azimuth is 147° . B represents “sunshade” pixels having aspects ranging from 238° through North to 56° . C represents “sunlit” pixels having aspects ranging from 57° to 237° . The term “sunshade” does not imply that the pixels are in shadow, only that the orientation is away from the incident solar beam.



Figure 2.2. Photograph within a fir wave near the summit of Mt. Cabot (Kilkenny, NH). Photograph shows the dead portion (High Mortality class) of the fir wave and was taken at the transition from the dead to live portion of the fir wave. Note the regenerating balsam fir (*Abies balsamea*) (A), and predominance of standing dead wood (B). Lack of significant canopy permits full light to reach the regenerating layer.

Definition	Description	Class
Canopy Cover > 10%	Forest	
Live:Dead Basal Area Ratio > 2.0	Live	
% Live Softwood Canopy Cover > 67%	Predominantly Live Softwoods	Softwoods
% Live Hardwood Canopy Cover > 67%	Predominantly Live Hardwoods	Hardwoods
% Live Hardwood Canopy Cover \geq 33% AND % Live Softwood Canopy Cover \leq 67%	Mixture of Live and Softwoods	Mixedwoods
Live:Dead Basal Area Ratio \leq 2.0	High Tree Mortality; fir waves	High Mortality
Canopy Cover \leq 10%	Non-forest	
Tree and Sapling Density > 0 stems/m ²	Tree and Sapling stems present, and greatly open canopy	Kampfzone
Tree and Sapling Density = 0 stems/m ²	Alpine	
Twisted/stunted softwood stems present	Alpine krummholz vegetation	Krummholz
Twisted/stunted softwood stems absent	Open rock dominates with some mostly herbaceous vegetation; includes slides and rocky outcrops	Open Rock

Table 2.2. Summary table of land cover class definitions and descriptions.

Treatment	khat (%)	Z score	p
I1	52.3	17.2	< 0.0001
I2	55.0	18.3	< 0.0001
I3	52.3	17.3	< 0.0001
I4	57.3	19.3	< 0.0001

Table 2.3. Table of results for testing the null hypothesis that image classifications were different than random (khat = 0). All classifications produced khat's significantly different from zero. Rejection of the null hypothesis occurs at $P < 0.05$.

Comparison	Z score	p
I1 and I2	-0.63	0.26
I1 and I3	0.00	0.50
I1 and I4	1.18	0.12
I2 and I3	-0.63	0.26
I2 and I4	0.55	0.29
I3 and I4	1.18	0.12

Table 2.4. Table of pairwise image classification comparisons using khat to test the null hypothesis that $I_j = I_k$. There was no image classification that was significantly different than the others. Rejection of the null hypothesis occurs at $P > 0.05$.

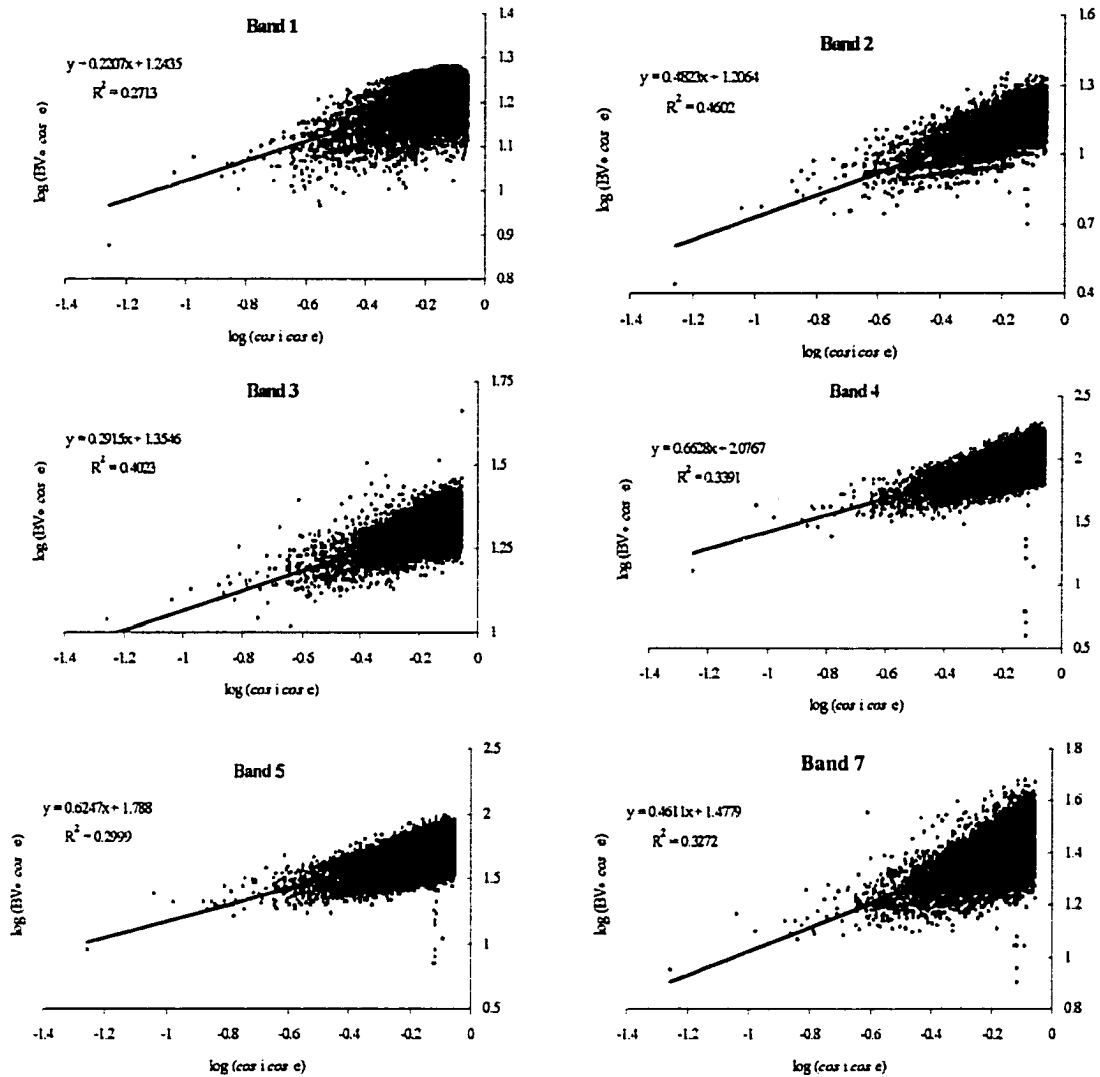


Figure 2.3. Scatterplots and regression equations for determining the Minnaert Constant $k(\lambda)$ for each band of the Enhanced Thematic Mapper Image. The slope of the regression line equals $k(\lambda)$.

Discrete Matrix		Reference							
		HW	MW	SW	KZ	KH	HM	OR	Totals
Image	HW	22	25	3	2	1	1	0	74
	MW	2	17	8	5	0	4	0	36
	SW	5	8	37	9	3	1	0	63
	KZ	0	0	2	33	2	14	1	30
	KH	0	0	0	1	57	6	11	56
	HM	0	0	0	22	5	24	0	51
	OR	1	0	0	0	1	0	38	40
Totals		50	50	50	50	50	50	50	350

Producer's Accuracy		User's Accuracy		
HW	84%	HW	57%	
MW	34%	MW	47%	Overall Accuracy
SW	74%	SW	59%	59%
KZ	22%	KZ	37%	
KH	76%	KH	68%	Khat 0.52
HM	48%	HM	47%	
OR	76%	OR	95%	

Fuzzy Matrix		Reference							
		HW	MW	SW	KZ	KH	HM	OR	Totals
Image	HW	42, 22	14, 11	0, 3	0, 2	0, 1	0, 1	0	74
	MW	0, 2	31, 17	3, 5	0, 5	0	0, 4	0	36
	SW	0, 5	6, 2	32, 37	9, 0	0, 3	1, 0	0	63
	KZ	0	0	1, 1	33, 16	2, 0	14, 0	0, 1	30
	KH	0	0	0	1, 0	57, 38	1, 5	10, 1	56
	HM	0	0	0	12, 10	0, 5	24, 24	0	51
	OR	0, 1	0	0	0	1, 0	0	38, 38	40
Totals		50	50	50	50	50	50	50	350

Producer's Accuracy		User's Accuracy		
HW	42/50 84%	HW	56/74 76%	Overall Accuracy
MW	37/50 74%	MW	20/36 56%	81%
SW	41/50 82%	SW	53/63 84%	
KZ	33/50 66%	KZ	28/30 93%	
KH	41/50 82%	KH	50/56 89%	
HM	40/50 80%	HM	36/51 71%	
OR	48/50 96%	OR	39/40 98%	

Figure 2.4. Discrete and Fuzzy Error Matrices for Classification of an image without Minnaert Correction or aspect partitioning (I1). Off diagonal elements in the fuzzy matrix are of the form a, b; where a equals the number of acceptable commission errors and b equals the number of unacceptable commission errors. a + b is the total number of commission errors of either type and equals the respective cell value in the discrete matrix above. Calculations for fuzzy accuracies are described in Congalton and Green (1999). Abbreviations are HW- Hardwoods, MW- Mixedwoods, SW -Softwoods, KZ- Kampfzone, KH- Krummholz, HM- High Mortality, OR- Open Rock.

Discrete Matrix		Reference							Totals
		HW	MW	SW	KZ	KH	HM	OR	
Image	HW	23	15	0	1	3	1	0	58
	MW	7	22	9	2	0	3	0	45
	SW	4	11	20	18	3	5	0	81
	KZ	0	0	1	10	3	7	1	22
	KH	0	0	0	0	25	7	8	50
	HM	0	0	0	19	5	25	0	51
	OR	1	0	0	0	1	0	41	43
Totals	50	50	50	50	50	50	50	350	

Producer's Accuracy		User's Accuracy		
HW	76%	HW	66%	
MW	48%	MW	53%	Overall Accuracy
SW	80%	SW	49%	61%
KZ	20%	KZ	45%	
KH	70%	KH	70%	Khat 0.55
HM	54%	HM	53%	
OR	82%	OR	95%	

Fuzzy Matrix		Reference							Totals
		HW	MW	SW	KZ	KH	HM	OR	
Image	HW	38	11, 4	0	0, 1	0, 3	0, 1	0	58
	MW	4, 3	22	4, 5	0, 2	0	0, 3	0	45
	SW	0, 4	8, 3	20	18, 0	1, 2	5, 0	0	81
	KZ	0	0	1, 0	10	3, 0	7, 0	0, 1	22
	KH	0	0	0	0	25	2, 5	7, 1	50
	HM	0	0	0	11, 8	0, 5	25	0	51
	OR	0, 1	0	0	0	1, 0	0	41	43
Totals	50	50	50	50	50	50	50	350	

Producer's Accuracy		User's Accuracy		
HW	42/50 84%	HW	49/58 84%	Overall Accuracy
MW	43/50 86%	MW	32/45 71%	85%
SW	45/50 90%	SW	72/81 89%	
KZ	39/50 78%	KZ	21/22 95%	
KH	40/50 80%	KH	44/50 88%	
HM	41/50 82%	HM	38/51 75%	
OR	48/50 96%	OR	42/43 98%	

Figure 2.5. Discrete and Fuzzy Error Matrices for Classification of an image without Minnaert Correction, but with aspect partitioning (I2). Off diagonal elements in the fuzzy matrix are of the form a, b; where a equals the number of acceptable commission errors and b equals the number of unacceptable commission errors. a + b is the total number of commission errors of either type and equals the respective cell value in the discrete matrix above. Calculations for fuzzy accuracies are described in Congalton and Green (1999). Abbreviations are HW- Hardwoods, MW- Mixedwoods, SW- Softwoods, KZ- Kampfzone, KH- Krummholz, HM- High Mortality, OR- Open Rock.

Discrete Matrix		Reference							Totals
		HW	MW	SW	KZ	KH	HM	OR	
Image	HW	25	2	2	1	2	0	78	
	MW	3	9	6	0	3	0	36	
	SW	1	9	10	4	4	0	65	
	KZ	0	0	2	0	8	0	21	
	KH	0	0	0	2	9	14	63	
	HM	0	1	0	19	6	0	50	
	OR	0	0	0	0	1	0	37	
	Totals	50	50	50	50	50	50	50	350

Producer's Accuracy		User's Accuracy			
HW	92%	HW	59%		
MW	30%	MW	42%	Overall Accuracy	
SW	74%	SW	57%	59%	
KZ	22%	KZ	52%	Khat 0.52	
KH	76%	KH	60%		
HM	48%	HM	48%		
OR	72%	OR	97%		

Fuzzy Matrix		Reference							Totals
		HW	MW	SW	KZ	KH	HM	OR	
Image	HW	14, 11	0, 2	0, 2	0, 1	0, 2	0	78	
	MW	2, 1	4, 5	0, 6	0	0, 3	0	36	
	SW	0, 1	4, 5	10, 0	1, 3	4, 0	0	65	
	KZ	0	0	1, 1	0	8, 0	0	21	
	KH	0	0	0	2, 0	2, 7	13, 1	63	
	HM	0	0, 1	0	11, 8	0, 6	0	50	
	OR	0	0	0	0	1, 0	0	37	
	Totals	50	50	50	50	50	50	50	350

Producer's Accuracy		User's Accuracy			
HW	48/50 96%	HW	60/78 77%	Overall Accuracy	
MW	33/50 66%	MW	21/36 58%	81%	
SW	42/50 84%	SW	56/65 86%		
KZ	34/50 68%	KZ	20/21 95%		
KH	40/50 80%	KH	55/63 87%		
HM	38/50 76%	HM	35/50 70%		
OR	49/50 98%	OR	37/37 100%		

Figure 2.6. Discrete and Fuzzy Error Matrices for Classification of an image with Minnaert Correction, but without aspect partitioning (I3). Off diagonal elements in the fuzzy matrix are of the form a, b; where a equals the number of acceptable commission errors and b equals the number of unacceptable commission errors. $a + b$ is the total number of commission errors of either type and equals the respective cell value in the discrete matrix above. Calculations for fuzzy accuracies are described in Congalton and Green (1999). Abbreviations are HW- Hardwoods, MW- Mixedwoods, SW- Softwoods, KZ- Kampfzone, KH- Krummholz, HM- High Mortality, OR- Open Rock.

		Reference							Totals
		HW	MW	SW	KZ	KH	HM	OR	
Image	HW	25	17	0	0	2	1	0	63
	MW	6	22	3	5	0	3	0	39
	SW	1	11	45	13	4	3	0	77
	KZ	0	0	2	10	4	6	1	23
	KH	0	0	0	1	5	8	9	51
	HM	0	0	0	21	6	2	0	56
	OR	0	0	0	0	1	0	40	41
	Totals	50	50	50	50	50	50	50	350

Producer's Accuracy		User's Accuracy			
HW	86%	HW	68%		
MW	44%	MW	56%	Overall Accuracy	
SW	90%	SW	58%	63%	
KZ	20%	KZ	43%		
KH	66%	KH	65%	Khat 0.57	
HM	58%	HM	52%		
OR	80%	OR	98%		

		Reference							Totals
		HW	MW	SW	Kz	Kh	FW	OR	
Image	HW	25	12, 5	0	0	0, 2	0, 1	0	63
	MW	4, 2	22	1, 2	0, 5	0	0, 3	0	39
	SW	0, 1	9, 2	45	13, 0	1, 3	3, 0	0	77
	Kz	0	0	0, 2	10	4, 0	6, 0	0, 1	23
	Kh	0	0	0	1, 0	5	1, 7	8, 1	51
	FW	0	0	0	11, 10	0, 6	2	0	56
	OR	0	0	0	0	1, 0	0	40	41
	Totals	50	50	50	50	50	50	50	350

Producer's Accuracy		User's Accuracy			
HW	47/50 94%	HW	55/63 87%	Overall Accuracy	
MW	43/50 86%	MW	27/39 69%	85%	
SW	46/50 92%	SW	71/77 92%		
Kz	35/50 70%	Kz	20/23 87%		
Kh	39/50 78%	Kh	43/51 84%		
FW	39/50 78%	FW	40/56 71%		
OR	48/50 96%	OR	41/41 100%		

Figure 2.7. Discrete and Fuzzy Error Matrices for Classification of an image with Minnaert Correction and aspect partitioning (I4). Off diagonal elements in the fuzzy matrix are of the form a, b ; where a equals the number of acceptable commission errors and b equals the number of unacceptable commission errors. $a + b$ is the total number of commission errors of either type and equals the respective cell value in the discrete matrix above. Calculations for fuzzy accuracies are described in Congalton and Green (1999). Abbreviations are HW- Hardwoods, MW- Mixedwoods, SW- Softwoods, KZ- Kampfzone, KH- Krummholz, HM- High Mortality, OR- Open Rock.

