

**LIDAR REMOTE SENSING FOR WILDLIFE HABITAT CHARACTERIZATION
AND MODELING: INCORPORATING REMOTELY SENSED VEGETATION
STRUCTURE INTO CURRENT ASSESSMENTS OF ANIMAL DISTRIBUTION AND
CONSERVATION**

A Dissertation

Presented in Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

with a

Major in Natural Resources

in the

College of Graduate Studies

University of Idaho

by

Sebastián Martinuzzi

June 2010

Major Professor: Lee A. Vierling, Ph.D.

UMI Number: 3424649

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

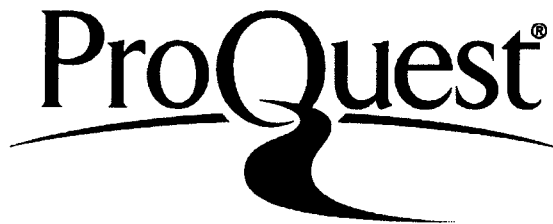
In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 3424649

Copyright 2010 by ProQuest LLC.

All rights reserved. This edition of the work is protected against unauthorized copying under Title 17, United States Code.

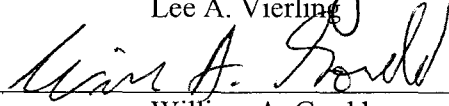


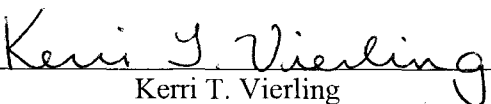
ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

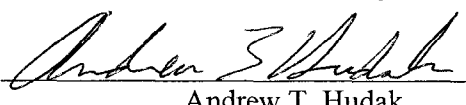
AUTHORIZATION TO SUBMIT DISSERTATION

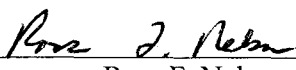
This dissertation of Sebastián Martinuzzi, submitted for the degree of Doctor of Philosophy with a major in Natural Resources and titled, "LiDAR Remote Sensing For Wildlife Habitat Characterization And Modeling: Incorporating Remotely Sensed Vegetation Structure Into Current Assessments Of Animal Distribution And Conservation", has been reviewed in final form, as indicated by the signatures and dates given below. Permission, as indicated by the signatures and dates given below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professor  Date 6/30/2010
Lee A. Vierling

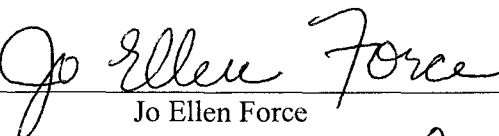
Committee Members  Date 7/14/2010
William A. Gould

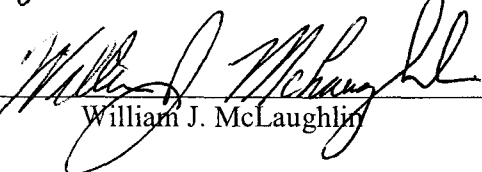
 Date 6/30/2010
Kerri T. Vierling

 Date 1 July 2010
Andrew T. Hudak

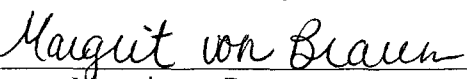
 Date 7/19/2010
Ross F. Nelson

 Date 7/1/2010
Eva K. Strand

Department Administrator  Date 7/30/10
Jo Ellen Force

Discipline's College Dean  Date 7/30/10
William J. McLaughlin

Final Approval and Acceptance by the College of Graduate Studies

 Date 8-4-10
Margrit von Braun

ABSTRACT

LIDAR REMOTE SENSING FOR WILDLIFE HABITAT CHARACTERIZATION AND MODELING: INCORPORATING REMOTELY SENSED VEGETATION STRUCTURE INTO CURRENT ASSESSMENTS OF ANIMAL DISTRIBUTION AND CONSERVATION