Class	Producer's Accuracy (%)	User's Accuracy (%)
HW	76 – 92	57 – 68
MW	34 – 48	42 – 56
SW	74 – 90	49 – 58
KZ	20 – 22	37 – 52
KH	66 – 76	60 – 70
HM	48 – 58	47 – 53
OR	72 – 82	95 – 98

Table 2.5. Table of ranges of Producer's and User's Accuracy for each land cover class. Abbreviations are HW- Hardwoods, MW- Mixedwoods, SW- Softwoods, KZ- Kampfzone, KH- Krummholz, HM- High Mortality, OR- Open Rock.

Treatments with Highest Class Specific Accuracy			Frequency each treatment produced the Highest Class Specific Accuracy			
Class	Producer's Accuracy	User's Accuracy	Image Treatment	Producer's Accuracy	User's Accuracy	Combined Accuracy
HW	I3	I2, I4	I1	2	3	5
MW	I2, I4	I2, I4	I2	4	4	8
SW	I4	I1, I3, I4	I3	3	3	6
KZ	I1, I2, I3, I4	I3	I4	5	5	10
KH	I1, I3	I1, I4				
HM	I2, I4	I2				
OR	I2, I4	I1, I2, I3, I4				

Table 2.6. Table of image treatment classifications yielding the highest discrete Producer's, User's, and combined accuracies. The left side of table lists by class and accuracy type, the image treatment that resulted in the highest accuracy. Treatments were considered tied for highest accuracy of their value was within 0.05 of the maximum for all treatments. The right side of the table summarizes the frequency at which a treatment classification resulted in the Highest Class Specific Accuracy. For example, I1 has two occurrences in the Producer's Accuracy column (KZ and KH), where it was or shared the ranking of highest accuracy. Abbreviations are HW- Hardwoods, MW- Mixedwoods, SW- Softwoods, KZ- Kampfzone, KH- Krummholz, HM- High Mortality, OR- Open Rock.

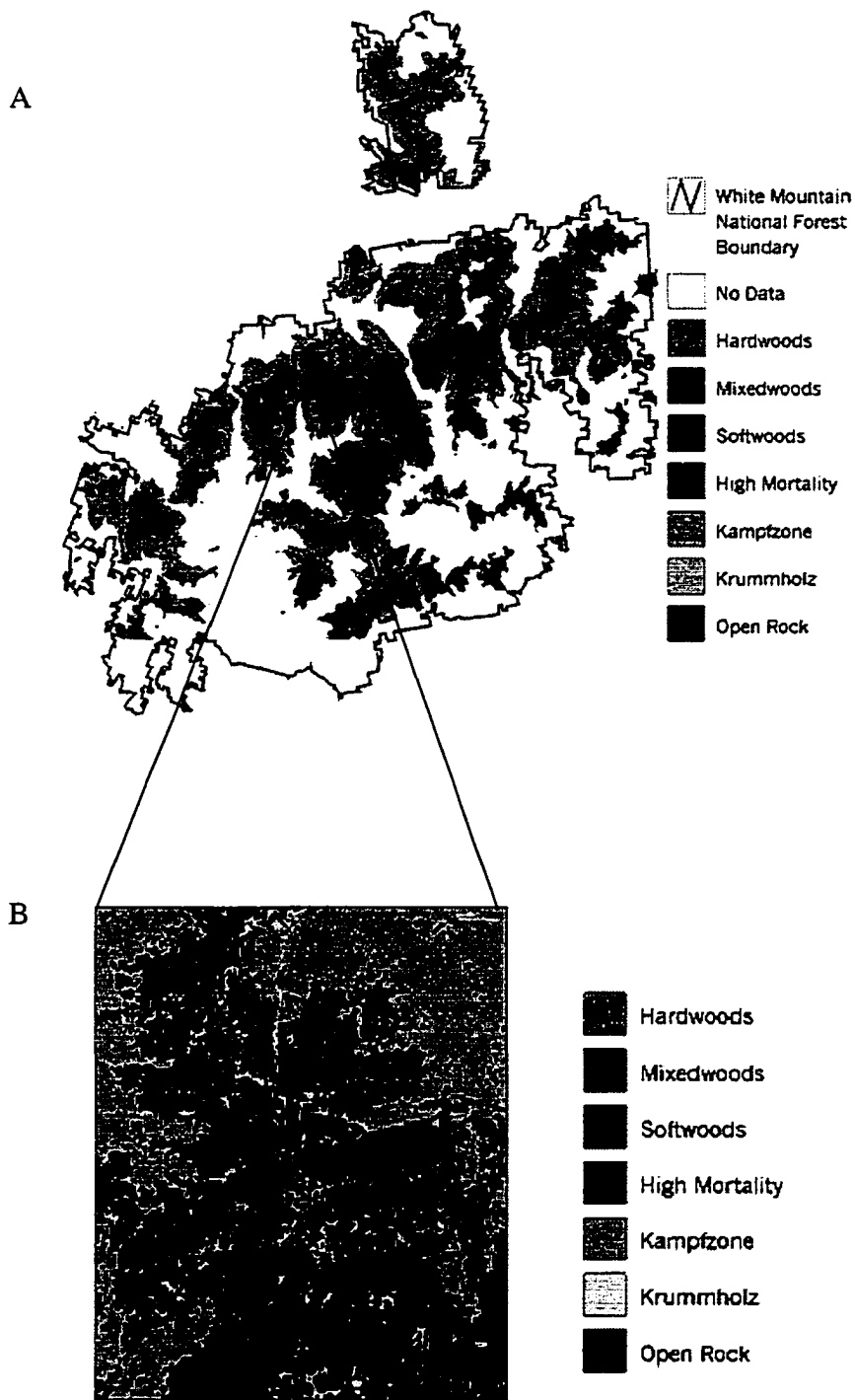


Figure 2.8. Treatment I4 thematic land cover classification of A) the White Mountain National Forest and B) the inset in A.

CHAPTER 3

IMPROVING LAND COVER CLASSIFICATION WITH ANCILLARY DATA

Introduction

Land cover classification using digital imagery is normally approached either by unsupervised or supervised classification using a pixel classifying algorithm (Jensen, 1996). These algorithms sort individual pixels based solely on their spectral similarity to a statistical measure of group spectral characteristics (unsupervised), or user defined spectral characteristics (supervised). However, thematic land cover classes are defined by criteria measurable by a human observer, and not by spectral characteristics. For example, “forested stand with greater than 66% canopy closure of softwood trees” could define a Softwoods class. The image analyst constructs an image-derived spectral signature as a proxy definition for the class, and applies it to determine class membership. Moreover, a complete classification contains several or many classes, which to avoid ambiguity, must have definitions that are exhaustive and mutually exclusive (i.e., each area on the ground must fulfill one and only one class definition). This requires classes to have discrete boundary definitions, which spectral signature definitions must reproduce.

Sorting continuous transitions in vegetation land cover into discrete classes denies the clinal properties of natural landscapes (Gopal and Woodcock, 1994). Thus, classification errors are expected at spectral feature space class boundaries, where transition between classes occurs. Given that these errors occur, the introduction of physical or biological ancillary information can improve upon the results of spectral classification (Hutchinson, 1982; Franklin and Wilson, 1992).

Adjusting spectra-based classification with ancillary data should be based on sound justification so that 1) it is expected that the ancillary information will perform better than the pixel classifier and 2) the ancillary information can be appropriately applied within the context of the map's purpose. In most cases, ancillary data sets were not produced with thematic classifications in mind, and may have severe limitations (Jensen, 1996). Three applications of ancillary information to the thematic land cover classification I4 (Chapter 1) were performed, 1) combining Kampfzone (KZ) and High Mortality (HM) into a single class, Fir Sapling (FS); 2) converting all Krummholz (KH) pixels below 1220 m to the FS class; and 3) inserting a new class, Water. All of these changes were considered in accordance with the overall purpose of the final classification. Specifically, the thematic map was intended for use in predictive mapping of the distribution of Bicknell's Thrush. Therefore, each proposed adjustment is attended by a justification related to this purpose. It is further noted that these changes may be inappropriate for alternative map uses for which it was not intended.

Merging Kampfzone and High Mortality classes

From the error matrices for I4 (Fig 2.7), it can be seen that KZ was poorly classified (Producer's Accuracy (PA) = 20%; User's Accuracy (UA) = 43%), rendering

analyses including KZ as a factor unreliable. The HM classification performed better (PA = 58%; UA = 52%), but still poorly for consistent identification of HM sites. Observe that 42% of the KZ reference polygons were mislabeled as HM, indicating that the pixel classifying algorithm frequently confused KZ as HM. This confusion is understandable, as the only difference between these two classes is the presence of standing dead fir and spruce stems in the HM class. The regenerating fir within the dead portions of fir waves is nearly structurally identical in all aspects with the fir of the KZ.

Provided no relationship exists between the presence or absence of BITH with the estimates of standing dead wood, then justification exists for combining KZ and HM into a new class, Fir Sapling. If a relationship were to exist, than preserving the identity of the HM class might be important. Testing this relationship (H_0 : slope = 0; $\alpha = 0.05$) using Logistic Regression found that the slope of the logits were not significantly different from zero for 1) Live:Dead Basal Ratio ($P = \chi^2_{(1, 1.38)} = 0.24$) or Basal Area of Standing Dead Wood ($P = \chi^2_{(1, 0.29)} = 0.59$). Because there was no association between BITH and the dead wood component at sampling locations, it is assumed that this component within fir waves is also unimportant. Therefore, the two classes were merged into the new class FS.

Changing Krummholz pixels below 1220 m to Fir Sapling

Throughout the range of elevation of the study area, there are many places with rocky exposures. Below 1220 m these exposures include wind, rain, and ice maintained outcrops and slides. Like krummholz, the vegetation in these exposed areas occurs in patches of shallow soils, and is mostly comprised of balsam fir. Unlike krummholz, the vegetation in exposed areas may have much greater species diversity, especially in the

form of small stature deciduous shrubs (e.g., *Vaccinium sp.*). When training sites for the Krummholz class were defined, the spectral signatures were from mixed pixels, which included 1) the krummholz vegetation type and 2) areas of exposed rock. Thus, spectra from these sites included features of vegetation with greater than normal reflectance, owing to the presence of exposed rock (Fig. 3.1). Using mixed pixels as the signature for KH was partly uncontrollable, and partly by design. It was uncontrollable because the patchwork nature of KH prohibited selection of these sites within a polygon of pixels without including rocky exposures. This property was used to advantage in that the elevated radiances in the signature allows the pixel classifying algorithm to separate KH from the similar KZ and SW classes (Fig. 3.1).

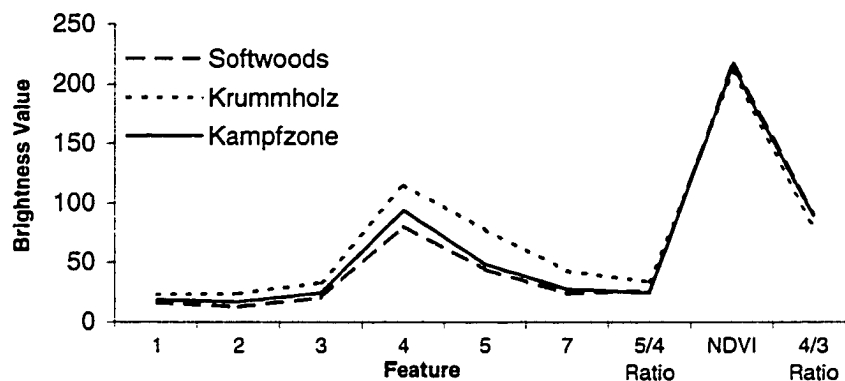


Figure 3.1. Graph of image feature mean Brightness Values for three land cover classes dominated by balsam fir (*Abies balsamea*). The presence of exposed rock in the training pixels for krummholz is evidenced by the elevated mean Brightness values in the 6 single ETM bands.

Below 1220 m krummholz vegetation does not occur, as the fir vegetation possesses neither the wind-worn, bent and twisted stem characteristics nor the stunted stature of KH. Instead the fir more closely approximates the structure found in the FS

class, namely high stem density with DBH less than 7.6 cm, and vegetation height from 1.5 to 5 m. Observations of BITH at elevations less than 1220 m have found them to occur in areas structurally similar to FS with large rocky exposures. Therefore, KH pixels below 1220 m were reclassified to its analog FS based on vegetation structure.

Insertion of the land cover class Water

Surface water features represent a minor fraction of the overall landcover of the study area, but where water does occur may play an important role in the distribution of BITH. Water features present include lakes, ponds, bogs and large and small stream drainages. While the remote sensing classification would have little problem classifying lakes, ponds and bogs, the streams would be more problematic. Streams are linear features within the landscape that 1) have widths substantially smaller than the image resolution (28.5 m) and 2) are obscured from direct sensor observation by overhanging adjacent trees. Spectral classification of forest streams would result in only haphazard placement of water pixels at locations of open canopy, leaving gaps in a stream's continuity. Therefore, because ancillary data for identifying surface water was available, attempts for spectral classification of water was not attempted.

The ancillary surface water data were from USGS 1:24000 surface hydrography quadrangles that were digitized and stored as ARC/INFO coverages with both polygon (lakes, ponds, bogs) and arc (streams) attributes, in accordance with US National Map Accuracy Standards. Coverage streams are represented as uninterrupted arcs, thus avoiding the problem encountered with spectral classifiers. However, these data were not compiled for use with digital remote sensing classifications, and georeferencing errors can occur with digitizing, projection transformation, non-planimetric conditions,

and registration with the base classification. Other errors result from vector to raster conversion of the data, causing streams to occupy full pixels even though it is known that higher order streams fall well below this dimension. This effect was ignored, because the variable of interest was distance to nearest surface water of any type, and the scale of this variable is much larger than the pixel dimension.

Methods

Merging Kampfzone and High Mortality Classes

To merge the KZ and HM classes, the KZ class value was recoded to equal the value of HM, and the label changed to FS. A potential side effect of this operation could be to alter the accuracy assessment error matrix, because combining the two classes may form a new FS majority and change a correctly labeled polygon to an incorrect one, or change an incorrectly labeled polygon to a correct FS label. New error matrices and accuracies were generated and evaluated using the same methods as in Chapter 1.

Changing Krummholz pixels below 1220 m to Fir Sapling

To change the KH label of pixels below 1220 m to FS, the DEM described in Chapter 1 was used to define the 1220 m threshold. A conditional expression was used at every pixel, so that if a pixel's elevation was less than 1220 m and it had the KH label, then the label was changed to FS. It was assumed that this transformation was correct, even though it was known that KH User's Accuracy was only 66%.

Inserting the class Water

USGS 1:24000 surface hydrography ARC/INFO vector coverages for the study area were tiled together to form a single GIS layer. This layer was then divided into two

coverages, one with only polygon attributes, and one with only arc attributes. There were two instances where streams forked high up on a mountain slope and merged again a great distance away. For the polygon coverage this resulted in two massive water bodies that were clearly wrong. These two polygons were deleted from the polygon coverage, but their identity as streams was preserved in the arc coverage.

Each coverage was rasterized and nearest neighbor resampled to the pixel size of the base classification. In Imagine, thematic class values were conditionally recoded. First, the thematic image was recoded using values from the water polygon image, so that if water was present at a pixel, the pixel value was recoded with the value for Water. This accomplished the insertion of lakes, ponds, and bogs. Second, the same method was applied with the rasterized image of streams. It was assumed that the insertion of the Water class was accurate.

Results

Discrete and fuzzy error matrices, accuracies, and the Khat statistic for the classification with merged KZ and HM classes into the single FS class are shown in Figure 3.2. Combining the two classes resulted in 100 reference polygons for the new FS class. The discrete overall accuracy of the thematic classification rose from 63% to 72%, and the Khat statistic rose significantly from 0.57 to 0.66 ($P_{(Z=2.10)} = 0.018$). The new FS class Producer's and User's Accuracies were 68% and 84%, respectively. Both of these class-wise measures of accuracy were substantially higher than for either of the original KZ or HM classes. The fuzzy accuracy assessment, including polygons that were acceptably correct, rose from 85% to 91%. The observed increases in accuracy were directly attributable to the merger of the KZ and HM classes. The effect of

converting KH pixels below 1220 m to FS and the insertion of the class Water are illustrated in Figure 3.3. Accuracy assessment of these adjustments was not performed, and it is assumed that the adjustments were correct.

Discussion

The I4 thematic classification was adjusted using ancillary data to improve classification accuracy and to insert land cover class information that would have been unsatisfactory from the spectra-based classifying algorithm. The combination of KZ and HM classes into FS was still based on spectral information only, but the decision and justification to combine the classes was based on outside information, namely that the difference between the class definitions was not relevant to the purpose of the study. Because the main difference between the KZ and HM classes was the presence of standing dead wood in the HM class, and there was no evidence of association between BITH and standing dead wood, then it was deemed justified to combine the classes. This resulted in an improved land cover classification, from which cover types could be used as factors in modeling the habitat of BITH. Changing KH pixels below 1220 m to FS further improves the classification, because it is known that KH rarely occurs below 1220 m. Although KH was classified with only 66% accuracy, the KH conversion to FS would be accomplished at a similar accuracy rate.

Insertion of the Water class cover type was completely exogenous to the spectral classification and brings advantages and disadvantages. The primary advantage from using the vector GIS coverage was representation of streams as continuous watercourses. However, several sources of error accompany this application including 1) errors in the source base map; 2) digital rendering of the base map; 3) georeferencing

thematic map registration; and 4) vector to raster data conversion. The insertion of the Water class was used to determine the minimum distance from a BITH station to any surface water feature. The scale of this variable across the study area is much greater than the variability associated with error in water placement. Thus, the advantage of continuous water representation outweighed the presence and magnitude of introduced spatial errors.

The operations using ancillary data were completed as final preparation of the thematic land cover map for use in the BITH habitat modeling project. The map is exhaustive, containing a class label for every location within the defined study area, and is acceptably accurate for all classes, in spite of the mountainous terrain.

Literature Cited

- Congalton, R. G. and K. Green. 1999. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. CRC Press. pp. 137.
- Franklin, S. E. and B. A. Wilson. 1992. A three-stage classifier for remote sensing of mountain environments. *Photogrammetric Engineering and Remote Sensing* 58(4): 449-454.
- Gopal, S. and C. Woodcock, 1994. Theory and methods for accuracy assessment of thematic maps using fuzzy sets. *Photogrammetric Engineering and Remote Sensing* 60(2): 181-188.
- Hutchinson, C. F. 1982. Techniques for combining Landsat and ancillary data for digital classification improvement. *Photogrammetric Engineering and Remote Sensing* 48(1): 123-130.
- Jensen, J. R. 1996. *Introductory Digital Image Processing*. Prentice Hall. Upper Saddle River, NJ. pp. 316.

Discrete Matrix		Reference						
		HW	MW	SW	FS	KH	OR	Totals
Image	HW	17	0	1	2	0		63
	MW	6	3	8	0	0		39
	SW	1	11	14	4	0		75
	FS	0	0	2	10	1		81
	KH	0	0	0	9	8		50
	OR	0	0	0	0	1		42
Totals		50	50	50	100	50	50	350

Producer's Accuracy		User's Accuracy		
HW	86%	HW	68%	
MW	44%	MW	56%	Overall Accuracy
SW	90%	SW	60%	72%
FS	68%	FS	84%	
KH	66%	KH	66%	Khat 0.66
OR	82%	OR	98%	

Fuzzy Matrix		Reference						
		HW	MW	SW	FS	KH	OR	Totals
Image	HW	12, 5	0	0, 1	0, 2	0, 1		64
	MW	4, 2	1, 2	0, 8	0	0		39
	SW	0, 1	9, 2	14, 0	1, 3	0		75
	FS	0	0	0, 2	6, 8	10, 0	0, 1	81
	KH	0	0	0	9, 0	6, 1		49
	OR	0	0	0	0	1, 0		42
Totals		50	50	50	100	50	50	350

Producer's Accuracy		User's Accuracy		
HW	47/50 0.94	HW	55/64 0.86	
MW	43/50 0.86	MW	27/39 0.69	Overall Accuracy
SW	46/50 0.92	SW	69/75 0.92	91%
FS	91/100 0.91	FS	78/81 0.96	
KH	45/50 0.90	KH	48/49 0.98	
OR	47/50 0.94	OR	42/42 1.00	

Figure 3.2. Discrete and fuzzy error matrices for classification of an image with Minnaert Correction and Aspect Partitioning (I4), with the Kampfzone and High Mortality classes combined into the single class, Fir Shrub. Off diagonal elements in the fuzzy matrix were of the form a, b ; where a equals the number of acceptable commission errors and b equals the number of unacceptable commission errors Congalton and Green (1999). The total number of commission errors of either type is $a + b$, and equals the respective cell value in the discrete matrix above.

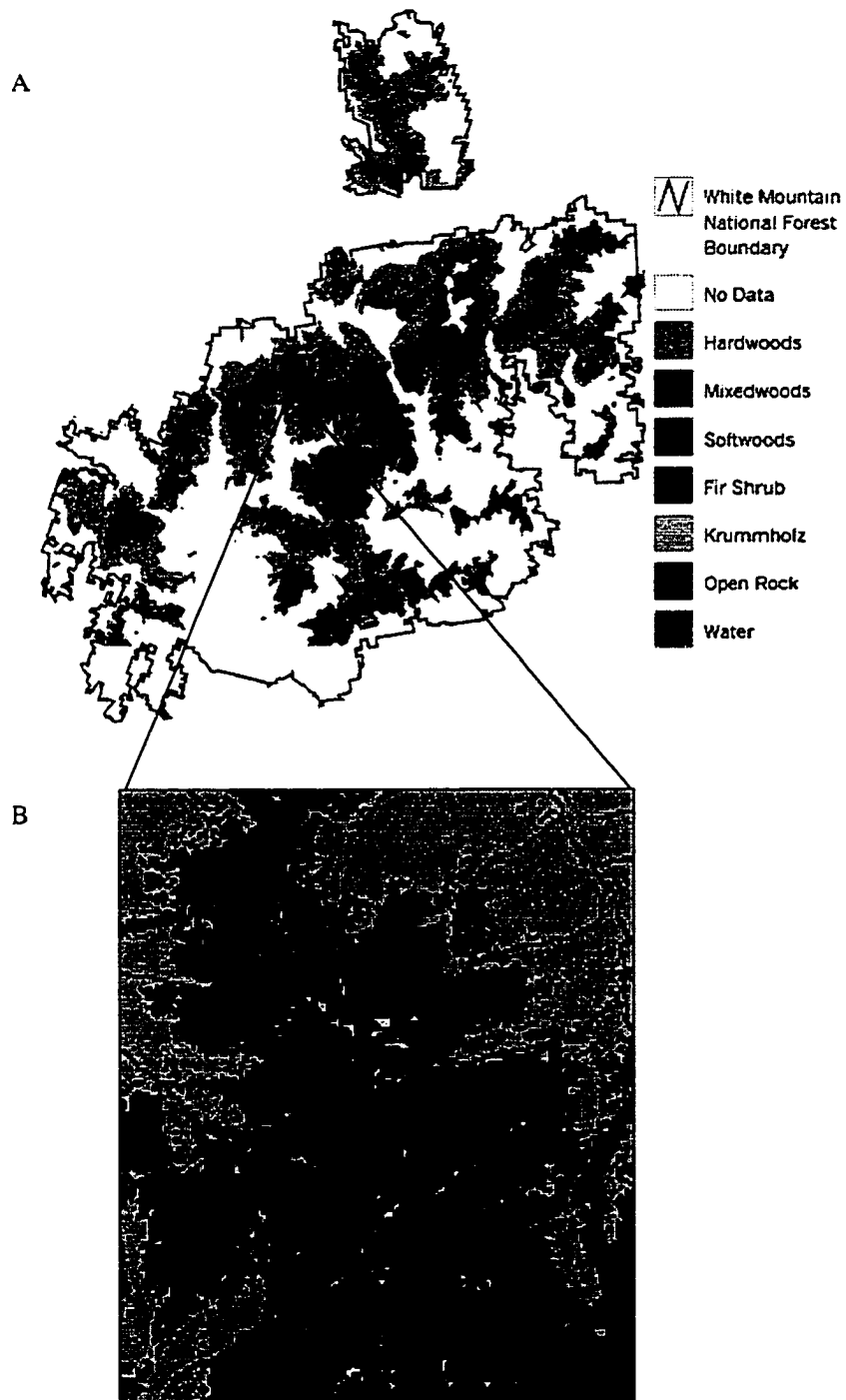


Figure 3.3. Land cover thematic classification of the White Mountain National Forest between the elevations of 610 and 1525 m. The geographical extent of inset (B) is depicted on the full map extent (A) as a black rectangle. This extent also precisely matches that of Figure 2.7B to facilitate direct examination of ancillary adjustments.

CHAPTER 4

MODELING BICKNELL'S THRUSH DISTRIBUTION USING ETM SATELLITE IMAGERY

Introduction

Modelers of forest bird habitat routinely measure biometric variables such as species, basal area, stem density, canopy closure, species composition, vegetation height, seral stage, forest damage, biomass and leaf-area-index (James and Shugart, 1970; James, 1971; Anderson and Shugart, 1974; Noon, 1981a; 1981b; Franzreb, 1983; Capen et al., 1986; Anthony et al. 1996; Petit and Petit, 1996). Unfortunately it is usually not possible to apply, in a spatially explicit way, models generated from these measurements, because variable estimates may not be available at locations not sampled. Remote sensing methods can solve this problem by providing spatially explicit habitat variable estimates at various spatial scales and across large geographic extents.

Satellite imagery comes in raster data layers composed of individual pixels, each containing a digital number representing sampled electromagnetic radiance from an area on the ground. Each data layer represents a band of the electromagnetic spectrum and each pixel's digital number (hereafter Brightness Value; BV) is proportional to the radiance detected at the sensor. Pixel BVs from spectral bands can be used as single features (i.e., one band) or to create additional features consisting of manipulations of multiple bands (i.e., band derivatives). Using satellite imagery, remote sensing analysts

of forest resources routinely estimate the same forest stand parameters and biometrics used by ornithologists, and do so over large geographic extents (Vogelmann and Rock, 1988; Sader et al., 1990; Spanner et al., 1990; Cohen and Spies, 1992; Franklin, 1994; Hall et al., 1995; Jakubauskas, 1996; Jakubauskas and Price, 1997). Therefore it follows that species-habitat models incorporating forest parameters, which can be estimated from satellite imagery, can be solved over large geographical extents.

Of published bird habitat modeling studies that use satellite imagery, none has utilized its potential for estimating model biometric variables. Most of these used Landsat-5 Thematic Mapper (TM) imagery to generate thematic classifications or maps of vegetative land cover as input into a geographical information system (GIS) (Green et al., 1987; Palmeirim, 1988; Franklin and Steadman, 1991; Herr and Queen, 1993). Analyses were performed with accompanying data layers to predict localities of species occurrence. Other studies have applied Bayes' Theorem to allow pixel BVs to be probabilistically recoded based on mean pixel BVs for sites with the species present versus those with the species absent (Aspinall and Veitch, 1993; Hepinstall and Sader, 1997). The products of these Bayesian-type methods are best described as species distribution maps rather than habitat maps, because no habitat model is used and no formal link between spectral and habitat variables is made.

The goal of this study was to determine if satellite imagery spectral features could be applied to a species habitat model (1) in a spatially explicit way and (2) over a large geographic extent. To effectively demonstrate this requires model predictions be validated using independent data, which were not included in model building. This study describes the construction of a probabilistic habitat map for Bicknell's Thrush

(BITH; *Catharus bicknelli*) on the White Mountain National Forest (WMNF) of New Hampshire. First, a multiple logistic regression habitat model using BITH presence/absence data was generated from habitat data of three types: (1) microhabitat forest biometrics, (2) a land-cover classification based upon Landsat-7 Enhanced Thematic Mapper data (ETM), and (3) elevation data acquired using Differential GPS (DGPS). This represents the familiar multivariate approach to species-habitat modeling, which returns coefficients and variables important in explaining the distribution of a species. However, in order to apply these models over large geographic extents, estimates for all model variables must be made at all locations. Estimates of forest biometric variables (e.g., dominant vegetation height) were not available for the study extent. Thus, the second step was to use ETM spectral features to generate estimates of forest biometric variables included in the habitat model.

A BITH habitat model was generated, applied, and validated on the WMNF. From independent sampling stations, the predictive map of BITH presence validated well up to a probability of 0.60. At probabilities above 0.60, the model failed to accurately represent BITH presence, but these values occurred in only 3% of the study area.

Methods

BITH Sampling

Presence or absence BITH data were derived from 5-minute point counts performed as part of the WMNF High Elevation Bird Monitoring Program (HEBMP). From 1993-1997, the HEBMP has amassed point-count observations at over 600

stations along 41 transects, performed annually in mid-June by observers trained in visual and acoustic identification of high elevation birds of the WMNF. Surveys were generally performed under conditions favorable for detecting birds, including minimal wind and rain, and constrained between 0600 to 1100 EDT. Local sunrise in June varied from 0457-0501 hours. Bird survey transects followed the existing trail system, with each station separated by 250m of trail distance (i.e., not straight line distance). Stations were not permanently marked, but were revisited by dragging a 62.5m nylon rope four lengths. Each transect had an identifiable landmark as a starting point. All visual and acoustic observations were recorded as (1) estimated >50 m from the point, (2) \leq 50 m from the point, or (3) flyover. The vast majority of observations were acoustic and assumed to represent singing males. Note, however, that female BITH may sing frequently when on the nest (Wallace, 1939).

Point-count stations from the HEBMP were included in this study if (1) they were surveyed a minimum of 4 of 5 years, (2) time constraints permitted visitation for field assessment, and (3) location was determined using Differential GPS (DGPS). A total of 471 stations on 34 transects was used, and over 80% of stations were surveyed in each of the 5 years. A station was scored as “BITH present” if in any of the survey years a BITH was recorded within 50 m of the station, or as “BITH absent” if no BITH was recorded within 50 m of the station. Birds recorded beyond 50 m from the station, or as flyovers were not considered.

Satellite Imagery and Raster Data

A cloud-free 31 August 1999 ETM image (path13/row29) with complete coverage of the WMNF was used to derive a thematic land cover classification, and

provide spectral data for estimating biometric variables. The image was reduced to include pixels both within an elevation range of 610-1525m, and within the New Hampshire portion of the WMNF. The raster format of the image served as the basis for habitat mapping and consisted of 28.5 x 28.5m pixels projected in New Hampshire Stateplane coordinates. From the image, a thematic land cover classification was derived (Chaps 2, 3), and spectral feature BVs were used to estimate biometric variables. Image spectral BVs were adjusted to minimize terrain effects by employing a scene-dependent Minnaert Correction on each spectral band (Chap2). The corrected six reflective ETM bands and band derivatives 5/4 ratio, Normalized Difference Vegetation Index (NDVI), and 4/3 ratio were used to evaluate and apply relationships between spectral features and forest biometric variables. A DEM was assembled from existing USGS Level 1 DEMs (Chap.2), nearest neighbor resampled from 31.4 to 28.5m, and twice smoothed with a low pass filter. Input, maintenance, and analysis of raster data sets were performed using ERDAS Imagine (v8.4) and ARC/INFO (v8.0) software.

Habitat Characterization

During summer 1998, all 473 stations were visited and geographic coordinate positions acquired using Forced Overdetermined DGPS (i.e., satellites ≥ 5 , PDOP < 6.0 , Signal/Noise ratio > 6.0) with a Trimble Pathfinder Pro XR receiver. Of these, 192 stations were randomly selected from within transects (about 6 stations per transect) for habitat model calibration, and the remaining 279 were set aside for independent model validation. Each calibration station, represented by a 50 m radius circle within the forested portion of the study area, was sampled for vegetation characteristics. Four 5m circular sampling plots were placed with stratified random locations within each station.