Remote sensing data play a key role for assessing wildlife habitat distribution and conservation. However, most efforts have depended on passive remote sensing data that poorly characterize the three-dimensional (3-D) structure of vegetation, an important variable influencing animal-habitat associations. In this thesis, we evaluated the consequences of integrating novel data of ecosystem 3-D structure from LiDAR (i.e. light detection and ranging) into current assessments of wildlife habitat distribution and conservation, with the main goal of quantifying LiDAR value for biodiversity and wildlife management. Using data from temperate and tropical landscapes (i.e. Idaho and Puerto Rico), this research exhibited the value of LiDAR data in characterizing key forest structure components for wildlife species, such as snags and understory shrub distribution, as well as for improved assessments of wildlife habitat suitability. In this sense, LiDAR helped to refine species-habitat models in ways not attained using traditional remote sensing technologies, making it possible to delineate known species associations with forest structure. In addition, LiDAR significantly improved the accuracy of current Landsat-based forest type classifications, which represent the principal source of geospatial data used in wildlife habitat studies. Finally, this research showed that incorporating remotely sensed data of vegetation structure can improve the results of regional conservation efforts such as Gap Analysis. This thesis demonstrated that LiDAR remote sensing has a great value for improved wildlife habitat assessments, providing unique opportunities to advance the way we manage and conserve biodiversity and habitats.

ACKNOWLEDGMENTS

To my major professor, Dr. Lee Vierling, for his guidance throughout this process, as well as to the members of my committee, including Dr. William Gould, Dr. Kerri Vierling, Dr. Andrew Hudak, Dr. Ross Nelson, and Dr. Eva Strand. I would like to thank all of you for the fruitful discussions, the freedom you gave me to pursue my ideas, and your guidance. I've learned a lot from you. I extend my acknowledgments to your institutions, including the University of Idaho, the USDA Forest Service, and the National Aeronautics and Space Administration (NASA).

I would like to acknowledge the USGS GAP Analysis Program for financing our research and trusting on us, and the University of Idaho Geospatial Laboratory for Environmental Dynamics (GLED) and the USDA Forest Service International Institute of Tropical Forestry for their technical and scientific support. In addition, several lab colleagues and collaborators provided frequent assistance with data analysis and helpful discussions, including Michael Falkowski, Jeffrey Evans, Steve Garrity, Jan Eitel, Jeffrey Lonneker, Alistair Smith, and Melanie Johnson. I extend my acknowledgments to the individuals who provided data or critical opinions in the different chapters of this dissertation.

Finally, I would like to thank Dr. Lee Vierling, Dr. William Gould, Dr. Jorge Frangi and Dr. Ariel Lugo for their motivation, encouragement, and opportunities during my scientific career.

DEDICATION

To my family and Maria Eugenia, for their support, patience and love.

TABLE OF CONTENTS

Title page	i
Authorization to submit dissertation.....	ii
Abstract.....	iii
Acknowledgements.....	iv
Dedication.....	v
Table of contents.....	vi
List of Tables	x
List of Figures	xii
Appendix.....	xiii
INTRODUCTION.....	1
References.....	4
CHAPTER 1. Improving the characterization and mapping of wildlife habitats with LiDAR data: Measurement priorities for the Inland Northwest, USA.....	6
Introduction.....	6
Methods.....	8
Results and Conclusion.....	9
Acknowledgements.....	11
References.....	11
Table	15
CHAPTER 2. Mapping snags and understory shrubs for a LiDAR-based assessment of wildlife habitat suitability.....	20
Abstract.....	20
Introduction.....	21
Background and rationale	22
Methods.....	26
Study area.....	26

Target wildlife species	26
Definition of understory shrub and snag classes.....	27
Field data acquisition and interpretation.....	28
LiDAR data acquisition and preprocessing	29
Classification tool	30
Understory shrub and snag mapping with LiDAR – Modeling approach and evaluation.....	31
Habitat Suitability Mapping.....	33
Results.....	35
Understory shrub distribution mapping	35
Snag distribution mapping	36
Distribution maps of understory shrubs and snags	36
Habitat suitability modeling.....	37
Discussion and conclusions	37
Acknowledgements.....	43
References.....	44
Tables.....	54
Figures.....	61