Each quadrant was constrained to hold only one plot, but randomly placed therein. Plot boundaries were determined using a Haglof Forestor DME 201 sonic rangefinder.

Within each plot, trees (diameter-at-breast-height [DBH] ≥ 7.6 cm) were exhaustively tallied by size classes 7.6-15.2 cm, 15.2-22.9 cm, 22.9-30.5 cm, and >30.5 cm, using Biltmore sticks (James and Shugart, 1971). All stems >30.5 cm were measured with a DBH tape. Diameter for each size class was taken as the midpoint of the diameter range (11.2 cm for the 7.6-15.2 cm class) and used to estimate basal area for each size class. Basal area was calculated directly for the >30.5 cm size class using measured diameters. All stems were also characterized by life status (live or dead) and species. Dominant vegetation height was estimated by measuring the height of the tallest living stem in each plot using a Suunto clinometer. Plot estimates were averaged to produce station estimates for each variable.

Sapling (DBH < 7.6 cm) density was estimated by traversing two perpendicular transects through the plot center and tallying stem hits to a meter-stick held at breast height (James and Shugart, 1971; Noon, 1981a). Saplings were further characterized by life status and species. Estimates from the four plots were averaged to produce a station estimate.

Canopy cover estimates were derived from foliage presence/absence counts using sighting tube densiometers. Presence of canopy was defined as a tree's leaf or needle seen at the cross hairs of the densiometer (GLOBE Program, 1997). The canopy was sampled from 10 locations at 4 m intervals along a randomly placed transect extending outward from the station center. Two transects were run at each station for a

total of 20 canopy sampling locations. Canopy was recorded as present or absent along with tree species contribution.

Landscape Characterization

A seven-class land cover thematic map for the study area was used to extract landscape scale variable estimates to be used as regressor variables. The seven land cover types included Hardwoods (HM), Mixedwoods (MW), Softwoods (SW), Fir Sapling (FS), Krummholz (KH), Open Rock (OR), and Water (W) (Chapters 2, 3). The thematic map was imported into ARC/INFO as a GRID. From this grid, landscape richness (Forman and Godron, 1986) variables were calculated at two scales. First, a 3 x 3 cell analysis window centered on the grid cell containing the station, and second a 7 x 7 analysis window were used to tally the total number of land cover classes within each window. A 3 x 3 analysis window has dimension 90 x 90 m and closely approximates the extent of the 50-meter radius circle of each station. Furthermore, the area of this window equals 0.81 ha and approximates the 0.6 ha territory size reported for BITH (Wallace, 1939; Rimmer et al., 1996). The 7 x 7 analysis window was used to determine if land cover diversity at a scale much greater than a territory was significant. Minimum distance variables from each station to the nearest polygon of each cover type were estimated after grid to vector conversion of the land cover map. The station was assigned a distance of zero for the cover type within which it resided. Distances were planimetric and not over-ground.

Slope, aspect, and elevation estimates were obtained from two different sources. Slope and aspect estimates at each station were derived from the DEM, and elevation estimates at each station center were obtained from the DGPS record of elevation.

Logistic Regression Analyses

Logistic Regression Analyses were performed following the procedures of Hosmer and Lemeshow (1989) using S-Plus 2000 statistical software. A total of 25 regressor variables was screened using univariate logistic regression with BITH presence/absence as the binary outcome variable (Table 4.1; see below for details). Variables were considered candidates for multivariate modeling if their resulting P-value was ≤ 0.10 . All regressor variables were treated as continuous variables, except for Aspect, which was grouped into 4 nominal classes representing the four cardinal directions. Aspect data are circular scale and thus possess arbitrary magnitudes (Zar, 1999).

After screening for candidate variables, two multivariate logistic regression models were developed. The first model represented the best fit considering all candidate variables. For this, a step-wise P-value criterion of 0.10 was used for inclusion of a regressor to the model, and a backward check for elimination P-value criterion of 0.15 was used for variable removal (Hosmer and Lemeshow, 1989). Because the goal of this study was to produce pixel level predictions of BITH presence, it was necessary that an estimate of all variables in the model be made at every pixel. For variable estimates garnered from plot sampling, this was not assured. A second model was produced with the step-wise ranking and P-value criteria as a guiding principle rather than a rule. This permitted alternative variables that could be mapped at every pixel to enter into the model. Thus, the second model can be viewed more as a management model, which can be applied and validated (Mosher et al., 1986).

Goodness-of-Fit

To assess the fit of the applied logistic model a grouping strategy was used to handle the binary nature of the BITH data (Hosmer and Lemeshow, 1989). Model output provided a probability estimate of BITH presence at each pixel based on the covariate pattern of coregistered pixels from all data layers. Field data of BITH presence/absence were collected from a circular area of 50-meter radius. A probability for an area comparable to that of the field data was calculated as the mean probability for a 3x3 pixel window centered on the station center. The Hosmer-Lemeshow Goodness-of-Fit statistic (\hat{c} ; Hosmer and Lemeshow, 1989) was calculated with stations grouped into deciles of probability for BITH presence to test the hypothesis that the observed number of BITH presence stations equals the expected number of BITH presence stations, with a critical alpha of 0.05. The statistic \hat{c} is chi-square distributed with degrees of freedom equal to the number of groupings minus 2 for assessing the fit of model calibration data, and degrees of freedom equal to the number of groupings for model validation data (Hosmer and Lemeshow, 1989). Chi-square values for predicted versus expected values below the critical alpha are deemed significantly different and represent a poor fit of the model to the reference data.

Results

Univariate Analyses

Univariate logistic regression using BITH presence/absence as the binary response variable produced 13 continuous variables as potential candidates for the multivariate logistic model (Table 4.2). The nominal scale Aspect variable was not

significant ($P > 0.48$). The amount of change in each variable that resulted in a doubling of the odds of BITH presence is shown in Table 4.3. For interpretation of the Double Odds Scale, the sign indicates an increasing or decreasing relationship and the variable's units must be referenced (Table 4.2). For example, every 2.38 m reduction in dominant vegetation height (VEGTHGT) produces a doubling of the odds associated with BITH presence. Likewise every 131 meter increase in elevation (ELEVATN) doubles the odds associated with BITH presence. The odds of incurring a BITH presence result rise (1) with increasing elevation, basal area of 7.6-11.2 cm DBH softwood stems, distance to the nearest hardwood land cover type, live softwood sapling density, and percentage of live softwood basal area; and (2) with decreasing dominant vegetation height, percentage of total canopy cover, basal area of softwood stems greater than 11.2 cm DBH, total live basal area of all trees, distance to nearest fir sapling land cover type.

Model Construction and Calibration

The best multivariate logistic model resulting from the 13 candidate variables included dominant vegetation height, elevation, and basal area of softwood trees with 22.9-30.5 cm DBH (Table 4.4). The logit equation for this model was

$$g(x) = - 0.156 * VEGTHGT + 0.00286 * ELEVATN - 0.1039 * BASW9-12 - 2.010 \quad (\text{eq. 4.1})$$

All three variables in this model were estimated from data sources that do not exhaustively cover the study area (i.e., plot sampling and DGPS). To apply the model for the entire study area, each variable had to be estimated by data sources that do cover the entire study area.

Elevation estimates from the DEM closely agree with the DGPS elevations taken at the 192 sampling stations (mean difference = 6.4 m; s.d. = 9.5 m; $R^2 = 0.996$; Fig. 4.1). VEGHGT and BASW9-12 estimates were derived from plot sampling and results for estimating these variables using spectral features the ETM imagery were mixed. To reduce the effects of within station variability presented by VEGHGT, a reduced set of stations ($n = 32$) having coefficients of variation (CV) less than 10% were used to model a relationship with imagery spectral features. Stations in this subset more closely approximate stands displaying VEGHGT homogeneity, which was important for constructing the relationship with spectral features. Using multiple linear regression, mean VEGHGT station values and spectral feature means from a 3x3 grid cell window centered on the cell containing the station center were used to estimate the relationship among VEGHGT spectral features. Both NDVI and Band1 explained significant amounts in VEGHGT variation (adjusted multiple $R^2 = 0.50$; Table 4.5) and resulted in the fitted equation

$$\text{VEGHGT} = 0.2107 * \text{NDVI} - 1.0882 * \text{Band1} - 15.82. \quad (\text{eq. 4.2})$$

No significant relationship between spectral features and BASW9-12 was found, and this variable had to be dropped. This resulted in an applicable two variable model including VEGTHGT and ELEVATN, which were somewhat correlated ($r = -0.58$). Thus, probabilities of BITH presence would rise with increasing elevation and the concomitant decrease in dominant vegetation height. The guiding P-value criterion for variable entry into the model was relaxed so that an additional variable could enter the model that was not correlated with the other regressors, and could be applied at all locations. The variable fitting these criteria was distance to nearest Fir Sapling cover

type (DIST2FS; partial $P < 0.15$). Inclusion of the DIST2FS regressor would increase estimated probabilities for BITH presence at low elevation areas where this cover type can occur, and thereby account for observations of BITH at lower elevations, where they sometimes occur. The DIST2FS variable was estimated from the land cover map with complete coverage of the study area, and therefore no additional data sets are necessary to provide estimates. The logit for the fitted management model (Fig. 4.2) was

$$g(x) = -\text{VEGTHGT} * 0.1885 + \text{ELEVATN} * 0.00270 \\ - \text{DIST2FS} * 0.00157 - 1.644 \quad (\text{eq. 4.3})$$

Because BITH is a forest bird and does not occur in areas devoid of vegetation, mapped probabilities for pixels with a VEGHGT estimate ≤ 0 were arbitrarily set to 0.001.

Solving for $g(x)$ from equation 4.3 at every pixel and following the conditional relationship

$$\pi(x) = 0.001 \quad \{\text{VEGHGT} \leq 0\} \\ \pi(x) = e^{g(x)} / e^{1 + g(x)} \quad \{\text{VEGHGT} > 0\} \quad (\text{eq. 4.4})$$

where $\pi(x)$ is the probability of BITH presence, provided the probability for BITH presence at every pixel.

The probability estimates should agree with the model calibration data from which they were derived. However, because the VEGHGT term in the model was estimated from a data source not used in model calibration, the fit of the model to the calibration stations had to be assessed. The model estimates were significantly different than the observed probabilities ($\hat{c} = 14.4$; $P = 0.01$), and systematically underestimated the probabilities of BITH presence at calibration stations by an average of 0.112 (Table 4.7). A new version of the model was generated that added a probability of 0.112 to all

pixels with VEGHGT estimates greater than zero. Pixel estimates were recalculated, and the fit at calibration stations was not significantly different than observed ($\hat{c} = 14.4$; $P = 0.87$; Fig. 4.3). Thus, the final model was

$$\begin{aligned} \pi(x) &= 0.001 && \{\text{VEGHGT} \leq 0\} \\ \pi(x) &= e^{g(x)} / e^{1 + g(x)} + 0.112 && \{\text{VEGHGT} > 0\} \quad (\text{eq. 4.5}) \end{aligned}$$

Model Validation

Data from 281 stations not included in the model calibration procedures were independently used to validate the final model. Considering validation stations collectively across the estimated decile groupings ranging from 0 to 0.60, model estimates were not significantly different than expected ($\hat{c} = 14.4$, $P = 0.10$; Table 4.8; Fig. 4.3). Within the deciles ranging from 0 to 0.60, significant deviation in the 0.40 to 0.50 decile was observed ($\hat{c} = 26.9$, $P = 0.0007$), with the model underestimating the number BITH presence stations. At probability estimates > 0.60 , the numbers of BITH presence records fell dramatically compared to model predictions. Thus, the model is ineffective at probabilities greater than 0.60. Note however, the area of pixels with probabilities greater than 0.60 occupied 3% of the total study area, and thus represents a minor fraction where the model failed. Model output at three scales are shown in Figure 4.5.

Discussion

Within the northeastern United States, BITH's summer distribution includes the montane coniferous forests from 915 m to tree line in the Adirondack Mountains of New York, Green Mountains of Vermont, White Mountains of New Hampshire, and Mt.

Katahdin in Maine (Wallace, 1939; Atwood et al., 1996; Rimmer et al., 1996).

Historically, BITH occurred over a wider range that included the Catskills of New York, Mt. Greylock of Massachusetts, and Mt. Monadnock of New Hampshire (Wallace, 1939), but recent surveys failed to reaffirm their presence (Atwood et al., 1996). In Canada the distribution of BITH is not as dependent upon elevation as in the northern Appalachian chain. Records from Canada include sea-level nesting observations along the coast and on islands of the Canadian Maritime provinces, the coniferous mountain zones of the Gaspé peninsula, and areas of second growth or recent burns in southern Quebec. (Wallace, 1939; Ouellet, 1993). The northern extent of their range appears to be the middle north shore of the Gulf of St. Lawrence (Wallace, 1939; Ouellet, 1993).

Descriptions of BITH's preferred habitat in the mountains of New York and New England vary little, and remark upon dense, nearly impenetrable thickets of balsam fir and red spruce (Wallace, 1939; Rimmer et al., 1996). Wallace (1939) noted BITH nesting in both dense fir-spruce thickets, and in "deeper woods." In the Green Mountains, BITH was associated with the uppermost elevations (> 1200 m) and predominated in areas of low canopy cover, low canopy height, low basal area of deciduous trees, and a high percentage of coniferous understory (Noon, 1981b). The results from this study agree closely with the others. This study found that the odds of a BITH presence observation rise with increasing elevation, basal area of 7.6 to 11.2 cm DBH softwood stems, distance to the nearest hardwood land cover type, live softwood sapling density, and percentage of live softwood basal area; and rise with declining dominant vegetation height, percentage of total canopy cover, basal area of softwood stems greater than 11.2 cm DBH, and total live basal area of all trees. On Mt.

Moosilauke, slope and aspect variables were found not to be important in explaining the distribution of subalpine birds including BITH (Sabo, 1980), a result confirmed by this study. There is no doubt about the preferred habitat of BITH described above, but they have been observed in the taller softwoods down to elevations of 762 m.

Habitat Modeling

BITH point count data were degraded to presence/absence by ignoring the count of BITH at each station and presences of more than one year. This resulted in stations with BITH presence being treated equally even though some stations had greater numbers of BITH, and multiple years of BITH presence. This was done to apply the data in a binary presence/absence format for use in the dichotomous logistic regression. The effects of ignoring the count were assumed to be minimal, because very few stations had greater than one individual per count, or more than one year of presence observations.

The goal of constructing a logistic multiple regression habitat model for BITH was to generate a spatially explicit map based on and derived from measured habitat variables important in explaining the distribution of this thrush. There was no *a priori* knowledge of the variables that would be included in the model, but because remote sensing studies routinely estimate the same suites of forest biometrics, it followed that if these variables were present in the model, they might be estimated over wide geographic extents. If estimates for all model variables were made, then the model could be applied at every location.

In the best fit model that included VEGHGT, ELEVATN, and BASW9-12 (eq. 4.1), it was determined that spectral features from ETM imagery could provide reliable estimates of VEGHGT, and the DEM could provide exceptionally reliable estimates of

ELEVATN. However, the ETM spectral features indicated no reliable way to estimate the BASW9-12 variable, and thus there was no way to apply this variable in a spatially explicit way. The VEGHGT variable was estimated from a multiple linear regression model that included the NDVI and Band1 spectral features (eq. 4.2). Spectral reflectance does not directly relate to VEGHGT, but represents the result of complex interaction between stand structure, composition, biomass, and geometry (Jakubauskas and Price, 1997). For example, stands with high dominant vegetation height may possess frequent shadows, high leaf-area-index (LAI), and prohibit light from reaching and reflecting from ground litter. Alternatively, stands with low dominant vegetation height tend toward high stem densities, cast fewer deep shadows, have lower LAI, and may allow sufficient radiation to pass through to the ground and reflect back into space. NDVI has been found to be highly correlated with LAI (Spanner et al., 1990; Jakubauskas and Price, 1997).

For the same reasons, one might expect to successfully estimate the basal area of 9-12 DBH softwood stems, and in some cases it may be possible. Here, this variable could not be resolved with the spectral features and ground data at the sampling stations. Pursuit of alternative data exploration methods or use of higher spectral and spatial resolution imagery could provide a means to estimate this variable. It is a difficult task to spectrally resolve stands of similar species composition into size classes, especially when stands possess mixtures of multiple classes. In this study, 22.9-30.5 cm DBH softwood trees were relatively rare and occurred mostly in the mixedwoods and toward the lower elevation limits of the softwoods. Even where they occurred there was much variability among the four sampling plots, and they were altogether absent from most

stations. Thus, a strong statistical relationship was not found, and the variable had to be dropped from the applied model.

The decision was made to include the variable representing the distance to the nearest Fir Sapling stand estimated from the land cover map, even though its contribution to the model was of minor significance. The Fir Sapling land cover type represents two structurally similar variants. The first variant is that of *kampfzone* vegetation, which occurs as the subtle transition from the softwoods at tree line to subalpine *krummholz*. It is characterized by high stem density, nearly pure stands of balsam fir lacking significant canopy cover. With increasing elevation, *kampfzone* vegetation transitions into the *krummholz* characterized by very short stature (less than breast height) balsam fir with twisted wind-worn stems and branches. The name Sapling is indicative of its size class and not its age, as many stems within the *kampfzone* are very old. The second variant of the Fir Sapling class represents the regenerating thickets of balsam fir found within the dead zones of fir waves (Sprugel, 1974). Fir waves occur throughout the WMNF, but are especially prevalent in the area of North and South Twin Mountains, Garfield Ridge, and Kinsman Ridge. Justification for including the Fir Sapling distance variable lies in its close correspondence as to the primary nesting habitat condition for BITH (Wallace, 1939; Ouellet, 1993; Atwood et al., 1996; Rimmer et al., 1996).

After applying the model at every pixel, the habitat map produced consisted of probability estimates for BITH presence. These probabilities reflect the BITH sampling data used to generate them, and specifically represent the probability of scoring at least one BITH presence from five 5-minute point-counts conducted in successive years from

1993-1997. It was expected that the resulting map would accurately estimate probabilities at the stations used in constructing the model. However, because estimates for the variables were generated from alternative sources (in the case of VEGHGT from spectral features rather than plot sampling), this had to be confirmed. The predictions were significantly different than observations at model calibration stations, and probabilities were systematically underestimated by a mean of 0.112. Elevation estimates from the DEM had a nearly 1:1 relationship with the DGPS elevation data, and estimates for DIST2FS were from the same data set for both calibration and application of the model. Comparison of VEGTHGT estimates from equation 4.2 versus those from plot sampling showed that the spectral data consistently overestimated VEGHGT, thereby causing the underestimation in probability. Adding the mean difference in probability to every pixel brought the predicted and observed probabilities at calibration stations into close agreement and represented the final model (eq. 4.5).

Atwood et al. (1996) presented a logistic multiple regression model for BITH presence throughout New York and New England that estimated greater probability of BITH presence with increasing elevation, amounts of softwood vegetation, latitude, and land area above 915 m elevation within 1km of the survey site. Sample size constraints prevented validation of this full model, but validation using elevation, amount of softwood vegetation, and index of latitude performed well. Atwood et al. (1996) did not generate a predictive map, but using GIS input coverages where available for their full model, a habitat map for BITH could be generated over extensive areas, and validated using independent data like those from the WMNF HEBMP.

Model Validation

Validation stations were used to independently assess the predictive capacity of the applied model. The model validated poorly at probabilities over 0.60 by overestimating the number of BITH presence observations. It is not clear why this happened in the 0.60 - <0.70 decile considering the model calibration data supported an increasing linear function including this range of probabilities, and the validations were not extrapolated greatly beyond ranges used for calibration. The height of the dominant vegetation was well within the range used for model calibration, but several stations were at higher elevations than used for calibration. Field notes from these stations suggested that the habitat was suitable for BITH, but BITH presence records from these stations were relatively few. BITH may become limited under conditions at high elevations where night time low temperatures even in summer can fall below freezing and wind speed rises, even though suitable vegetation characteristics persist. Even though BITH requires smaller metabolic response to low ambient temperatures in order to maintain homeostatic temperatures than other forest thrushes of the northeastern US (Holmes and Sawyer, 1975), climatic conditions with increasing elevation may directly impact the birds energy reserves, or indirectly impact them by limiting insect prey.

Probability estimates over 0.70 occur at pixels in krummholz and the alpine zone, where trees no longer persist and should be disregarded. The model calibration had no alpine zone representation, but running the model in these areas produced exceptionally high probabilities greater than 0.75. A crude stopgap measure was included in the model to resist this problem by arbitrarily setting all pixels with ≤ 0 VEGHGT to a fixed probability of 0.001. However, where krummholz persists up to

1525 m, the estimate of VEGHGT could be very low, but still greater than zero. Hence, very high, unrealistic probabilities result at pixels with covariate patterns consisting of elevations greater than 1370 m, dominant vegetation heights between 0 and 1 meter, and in or near the Fir Sapling cover type. Thus, pixels far above tree line have very high probability estimates even though it is known that BITH does not frequent these areas. These inaccuracies in the model represent less than 1.5% of the total study area.

Considering validation stations collectively across estimated deciles ranging from 0 - <0.60, model estimates were not significantly different than expected. Within the 0.40 - <0.50 decile alone, however, model estimates were significantly different than observed, and underestimated BITH expected proportions. It is unclear why this decile fit well at calibration stations, but failed at validation stations. There may be alternative variables not included in this study that were important in explaining BITH in areas exhibiting this range of probabilities.

There are several factors in the study design that suggested areas included in the 0.10 - <0.20 decile were greatly overestimated, even though it validated well with independent data. First, the layout of point count stations was not random with respect to habitat type. Stations were located along transects with the condition that transects terminated when entering areas characterized by significant fractions of hardwoods. Second, although validation stations were not included in model calibration, they did come from the same survey transects. Third, the survey transects follow the established trail system of the WMNF and at lower elevations these trails tend to follow cool stream valleys, which support softwood species at elevations well below normal. Fourth, the habitat model contained no regressors that distinguished hardwoods from softwoods, and

it is generally known and supported by the univariate analyses here that BITH does not occur in hardwood forest types. All factors taken together suggest at lower elevations (610-762 m) the model predicts well only in the cool stream valleys where softwoods extend downward. These softwood extensions represent a minor fraction of the total area from 610-762 m that are otherwise dominated by nearly pure hardwood stands composed of hardwood species, where the probability of detecting a BITH likely diminishes to near zero.

Implications of Habitat Modeling using Satellite Imagery

Predictive species-habitat maps represent a valuable tool for resource management and study (Dettmers and Bart, 1999). They can be used to identify areas where a species is likely found (i.e., suitable habitat), estimate the area of suitable habitat, and with some simplifying assumptions provide estimates of population size. Moreover, maps derived from habitat modeling offer the added advantage of simulating alternative scenario conditions like climate change, fire suppression, or forest management. Models and map products should undergo validation procedures so areas of unreliability can be identified and results evaluated accordingly. A map that does not validate well under certain conditions does not mean that the map is of little or no value. It is crucial, however, to know the conditions and locations where the map is appropriate for use.

Until now no study has used satellite imagery to estimate a variable included in a wildlife habitat model so that the model could be solved across large geographic extents. Here, probability estimates for BITH presence on 160,000 ha (400,000 acres) of the WMNF were generated and validated, providing estimates at over 2.2 million pixels

with 28.5 x 28.5 meter spatial resolution. No wildlife habitat mapping study has accomplished this level of detail of such a large geographic extent. Moreover, the ETM imagery used in this study includes full coverage of the Green Mountains in neighboring Vermont, which combined with the White Mountains represent the two greatest distribution centers for BITH (Wallace, 1939; Atwood et al., 1996). The model developed here could be readily applied to the Green Mountains and greatly increase the mapped range of BITH in New England. An independent model validation exercise would still be required to assure the model could be extrapolated beyond the WMNF.

This mapping study of BITH has demonstrated that satellite imagery can (1) produce accurate probability estimates of species presence; (2) provide cost efficient results in terms of cost (money and time) per unit area; (3) utilize habitat variables typically measured by ornithologists and forest managers, including those important in determining species distribution; and (4) be applied over large geographical extents. The use of satellite imagery for habitat modeling and mapping most closely reaches the goals set forth by Mosher et al. (1986), than any other mapping technique.

Literature Cited

- Anderson, S. H. and H. H. Shugart, Jr. 1974. Habitat selection of breeding birds in an east Tennessee deciduous forest. *Ecology* 55: 828-837.
- Anthony, R. G., G. A. Green, E. D. Forsman, and S. K. Nelson. 1996. Avian abundance in riparian zones of three forest types in the Cascade Mountains, Oregon. *Wilson Bull.* 108(2): 280-291.
- Aspinall, R. and N. Veitch. 1993. Habitat mapping from satellite imagery and wildlife survey data using a Bayesian modeling procedure in a GIS. *Photogrammetric Engineering & Remote Sensing* 59(4): 537-543.
- Atwood, J. L., C. C. Rimmer, K. P. McFarland, S. H. Tsai, and L. R. Nagy. 1996. Distribution of Bicknell's Thrush in New England and New York. *Wilson Bull.* 108(4): 650-661.

- Capen, D. E., J. W. Fenwick, D. B. Inkley, and A. C. Boynton. 1986. Multivariate models of songbird habitat in New England forests. Pp. 171-175. in *Wildlife 2000: Modeling Habitat Relationships of Terrestrial Vertebrates* (Verner, J., M. L. Morrison, and C. J. Ralph, eds.). University of Wisconsin Press.
- Cohen, W. B. and T. A. Spies. 1992. Estimating structural attributes of Douglas Fir/western hemlock forest stands from Landsat and SPOT imagery. *Remote Sensing of Environment* 41: 1-17.
- Dettmers, R. and J. Bart. 1999. A GIS modeling method applied to predicting forest songbird habitat. *Ecological Applications* 9(1): 152-163.
- Forman, R. T. T. and M. Godron. 1986. *Landscape Ecology*. John Wiley. New York. pp. 619.
- Franklin, S. E. 1994. Discrimination of subalpine forest species and canopy density using digital CASI, SPOT PLA, and Landsat TM data. *Photogrammetric Engineering & Remote Sensing* 60(10): 1233-1241.
- Franklin, J. and D. W. Steadman. 1991. The potential for conservation of Polynesian birds through habitat mapping and species translocation. *Conservation Biology* 5(4): 506-521.
- Franzreb, K. E. 1983. A comparison of avian foraging behavior in unlogged and logged mixed-coniferous forest. *Wilson Bull.* 95(1): 60-76.
- GLOBE Program. 1997. *The GLOBE Teacher's Guide*. Washington D. C.
- Green, K. M., J. F. Lynch, J. Sircar, and L. S. Z. Greenberg. 1987. Landsat remote sensing to assess habitat for migratory birds in the Yucatan Peninsula, Mexico. *Vida Silvestre Neotropical* 1(2): 457-471.
- Hall, F. G., Y. E. Shimabukuro, and K. F. Huemmrich. 1995. Remote sensing of forest biophysical structure using mixture decomposition and geometric reflectance models. *Ecological Applications* 5(4): 993-1013.
- Hepinstall, J.A., & S.A. Sader. 1997. Using Bayesian statistics, Thematic Mapper satellite imagery, and Breeding Bird Survey data to model bird species probability of occurrence in Maine. *Photogrammetric Engineering & Remote Sensing* 63(10): 1231-1237.
- Herr, A. M. and L.P. Queen. 1993. Crane Habitat Evaluation Using GIS and Remote Sensing. *Photogrammetric Engineering & Remote Sensing* 59(10): 1531-1538.

- Holmes, R. T. and R. H. Sawyer. 1975. Oxygen consumption in relation to ambient temperatures in five species of forest-dwelling thrushes (*Hylocichia* and *Catharus*). *Comp. Biochem. Physiol.* 50A: 527-531.
- Hosmer, D. W. and S. Lemeshow. 1989. *Applied Logistic Regression*. John Wiley. New York. pp. 307.
- Jakubauskas, M. E. 1996. Canonical correlation analysis of coniferous forest spectral and biotic relations. *Int. J. Remote Sensing* 17(12): 2323-2332.
- Jakubauskas, M. E. and K. P. Price. 1997. Empirical relationships between structural and spectral factors of Yellowstone Lodgepole pine forests. *Photogrammetric Engineering & Remote Sensing* 63(12): 1375-1381.
- James, F. C. 1971. Ordinations of habitat relationships among breeding birds. *Wilson Bulletin* 83(3): 215-236.
- James, F. C. and H. H. Shugart, Jr. 1970. A quantitative method of habitat description. *Audubon Field Notes* 24(6): 727-736.
- Mosher, J. A., K. Titus, and M. R. Fuller. 1986. Multivariate models of songbird habitat in New England forests. Pp. 171-175. *in Wildlife 2000: Modeling Habitat Relationships of Terrestrial Vertebrates* (Verner, J., M. L. Morrison, and C. J. Ralph, eds.). University of Wisconsin Press.
- Noon, B. R. 1981a. Techniques for sampling avian habitats. Pp. 42-52. *in The Use of Multivariate Statistics in Studies of Wildlife Habitat* (Capen, D. E., ed.). United States Forest Service General Technical Report RM-87.
- Noon, B. R. 1981b. The distribution of an avian guild along a temperate elevation gradient: The importance and expression of competition. *Ecological Monographs* 51(1): 105-124.
- Ouellet, H. 1993. Bicknell's Thrush: Taxonomic status and distribution. *Wilson Bull.* 105(4): 545-574.
- Palmeirim, J. M. 1988. Automatic mapping of avian species habitat using satellite imagery. *Oikos* 52:59-68.
- Petit, L. J. and D. R. Petit. 1996. Factors governing habitat selection by Prothonotary Warblers: Field tests of the Fretwell-Lucas models. *Ecological Monographs* 66(3): 367-387.
- Rimmer, C. C., J. L. Atwood, K. P. McFarland, and L. R. Nagy. 1996. Population density, vocal behavior, and recommended survey methods for Bicknell's Thrush. *Wilson Bulletin* 108(4): 639-649.

- Sabo, S. R. 1980. Niche and habitat relations in subalpine bird communities of the White Mountains of New Hampshire. *Ecological Monographs* 50(2): 241-259.
- Sader, S. A., R. B. Waide, W. T. Lawrence, and A. T. Joyce. 1990. Tropical forest biomass and successional age class relationships to a vegetation index derived from Landsat TM data. *Remote Sensing of Environment* 28: 143-156.
- Spanner, M. A., L. L. Pierce, D. L. Peterson, and S. W. Running. 1990. Remote sensing of temperate coniferous leaf area index: The influence of canopy closure, understory vegetation, and background reflectance. *International Journal of Remote Sensing* 11(1): 95-111.
- Sprugel, D. G. 1974. Dynamic structure of wave-regenerated *Abies balsamea* forests in the North-eastern United States. *Journal of Ecology* 64: 889-911.
- Vogelmann, J. E. and B. N. Rock. 1988. Assessing forest damage in high-elevation coniferous forests in Vermont and New Hampshire using Thematic Mapper data. *Remote Sensing of Environment* 24: 227-246.
- Wallace, G. J. 1939. Bicknell's Thrush, its taxonomy, distribution, and life history. *Proc. Boston Soc. Nat. Hist.* 41(6): 211-402.
- Zar, J. H. 1999. *Biostatistical Analysis*. 4th edition. Prentice Hall. Upper Saddle River, NJ. pp. 929.

Code	Description	Units	Mean	Range	Source
ASPECT	Azimuthal aspect orientation	degrees	-	N,E,S,W	DEM
BASW3-6	Basal area of live softwood trees 3-6" DBH	m ² /ha	10.0	0.656 - 47.9	PS
BASW6-9	Basal area of live softwood trees 6-9" DBH	m ² /ha	6.83	0 - 22.7	PS
BASW9-12	Basal area of live softwood trees 9-12" DBH	m ² /ha	3.32	0 - 21.4	PS
BASW12+	Basal area of live softwood trees >12" DBH	m ² /ha	1.38	0 - 14.5	PS
DIST2FS	Distance to nearest Fir Shrub cover type	m	43.6	0 - 287	LCM
DIST2HW	Distance to nearest Hardwoods cover type	m	162	0 - 655	LCM
DIST2KH	Distance to nearest Krummholz cover type	m	2720	0 - 13200	LCM
DIST2MW	Distance to nearest Mixedwoods cover type	m	29.6	0 - 308	LCM
DIST2OR	Distance to nearest Open Rock cover type	m	823	1.7 - 311	LCM
DIST2SW	Distance to nearest Softwoods cover type	m	18.4	0 - 335	LCM
DIST2WT	Distance to nearest Water cover type	m	122	1.8 - 268	LCM
ELEVATN	Elevation above mean sea level	m	1130	740 - 1438	DGPS
LSRICH3	Number of cover types in 3x3 cell window	covers	2.40	1 - 5	LCM
LSRICH7	Number of cover types in 7x7 cell window	covers	3.45	1 - 5	LCM
LVDDBAR	Live to Dead wood basal area ratio	unitless	7.73	0.45 - 134	PS
PCLSWBA	% live basal area from softwood trees	%	77.7	1.51 - 100	PS
PCLSWCC	% live canopy cover from softwood trees	%	58.6	0 - 100	PS
PCTOTCC	% canopy cover by live trees	%	51.3	0 - 95	PS
PTCAREA	Land cover patch area	ha	518	0.081 - 7000	LCM
SHDLVSW	Density of Live softwood shrubs	stems/m ²	0.219	0 - 1.66	PS
SLPDEGR	Slope in degrees	degrees	14.8	3 - 37	DEM
TOTBADD	Total basal area of all dead stems	m ² /ha	8.75	0 - 42.9	PS
TOTBALV	Total basal area of all live trees	m ² /ha	28.4	2.62 - 53.1	PS
VEGTHGT	Height of dominant tree canopy	meters	9.55	2.13 - 19.3	PS

Table 4.1. Table of logistic regressor variables. Table includes variable abbreviations, descriptions, units, mean, range, and data source. Abbreviations for Source are DEM- Digital Elevation Model, DGPS- Differential GPS, LCM- Land cover map, PS- Plot sampling. No mean is given for aspect, because it is a circular scale variable.

Variable	Coefficient	Standard Error	Deviance	P
VEGTHGT	-0.291	0.0574	31.7	< 0.0001
ELEVATN	0.00529	0.00117	23.7	< 0.0001
PCTOTCC	-0.0246	0.00609	17.7	< 0.0001
BASW9-12	-0.203	.0555	17.7	< 0.0001
DIST2KH	-0.000152	0.00000482	11.9	0.0006
BASW3-6	.0802	.0249	11.9	0.0006
TOTBALV	-0.0495	.0153	11.2	0.0008
DIST2FS	-0.00100	0.000350	10.4	0.0012
DIST2HW	.000246	0.0000920	7.32	0.0068
SHDLVSW	1.78	0.699	7.27	0.0070
PCLSWBA	0.0185	0.00745	6.89	0.0086
BASW6-9	-0.0764	.0346	5.21	0.022
DIST2WT	0.000400	0.000241	2.71	0.10

Table 4.2. Table of results from univariate logistic regression using the continuous variables. Analyses were performed with Bicknell's Thrush presence/absence as the binary response variable. Logit coefficients and standard errors are presented, along with the model Deviance (D), which is chi-square distributed with one degree of freedom for continuous variables. Variables are ranked by P-values, where $P = X^2_{(df=1, D)}$. Variables with P-values < 0.01 were considered candidates in the multiple logistic regression model. Reported values are to 3 significant digits.

Variable	Double Odds Scale	95%CI low	95%CI high
VEGTHGT	-2.38 m	1.53	2.61
ELEVATN	131 m	1.48	2.70
PCTOTCC	-28.2 %	1.43	2.80
BASW9-12	-3.42 m ² /ha	1.38	2.90
DIST2KH	-424 m	1.92	2.09
BASW3-6	8.64 m ² /ha	1.31	3.05
TOTBALV	-14.0 m ² /ha	1.31	3.04
DIST2FS	-64.2 m	1.24	3.22
DIST2HW	262 m	1.20	3.33
SHDLVSW	0.389 stems/m ²	1.17	3.41
PCLSWBA	37.4 %	1.16	3.46
BASW6-9	-9.08 m ² /ha	1.08	3.70
DIST2WT	162 m	0.874	4.58

Table 4.3. Table of variable scales that result in the doubling of the odds of a Bicknell's Thrush (BITH) present result. Double Odds Scale represents the amount of change in the regressor variable that doubles the odds of detection of a BITH. For example, every decrease of 2.38 m in dominant vegetation height doubles the odds of a BITH being present at least once in 4 survey visits. Likewise, the odds double for every 131 meter increase in elevation. The 95% confidence intervals (CI) bound a 2:1 odds-ratio for each variable.

Source	Coefficient	Standard Error	Deviance	P-value
VEGHGT	-0.156	0.0721	30.4	< 0.0001
ELEVATN	0.00286	0.00138	4.08	0.043
BASW9-12	-0.104	587	3.49	0.062
Intercept	-2.01	1.98		

Table 4.4. Table of step-wise Logistic Multiple Regression. A variable entered the model if its P-value was less than 0.10. The model represents the best fit of the 25 regressor variables to the binary outcome variable of Bicknell's Thrush presence/absence. Abbreviations are VEGHGT-dominant vegetation height, ELEVATN- elevation, and BASW9-12- basal area of 22.9-30.5 cm DBH softwood stems.

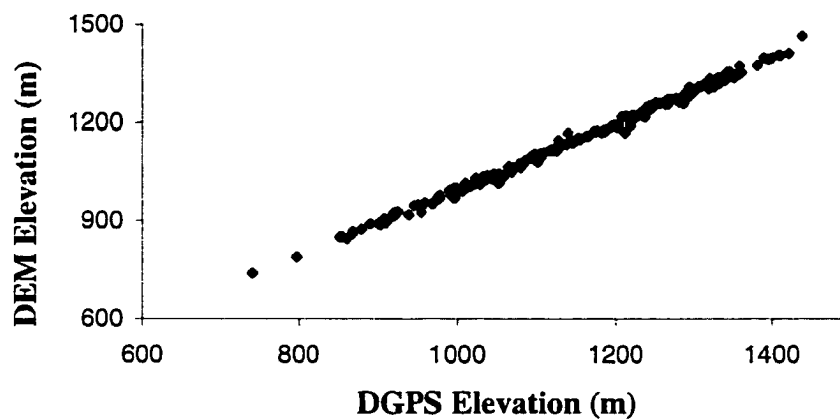


Figure 4.1. Scatterplot of elevation (meters) derived from a USGS Level 1 Digital Elevation Model (DEM) and from Differential GPS (DGPS). The equation for the fitted least-squares regression line through the data points is $y = 1.007x - 14.074$ ($R^2 = 0.996$).

Source	df	Sum of Sq	Mean Sq	F	P-value
NDVI	1	200.48	200.48	27.26	< 0.0001
B1	1	39.21	39.21	5.33	0.029
Residuals	29	213.28	7.35		

F-statistic: 16.3 on 2 and 29 degrees of freedom, the P-value is 0.000018
Adjusted Multiple R² = 0.50
Adjusted Multiple R = 0.71

Table 4.5. Analysis of Variance Table for multiple regression of Dominant Vegetation Height (VEGHGT) as the response variable and the ETM spectral features Normalized Difference Vegetation Index (NDVI) and Band1 (B1). Data points included thirty-two stations exhibiting Coefficients of Variation less than 10%. The regression model was $VEGHGT = 0.2107 * NDVI - 1.0882 * Band1 - 15.82$.

Source	Coefficient	Standard Error	Deviance	P-value
VEGHGT	-0.188	0.0688	30.4	< 0.0001
ELEVATN	0.00270	0.00138	4.08	0.043
DIST2FS	-0.00157	0.00125	1.74	0.18
Intercept	-1.64	1.96		

Table 4.6. Table of Logistic Multiple Regression for the applied Bicknell's Thrush (BITH) habitat model. The P-value criterion of 0.10 was relaxed to permit a regressor variable to enter that did not correlate with elevation. Abbreviations are VEGHGT- dominant vegetation height, ELEVATN- elevation, and DIST2FS- distance to Fir Sapling cover type.

No. of Stations	Range	Estimated	Observed	Difference
22	0 - < 0.1	0.05	0.04	+0.01
39	0.1 - < 0.2	0.15	0.26	-0.11
53	0.2 - < 0.3	0.25	0.38	-0.13
43	0.3 - < 0.4	0.35	0.40	-0.05
22	0.4 - < 0.5	0.45	0.68	-0.23
9	0.5 - < 0.6	0.55	0.73	-0.18
4	0.6 - < 0.7	0.65	0.75	-0.10

Table 4.7. Table of estimated and observed probabilities from application of the logit $g(x) = -VEGHGT * 0.1885 + ELEVATN * 0.00270 - DIST2FS * 0.00157 - 1.644$. The estimated probabilities were systematically underestimated by an average of 0.112. Over the entire range of deciles, the estimated decile was significantly different than observed ($X^2_{(df=5)} = 14.4$; $P = 0.04$).

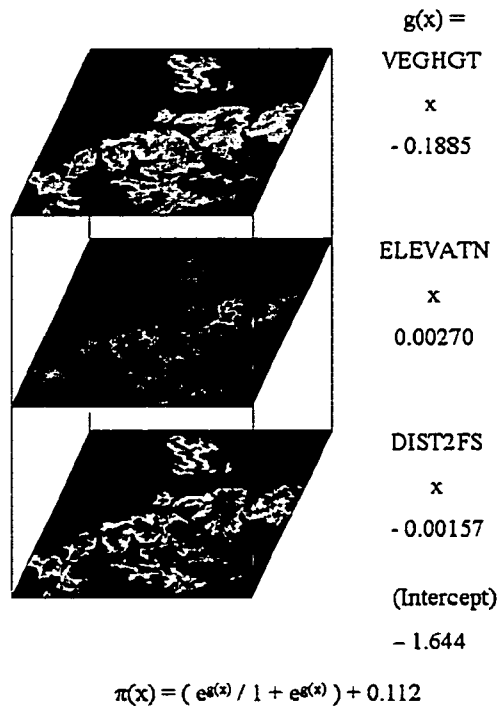


Figure 4.2. Diagrammatic representation of the solution for the logit $g(x)$ from grayscale image data layers. Coregistered pixels within each layer were substituted into the logit to provide a solution at every pixel. The logit is then used to solve for $\pi(x)$, the probability of a BITH presence in at least 1 of 5 years. An adjustment of 0.112 is added in this case to account for an overestimation of dominant vegetation height (VEGHGT). In all variable-image layers, lighter tones represent higher variable values. ELEVATN = elevation and DIST2FS = planimetric distance to nearest Fir Sapling Cover type.

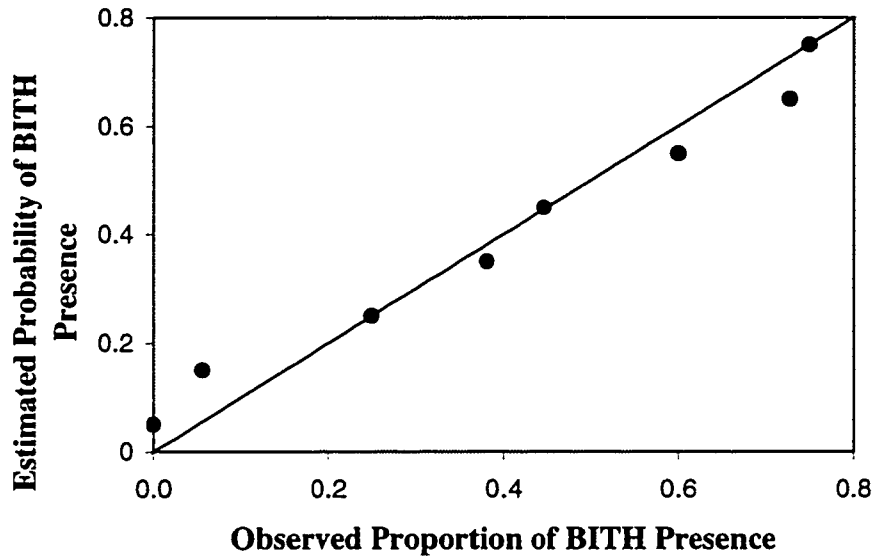


Figure 4.3. Scatterplot of the fit of estimated probability and the observed proportion of Bicknell's Thrush (BITH) presence at model calibration stations. Probabilities were calculated from $\pi(x) = e^{g(x)} / e^{1 + g(x)} + 0.112$, where $g(x) = -\text{VEGHGT} * 0.1885 + \text{ELEVATN} * 0.00270 - \text{DIST2FS} * 0.00157 - 1.644$. The solid black line represents the line of equal prediction.

Decile Midpoint	N	Observed	Predicted	p
0.05	2	0	0.1	0.80
0.15	15	2	2.6	0.69
0.25	61	13	15.5	0.46
0.35	68	28	24.0	0.31
0.45	62	38	27.4	0.0007
0.55	50	23	27.4	0.21
0.65	15	5	9.6	0.014
0.75	6	1	4.4	0.001

Table 4.8. Table of observed and predicted numbers of Bicknell's Thrush (BITH) presence at model validation stations. Stations were grouped into deciles of probability. Deciles with midpoints 0.45, 0.65, and 0.75 deviated significantly (critical alpha = 0.05) in their predictions of BITH presence.

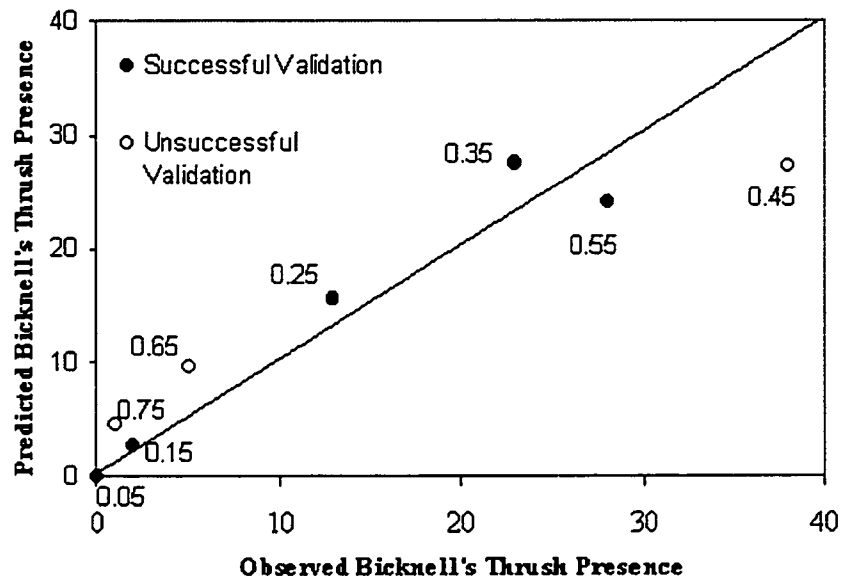


Figure 4.4. Scatterplot of the predicted count of Bicknell's Thrush (BITH) presence and observed count of BITH presence at model validation stations. The solid black line represents the line of equal prediction. Data point labels indicate the plotted probability decile. Validation success was determined by failure to reject hypothesis that predicted and observed presence counts were equal using a critical alpha of 0.05.

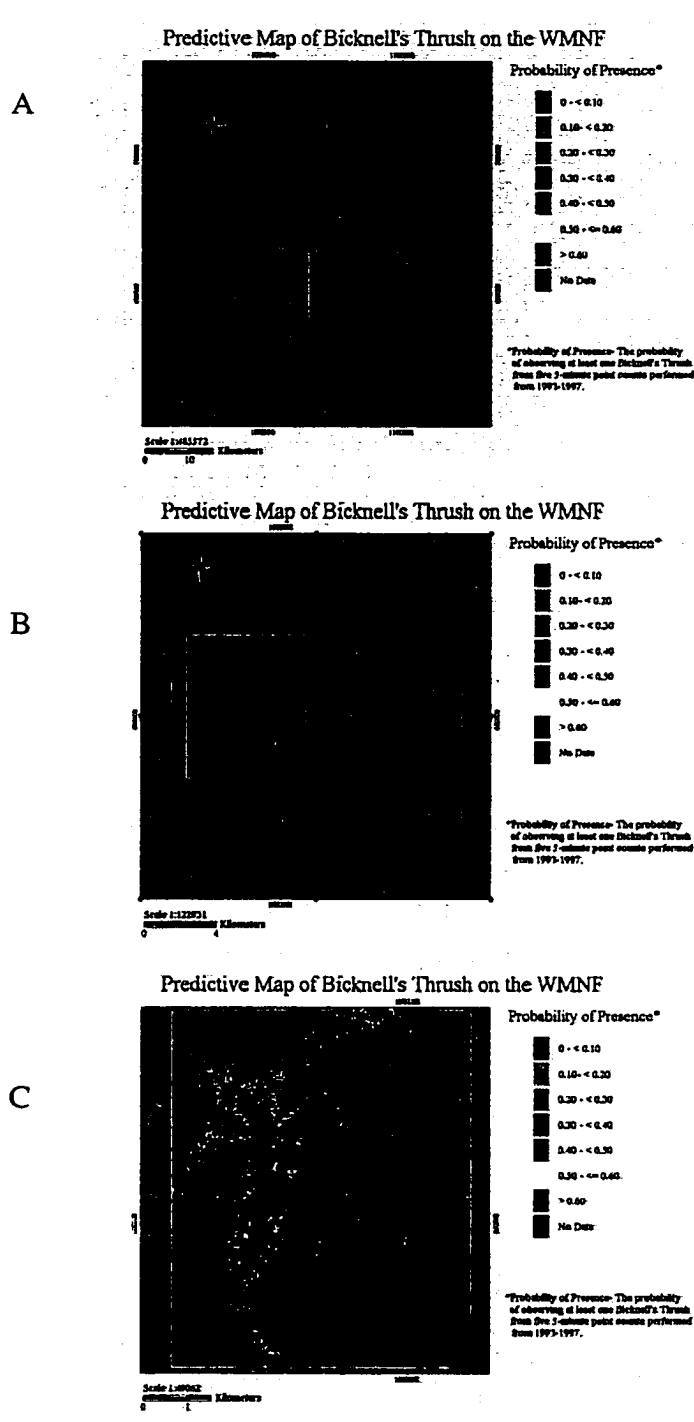


Figure 4.5. Predictive Maps of Bicknell's Thrush Distribution at 3 scales. Panel A) Full extent of the White Mountain National Forest; Panel B) Inset from panel A. Pemigewasset Wilderness Area. Panel C) Inset from panel B. Franconia Ridge, and Owl's Head toward the lower right. The black island polygon includes Mt. Lafayette, which is masked out of the study because it is over 5000 ft. in elevation. Locations in red are probabilities significantly greater than observed with validation data.

CHAPTER 5

ESTIMATING BICKNELL'S THRUSH (*Catharus bicknelli*) DISTRIBUTION, DENSITY, AND ABUNDANCE FROM A HABITAT MODEL DERIVED FROM SATELLITE IMAGERY

Introduction

Wildlife habitat studies that use geospatial technologies (i.e., Global Positioning System (GPS), geographic information systems (GIS) and remote sensing) have increased rapidly over the past decade. It is now commonplace for studies of birds for instance to include precise location information using GPS, and spatially explicit data layers as input into GIS databases. However, satellite imagery data have been utilized to a far lesser extent, and the full potential for satellite imagery in ornithological applications has gone untested and therefore unproven.

Two basic strategies for digital processing of satellite imagery can provide information about bird habitat. In one strategy, image pixels are sorted into discrete classes based upon their spectral similarity to areas of known land cover type (e.g., deciduous forest, fields, water bodies), or based upon their similarity in forming discrete spectral clusters that can be matched to known land cover types. Pixels coded according to class membership, produce a thematic map that can be used as input in a GIS. This process, termed image classification, represents the most common approach to using satellite imagery in ornithological studies. In the second strategy, digital Brightness

Values (BVs) for spectral features at every pixel are used to estimate the condition or value of some variable within the pixel. For example, if vegetative biomass is associated with the amount of spectral reflectance, then vegetative biomass may be estimated or predicted from spectral reflectance.

Almost all applications of satellite imagery in ornithology have used land cover maps derived from Landsat-5 Thematic Mapper (TM) or SPOT (Système Probaatoire d'Observation de la Terre) data. Land cover maps were used to map the habitats of Hooded Warbler (*Wilsonia citrina*) in the Yucatan (Green et al., 1987); White-breasted Nuthatch (*Sitta carolinensis*), Red-eyed Vireo (*Vireo olivaceus*), and Indigo Bunting (*Passerina cyanea*) in Kansas (Palmeirim, 1988); birds of the Cook Islands (Franklin and Steadman, 1991); Sandhill Cranes (*Grus canadensis*) in Minnesota (Herr and Queen, 1993); Stone Curlews (*Burhinus oedicnemus*) in Britain (Green and Griffiths, 1994); and to inventory Harlequin Ducks (*Histrionicus histrionicus*) in northern Quebec (Thibault et al., 1998). Adelle Penguin (*Pygoscelis adeliae*) colonies were mapped based on the spectral contrast between ice and their guano deposits (Schwaller et al., 1989). The second strategy has been used in combination with a Bayesian statistical approach, such that spectral BVs from locations with known species presence and absence were identified and used to predict locations where species were likely found (Aspinall and Veitch, 1993; Hepinstall and Sader, 1997). In the absence of ground data, however, this type of Bayesian approach does not include any direct linkage to habitat data.

Spectral BVs have been used by remote sensing analysts to estimate the same types of forest stand parameters used by ornithologists in habitat modeling, including basal area, stem density, canopy closure, species composition, vegetation height, seral

stage, forest damage, biomass and leaf-area-index (Vogelmann and Rock, 1988; Sader et al., 1990; Spanner et al., 1990; Cohen and Spies, 1992; Franklin, 1994; Hall et al., 1995; Jakubauskas, 1996; Jakubauskas and Price, 1997). In Chapter 4, macrohabitat and landscape characteristics were integrated to construct and validate a Bicknell's Thrush (BITH; *Catharus bicknelli*) habitat model using logistic multiple regression and Landsat-7 Enhanced Thematic Mapper (ETM) imagery. Modeling results produced dominant vegetation height (VEGHGT) as one of three variables important in explaining the presence/absence of BITH on the White Mountain National Forest (WMNF) of New Hampshire. Spatially explicit data for VEGHGT were not available, so ETM BVs were used to estimate VEGHGT at every pixel within the study area. Combined with a digital elevation model (DEM) and land cover map, the logistic habitat model was solved at every pixel, which resulted in a predictive, probabilistic map of BITH presence at approximately 2.2 million 28.5 x 28.5 m. pixels within the 162,000 ha study area.

The purpose of this study was to examine the output from a satellite imagery derived and deployed BITH habitat model for the WMNF (Chapter 4). The BITH habitat model was used to estimate areas of habitat suitability, BITH relative density, and BITH relative abundance. Model results predicted greater relative abundance of BITH in areas with lower relative densities, because the amount of area supporting lower relative densities was much greater than in areas supporting greater relative densities. This occurred because of the dependence of preferred BITH habitat on increasing elevations and the reduction of land area with increasing elevation. Estimates of relative density for areas considered the most suitable habitat (i.e., had the greatest likelihood of a BITH

presence) were approximately the same as densities derived from spot mapping within high BITH density habitat on Mt. Mansfield, Vermont.

Methods

About the BITH Habitat Model

Spatially explicit output from the logistic multiple regression BITH habitat model developed in Chapter 4 were used to quantify areas of BITH presence on the WMNF. The probability of BITH presence increased with decreasing VEGHGT, increasing elevation (ELEVATN), and decreasing distance to the nearest Fir Sapling cover type. Vegetation composition and structure of the Fir Sapling cover type match descriptions of BITH nesting habitat (Wallace, 1939; Rimmer et al., 1996). Modeled output consisted of probability estimates for observing BITH at every 28.5 x 28.5 m pixel within an elevation range of 610-1525 m. Modeled probability estimates represent the likelihood of scoring at least one BITH presence from five 5-minute point counts conducted annually (in mid-June) from 1993-1997. Point count records were dominated by song and call detections, and therefore may more accurately reflect the number of singing males. Note, however, Wallace (1939) reported that female BITH sing frequently while on the nest.

Validation of the habitat model using point count stations not included in the model construction produced agreement between observed and estimated numbers of BITH presences across the four probability deciles ranging from 0 to < 0.40 and the decile ranging from 0.50 to <0.60. The probability decile from 0.40 to <0.50 significantly underestimated the observed number of validation stations with BITH presence and the decile from 0.60 to 0.70 significantly overestimated the number of validation stations with BITH presence. Estimated probabilities greater than 0.70

resulted at locations outside of the ranges of model variables and represent an inappropriate extrapolation of the model. Values above 0.70 were estimated at pixels with both ELEVATN greater than 1220 m and VEGHGT less than 2 m. These pixels occur above tree line in areas of krummholz and alpine tundra, where the likelihood of BITH observation falls to zero. Because all pixels with probabilities greater than 0.60 were either underestimated or beyond the range of variables for which BITH naturally occur, these pixels (accounting for 3.0% of the entire study area) were not considered in the analysis. Although the 0.10 to <0.20 decile validate well with independent survey station data, there were factors in the study design that suggested this decile greatly overestimated the probability of BITH presence when applied to areas comprised of predominantly hardwood forests (Chapter 4). The interested reader should see Chapter 4 for full treatment and discussion of the BITH model.

Revised Study Area

As indicated above, the area estimated by the model for the 0.10 to <0.20 decile was considered greatly overestimated due to inclusion of vast hardwood stands not accounted for in model calibration and validation. To estimate a more appropriate area for this decile, a land cover map that included hardwoods was used (Chapters 1 and 2), so that Hardwood pixels within this decile were removed from this analysis. Elimination of Hardwoods pixels from the 0.10 to <0.20 decile reduced the study area extent from 160,000 ha. to approximately 91,000 ha.

Index of Relative Abundance

The area of each probability decile (e.g., 0 to <0.10, 0.10 to <0.20, etc.) was calculated by summing the number of pixels within each decile and multiplying by the

area of one pixel (0.0812 ha.). All areas are planimetric rather than over ground. The midpoint value of each probability decile was treated as an index of relative abundance (IRA). In the case of the 0.20 to <0.30 decile, the IRA was 0.25. An IRA of 0.25 suggests that 1 in 4 pixels within this decile would produce at least one BITH presence from the five annual point counts.

Estimation of Bicknell's Thrush Abundance and Density

To estimate the relative abundance and relative density of BITH, probability estimates of at least one BITH presence record in 5 years, $P(\text{BITH})$, where $P(\text{BITH}) = \text{IRA}$, were transformed to a probability estimate of a BITH presence in 1 year, $p(\text{BITH})$, using relationships for binomially distributed variables. Assumptions for this transformation include that within probability deciles and between years: surveys were independent; population size did not vary, and BITH detection by trained observers did not vary. The probability of at least one BITH presence in n trials is given as $P(\geq 1\text{BITH}) = P(1\text{BITH}) + P(2\text{BITH}) + \dots + P(n\text{BITH})$. Equally, the probability of at least one BITH presence in n trials is given as

$$P(\geq 1\text{BITH}) = 1 - P(\text{noBITH}), \quad (\text{eq. 5.1})$$

where $P(\text{noBITH})$ equals the probability of no BITH presence observations in n trials. $P(\text{noBITH})$ can be found as $(n!/(n-r)!)$ multiplied by $p(\text{BITH})^0$ multiplied by $(1 - p(\text{BITH}))^n$, where r equals the number of BITH presences. When solving for $P(\text{noBITH})$, $r = 0$ and the first two terms in the latter equation equate to 1. The equation can then be simplified to $P(\text{noBITH}) = (1 - p(\text{BITH}))^n$, and substituted into equation 5.1 to yield $P(\geq 1\text{BITH}) = 1 - (1 - p(\text{BITH}))^n$. A lookup table was constructed that calculates $P(\geq 1\text{BITH})$ from values of $p(\text{BITH})$ starting at 0.001 upward to 1.00 at 0.001 intervals,

and $n = 5$. The \overline{IRA} from model output $P(\geq i | \overline{BITH})$ was then used to look up the probability of a BITH presence $p(\overline{BITH})$ for one trial (year).

To estimate relative abundance (i.e., the number of BITH presences in one year), the original decile groupings were maintained so that the relative number of BITH presences was estimated as $p(\overline{BITH})$ multiplied by the number of pixels within the decile corresponding to $p(\overline{BITH})$. In the case of the decile from 0.50 to < 0.60 , the \overline{IRA} was 0.55 and its corresponding $p(\overline{BITH})$ was 0.147. Density estimates resulted from the estimate of BITH abundance per unit area.

It is important to note that the relative abundance and densities were solved for on a per pixel basis, and only indirectly reflect the BITH sampling data acquired over an area of a 3x3 pixel window. That is, the habitat model and covariate pattern of model regressors were used to estimate the probability of BITH presence within each pixel. However, because the model was derived from binary BITH presence/absence and averaged habitat data from a sampling unit of a 50-meter radius circle, or approximately a 3x3 pixel window, these sampling units do not agree.

Results

The distribution of BITH presence probability deciles, area, BITH abundance, and BITH density estimates are presented in Tables 5.1-5.2. The probability decile 0.50 to < 0.60 included pixels whose pattern of model covariates represents the greatest validated probability of scoring a BITH presence. The combined area of this decile was approximately 2800 ha, and accounted for 3.0% of the approximately 91,000 ha within the revised study area (i.e., excluding Hardwood forest pixels within the 0.10 to < 0.20 decile). Except for the 0 to < 0.10 and 0.10 to < 0.20 probability deciles, the amount of

land area within each decile decreases with increasing probability of BITH presence (Table 5.1). The 0 to <0.10 decile consists of areas devoid of vegetation such as wind and ice eroded rocky outcrops, slides and water bodies, where estimated VEGHGT is zero. The 0.10 to <0.20 decile was adjusted from about 100,000 ha to about 31,000 ha, where this decile more accurately represents areas of cool stream valleys that support softwood vegetation to elevations much lower than normal. All other deciles indicate greater probabilities of BITH presence within zones of decreasing area. This result can be attributed to the strong association between BITH presence and elevation, and the concomitant decrease in land area with increasing elevation.

The effect from reduction of area with increasing elevation resulted in lower estimated relative abundances of BITH in zones of greatest IRA than in zones of lower IRA (Table 5.1). The model estimated nearly 13 times more area within the 0.20 to <0.30 decile than within the 0.50 to <0.60 decile, which resulted in 4.8 times more BITH presences within the 0.20 to <0.30 decile. Combining estimated abundances across all deciles produced a total estimate of approximately 67,000 singing male BITH for the entire study area. Density estimates ranged from 4.9 to 72.4 BITH singing males per 40 ha , and increased with increasing decile (Table 5.2).

Discussion

Findings derived from the habitat model suggested habitats with relatively low densities of BITH accounted for such large areas compared to higher density habitats that the abundance of individuals within low density habitats may actually exceed the abundance from habitats with higher densities. Confirming the pattern of decreasing density with decreasing elevation, BITH density estimates from territory mapping ranged

from 36-65 breeding pairs per 40 ha (bp/40ha) at upper elevation sites, and 4-21 bp/40ha at two lower elevation sites on Mt. Mansfield in the Green Mountains of Vermont, (Rimmer and McFarland, 1995; Rimmer et al., 1996). Moreover, the BITH habitat model estimated the relative density for the 0.50 - <0.60 probability decile at 72.4 singing males per 40 ha, which assuming a singing male represents a breeding male on territory closely approximates the density at upper elevation sites from Mt. Mansfield.

It was encouraging to find model estimates of relative density for the highest decile of BITH in such close agreement with values found in the field for Mt. Mansfield. Assuming that habitats within the highest probability decile were saturated with BITH territories, and that, on average, territories approximated the size and shape of the 50-meter radius sampling circle, then the expected likelihood of any pixel containing a territory would be 1 in 9 = 0.111, where 9 equals the number of pixels contained within the sampling circle. In comparison, the modeled probability of a BITH in a given pixel within the highest decile was 0.147. Of further interest was to use the midpoint of the range of territory densities, 50.5 bp/40ha within high density habitat on Mt. Mansfield (Rimmer et al., 1996). At this density, the expected mean territory size would equal 0.79 ha, and this value very closely approximates the 0.75 ha area of the 50-meter radius sampling circle. Thus, the relative density estimates for the 0.50 to <0.60 decile from the model may reflect actual densities estimates obtained from territory mapping. If this is so, the model estimates approximately 5000 breeding pairs within highest density habitat for the WMNF.

Further research to calibrate the other probability deciles with on-the-ground estimates of breeding density could potentially provide robust estimates of population

size. A logical and practical approach would be to apply the model to the Green Mountains, where active research is ongoing on population level reproductive and demographic parameters (Chris Rimmer, personal communication). The Green Mountains are included within the ETM image used to derive and apply the BITH habitat model and the model could readily be applied to those sites.

The modeled result that more individuals occur within more extensive, lower density habitats carries important implications for the conservation and management of this species. The model predicted that 44% of singing males were located within lower elevation habitats of lower individual densities and 7.5% within habitats supporting the greatest densities. If the single inference were made that habitats supporting the greatest densities of BITH were most important in maximizing population size, then conservation efforts aimed toward protecting habitats supporting the greatest densities of individuals would be guided toward the smallest fraction of the total population.

It is often implied that habitats hosting greater densities of individuals of a species are of greater habitat quality for supporting their overall resource requirements, including reproduction. However, density by itself has been demonstrated as an unreliable predictor of habitat quality (Van Horne, 1983; Vickery et al., 1992). By itself, individual density carries no ancillary information pertaining to reproductive output (e.g., clutch size, fledge rates, or survivorship). Therefore, only tenuous conclusions about habitat importance or quality can be derived from density estimates.

Of fundamental importance to stability of populations is the distribution and size of population sources and sinks (Pulliam, 1988; Pulliam and Danielson, 1991; Robinson, 1992). Habitat sinks result when individual mortality within a habitat outpaces

reproduction. Identification of population sources is essential for guiding conservation efforts toward the maintenance of stable population sizes. Assume that all habitats (i.e., probability deciles) provide equal fitness potential in terms of each breeding pair's ability to reproduce, and that within habitat types singing males are equally likely to attract mates. This reflects the Ideal-free Distribution where individuals saturate a habitat type of highest quality until resource depletion impels individuals to utilize habitats of lesser quality, that impart less competition and concomitant equal fitness potential (Fretwell and Lucas, 1970). Under the Ideal-free Distribution, the greatest contribution of BITH each year would come from the lower elevation habitats that support lower individual densities, simply because of the greater proportion of habitat area. Loss or alteration of lower density habitats, without some mitigating increasing in higher density habitats, would carry more deleterious effects on total population size than loss of the highest density habitat.

Let us now assume that there is a strong correlation between habitat density and habitat quality in terms of reproductive fitness, and maintain the assumption that within habitat types singing males are equally likely to attract mates. Given the correlation, the relative importance of habitats remains unclear, because the magnitude of increased fitness must be known that offsets the contribution from lower habitat quality areas. For example, the model estimates there would be approximately 5 times more breeding pairs in lower elevation habitats, and therefore on average these would need only one-fifth the reproductive fitness to match the contribution of the highest quality areas. Furthermore, the assumption of equal nesting of singing males could be discarded, such that only some fraction needs to attract mates and have an intermediate reproductive fitness.

Thus, it cannot be assumed that areas of greatest density, even if they are of greatest habitat quality represent the greatest source of individuals to the population. The area of lower quality habitats occurring over greater land areas has the potential to outpace productivity of higher quality habitat. In order to further resolve the importance of various habitats, reproductive fitness data are need for various habitats to confirm if lower density habitats are in fact areas of diminished reproductive potential to be considered as population sinks (Martin, 1992).

Conclusions

The BITH habitat model has illustrated that habitats of lower BITH density could potentially be more important population sources than areas of higher BITH density. This relationship could be expected to hold for any population of organisms that follows a similar pattern. Any montane species that increases in density with increasing elevation will experience lesser amounts of habitat with increasing elevation. Or on flat terrain, any species with a limited patchy distribution of habitat that decreases in quality from the patch center, such that the outer reaches of the patch account for a greater area than the patch core. In any case, some measure of reproductive potential for the various grades of habitat are required.

This study has demonstrated that satellite imagery combined with habitat modeling can provide important information in understanding the distribution, density, and abundance for BITH. Rimmer et al. (1996) recommended that 1) population estimates for BITH be derived from calculations using remote sensing methods to quantify the extent of suitable habitat, and 2) density estimates be obtained from representative sites. The total population estimate of 67,000 singing males was presented

to establish an initial baseline estimate for the WMNF. The satellite imagery-derived model can be easily applied to sites in the Green Mountains, but should be further validated with independent data. Assuming reasonable validation at alternative sites, then distribution, density, and abundance estimates can be obtained for vast spatial extents, more focused mountain-top regions, or specific areas targeted for land use change.

Literature Cited

- Aspinall, R. and N. Veitch. 1993. Habitat mapping from satellite imagery and wildlife survey data using a Bayesian modeling procedure in a GIS. *Photogrammetric Engineering & Remote Sensing* 59(4): 537-543.
- Cohen, W. B. and T. A. Spies. 1992. Estimating structural attributes of Douglas Fir/western hemlock forest stands from Landsat and SPOT imagery. *Remote Sensing of Environment* 41: 1-17.
- Franklin, J. and D. W. Steadman. 1991. The potential for conservation of Polynesian birds through habitat mapping and species translocation. *Conservation Biology* 5:506-521.
- Franklin, S. E. 1994. Discrimination of subalpine forest species and canopy density using digital CASI, SPOT PLA, and Landsat TM data. *Photogrammetric Engineering & Remote Sensing* 60(10): 1233-1241.
- Fretwell, S. D. and H. L. Lucas. 1970. On territorial behavior and other factors influencing habitat distribution in birds. *Acta Biotheoretica* 19: 16-36.
- Green, K. M., J. F. Lynch, J. Sircar, and L. S. Z. Greenberg. 1987. Landsat remote sensing to assess habitat for migratory birds in the Yucatan Peninsula, Mexico. *Vida Silvestre Neotropical* 1(2): 457-471.
- Green, R. E. and G. H. Griffiths. 1994. Use of preferred nesting habitat by Stone Curlews *Burhinus oedicnemus* in relation to vegetation structure. *J. Zool., Lond.* 233:457-471.
- Hall, F. G., Y. E. Shimabukuro, and K. F. Huemrich. 1995. Remote sensing of forest biophysical structure using mixture decomposition and geometric reflectance models. *Ecological Applications* 5(4): 993-1013.

- Hepinstall, J.A., & S.A, Sader. 1997. Using Bayesian statistics, Thematic Mapper satellite imagery, and Breeding Bird Survey data to model bird species probability of occurrence in Maine. *Photogrammetric Engineering and Remote Sensing* 63(10): 1231-1237.
- Herr, A. M. and L.P. Queen. 1993. Crane Habitat Evaluation Using GIS and Remote Sensing. *Photogrammetric Engineering & Remote Sensing* 59(10): 1531-1538.
- Jakubauskas, M. E. 1996. Canonical correlation analysis of coniferous forest spectral and biotic relations. *Int. J. Remote Sensing* 17(12): 2323-2332.
- Jakubauskas, M. E. and K. P. Price. 1997. Empirical relationships between structural and spectral factors of Yellowstone Lodgepole pine forests. *Photogrammetric Engineering & Remote Sensing* 63(12): 1375-1381.
- Martin, T. E. 1992. Breeding productivity considerations: What are the appropriate habitat features for management? Pp.455-473. in *Ecology and Conservation of Neotropical Migrant Landbirds* (Hagan, J. M. III and Johnston, D. W., eds.). Smithsonian Institution Press. Washington D. C.
- Palmeirim, J. M. 1988. Automatic mapping of avian species habitat using satellite imagery. *Oikos* 52:59-68.
- Pulliam, H. R. 1988. Sources, sinks, and population regulation. *American Naturalist* 132: 652-661.
- Pulliam, H. R. and B. J. Danielson. 1991. Sources, sinks, and habitat selection: A landscape perspective on population dynamics. *American Naturalist* 137: S50-S66.
- Robinson, S. K. 1992. Population dynamics of breeding Neotropical migrants in a fragmented Illinois landscape. Pp.408-418. in *Ecology and Conservation of Neotropical Migrant Landbirds* (Hagan, J. M. III and Johnston, D. W., eds.). Smithsonian Institution Press. Washington D. C.
- Rimmer, C. C. and K. P. McFarland. 1997. Population density and demographic studies of Bicknell's Thrush on Mt. Mansfield, Vermont and other northeastern United States Peaks. Pp. 227-246. in *Vermont Forest Ecosystem Monitoring: Annual Report for 1995* (Wilmot, S. H. and T. D. Scherbatskoy, eds.). VforEM Ann. Rep. No. 5. Vermont Department of Forests, Parks, and Recreation. Waterbury, VT.
- Rimmer, C. C., J. L. Atwood, K. P. McFarland, and L. R. Nagy. 1996. Population density, vocal behavior, and recommended survey methods for Bicknell's Thrush. *Wilson Bulletin* 108(4): 639-649.

- Sader, S. A., R. B. Waide, W. T. Lawrence, and A. T. Joyce. 1990. Tropical forest biomass and successional age class relationships to a vegetation index derived from Landsat TM data. *Remote Sensing of Environment* 28: 143-156.
- Schwaller, M. R., C. E. Olson Jr., Z. Ma, Z. Zhu, and P. Dahmer. 1989. Remote sensing analysis of Adelie Penguin rookeries. *Remote Sensing of Environment* 28:199-206.
- Spanner, M. A., L. L. Pierce, D. L. Peterson, and S. W. Running. 1990. Remote sensing of temperate coniferous leaf area index: The influence of canopy closure, understory vegetation, and background reflectance. *International Journal of Remote Sensing* 11(1): 95-111.
- Thibault, D. S., S. Chalifoux, and M. Laperle. 1998. Using satellite imagery as a planning tool for Harlequin Duck inventory. *Int. J. Remote Sensing* 19: 5-9.
- Van Horne, B. 1983. Density as a misleading indicator of habitat quality. *J. Wildl. Manage.* 47(4): 893-901.
- Vickery, P. D., M. L. Hunter, Jr., and J. V. Wells. 1992. Is density an indicator of breeding success? *Auk* 109(4): 706-710.
- Vogelmann, J. E. and B. N. Rock. 1988. Assessing forest damage in high-elevation coniferous forests in Vermont and New Hampshire using Thematic Mapper data. *Remote Sensing of Environment* 24: 227-246.
- Wallace, G. J. 1939. Bicknell's Thrush, its taxonomy, distribution, and life history. *Proc. Boston Soc. Nat. Hist.* 41(6): 211-402.

Decile	IRA	Pixels	Area	% Area
0 - <0.10	0.05	19534	1587	1.7
0.10 - <0.20	0.15	379253	30805	33.7
0.20 - <0.30	0.25	434990	35332	38.7
0.30 - <0.40	0.35	169790	13791	15.1
0.40 - <0.50	0.45	71608	5816	6.4
0.50 - <0.60	0.55	34265	2783	3.0

Table 5.1. Table of Bicknell's Thrush (BITH) estimates of the amount of area occupied by the six probability deciles representing BITH presence for the entire study area. Values include the probability deciles of BITH presence in at least one of 5 years of point count sampling; IRA- Index of Relative Abundance, the midpoint of the decile range used to represent all pixels within the decile; Pixels- the number of pixels with values falling within the decile; Area- the area in hectares; and percentage of total area occupied by each decile. The dimensions of a single pixel are 28.5 x 28.5 m.

Decile	IRA	p(BITH)	Abundance	% Abundance	Density
0 - <0.10	0.05	0.01	195	0.3	4.9
0.10 - <0.20	0.15	0.032	12136	18.1	15.8
0.20 - <0.30	0.25	0.056	24359	36.4	27.6
0.30 - <0.40	0.35	0.083	14093	21.1	40.9
0.40 - <0.50	0.45	0.113	8092	12.1	55.6
0.50 - <0.60	0.55	0.147	5037	7.5	72.4

Table 5.2. Table of Bicknell's Thrush (BITH) estimates of relative abundance and density for the six probability deciles representing BITH presence for the entire study area. Values include the probability deciles of BITH presence in at least one of 5 years of point count sampling; IRA- Index of Relative Abundance, the midpoint of the decile range used to represent all pixels within the decile; p(BITH)- the probability of a BITH presence in one point count; Abundance- the estimated number of BITH presences within the decile; % Abundance- the percentage of all BITH within the decile; and Density- the abundance of BITH per 40 ha.

CHAPTER 6

SUMMARY OF FINDINGS

The primary finding of this research was the proof-of-concept that satellite imagery can effectively be used in the derivation and application of species-habitat models. For the first time, the spectral feature data from satellite imagery were used to estimate habitat model variables that were found to be important in explaining the distribution of a species. Solving the model at every pixel within the study area produced results that were successfully validated over most of predicted ranges of probabilities for which Bicknell's Thrush (BITH) was expected to occur. This satisfied the remaining two model criteria (see Introduction) of high accuracy and applied with variables found important in explaining the species' distribution set forth by Mosher et al., (1986). A map product was generated that included predicted probabilities of BITH presence at a fine spatial scale (28.5 meter pixels) and over a great geographic extent (91,000 ha).

Model output predicted the distribution of BITH habitats supporting various densities. The relationship was revealed that habitats supporting the greatest densities of BITH decline in total area, because of decreasing area with increasing elevation. This relationship carries various implications for the population dynamics and conservation of this species. Assuming the highest density habitat to be the most important overall source of new BITH with each breeding season overlooks potentially greater

contributions of individuals from habitats supporting lower densities, which occur over much greater total areas. Therefore, the modeling results illuminate directions for further research, which includes the need for reproductive fitness data within the full range of BITH habitats.

The potential for satellite imagery in habitat modeling to provide estimates of population size was revealed. The coincidental finding that the study sampling unit may approximately equal territory size of BITH and produce density estimates comparable to those found with territory mapping suggests that this method could be used to estimate breeding density. Further research, where various habitats are calibrated with breeding densities and measures of reproductive fitness estimated using territory mapping, would further strengthen our understanding of the population dynamics of BITH at a variety of spatial scales.

This study was conducted under the unfavorable conditions of extreme topographic relief, which can confound the use of satellite imagery. As a near worst case scenario, the methodology used in this study performed well. The topographic normalization technique, Minnaert Correction was used to adjust the ETM spectral features, so that slope and aspect effects on incident illumination and non-Lambertian reflectance were minimized. Adjusted reflectance from the Normalized Difference Vegetation Index and Band 1 were then used to estimate the dominant vegetation height variable at every pixel.

The ETM imagery was also used to construct a land cover map for the study area. Steps were taken to maximize the accuracy of the map to be included in developing model. The first step was to test the accuracies of maps generated using a variety of

topographic normalization techniques, including band-ratios, Minnaert Correction, and aspect partitioning. Little was known about these techniques' ability to deliver high accuracy land cover classifications. It was found that in terms of Overall Accuracy from analyses of error matrices that no technique outperformed the others. However, the combination of all three techniques (I4) did produce the greatest number of highest scored accuracies among the Producer's and User's Accuracy measures for individual land cover classes. The classification I4 was selected for use in habitat modeling.

Before employing the I4 classification in habitat modeling, it was examined for systematic errors. The confusion of the Kampfzone and High Mortality classes, and justification that these cover types were structurally very similar led to their being combined into the single class Fir Sapling, with a resulting rise in the accuracy of the classification. Distance to the nearest Fir Sapling cover type eventually became included in the final habitat model, suggesting that this was an important adjustment.

Literature Cited

Mosher, J. A., K. Titus, and M. R. Fuller. 1986. Developing a practical model to predict nesting habitat of woodland hawks. Pp. 31-35. in *Wildlife 2000: Modeling Habitat Relationships of Terrestrial Vertebrates* (Verner, J., M. L. Morrison, and C. J. Ralph, eds.). University of Wisconsin Press. Madison, WI.