CHAPTER 3. Incorporating remotely sensed tree canopy cover data into broad scale assessments of wildlife habitat distribution and conservation	67
Abstract.....	67
Keywords	68
Introduction.....	68
Material and methods.....	70
GAP predictions of species distribution and the NLCD2001 tree canopy cover product	70
Study area.....	71
Data.....	71
Approach.....	72
Accuracy assessment of the new maps of bird species distribution	73

Results.....	74
Observed species-habitat associations based on NLCD_TCC and expected distribution patterns	74
Predicted species distribution incorporating NLCD_TCC data.....	75
Model evaluation	75
Species representation within the network of protected lands.....	76
Patterns of species richness after adding NLCD_TCC data.....	76
Discussion and conclusions	78
Acknowledgements.....	82
References.....	83
Tables.....	88
Figures.....	91
Appendix.....	95

CHAPTER 4. Classifying forest type and succession in a tropical dry forest: a LiDAR perspective.....

Abstract.....	96
Introduction.....	97
Methods.....	99
Study area.....	99
Vegetation and land use history	100
Data.....	101
Data preprocessing	103
Data analysis	104
Results.....	105
Discussion and conclusions	107
Acknowledgements.....	111
References.....	112
Figures.....	119
Tables.....	125

CONCLUSIONS AND FUTURE WORK130
References.....134

LIST OF TABLES

CHAPTER 1

Table. Species and (potential) lidar-derived habitat variables	15
---	----

CHAPTER 2

Table 1. Target avian species and habitat characteristics, based on the corresponding habitat suitability models (i.e. Schroeder, 1982; Souza, 1982, 1987; Roloff 2001)	54
--	----

Table 2. Understory shrub cover characteristics	55
--	----

Table 3. Snag distribution within forest successional stages. Inventory data was not available for old-growth forest (only GPS locations)	56
--	----

Table 4. LiDAR-derived metrics of canopy height (top group) and topography (bottom group). The list does not include multi-collinear variables, as they were initially identified and removed using QR-Decomposition (Becker et. al., 1988) in Falkowski et al. (2009). Variables selected for understory shrub and snag map predictions are identified with an “X”. Since snags were mapped using two different approaches (i.e. with and without image segmentation), an additional “X” was added when the variable was selected in both methods	57
---	----

Table 5. Minority/majority ratios for the different presence/absence classifications. The number of samples per class (P denotes presence and A denotes absence) used to derive the ratios is shown between parentheses.....	58
---	----

Table 6. Accuracy statistics for the model of understory shrub presence/absence	59
--	----

Table 7. Accuracy statistics for the different models of snags presence/absence, including with and without segmentation	60
---	----

CHAPTER 3

Table 1. Predicted habitats before and after adding the NLCD_TCC. The table includes estimates of total habitat area and area predicted within protected lands (i.e. reserves). Area unit corresponds to thousands of km ² . The species are presented by groups (denoted by the letter “G”) similar to Fig. 2. The scientific names of the species is provided in Appendix A	88
---	----

Table 2. Accuracy assessment of the initial (i.e. ID-GAP) and refined (i.e. ID-GAP + NLCD_TCC) predicted species distributions, including the number of sites used for evaluation, the % of correct predictions (CP%), and the % of omissions (OM%)	89
--	----

Table 3. Summary table of the consequences for GAP assessments brought about by the inclusion of broad scale remote sensing data about vegetation structure.90

CHAPTER 4

Table 1. Forest types in Guánica based on Lugo et al. (1978), Lugo 2005, Murphy and Lugo (1986), Farnsworth (1993), and Gould et al. (2008). The number of field samples used by this study (N) is also presented.....125

Table 2. Remote-sensing based explanatory variables.....126

Table 3. Forest type classification models, including sources of predictor variables (top section), accuracy statistics (center) and variables included in the final models (bottom). Models 1-5 developed from single sources of predictor variables; models 6-9 from multiple sources, and models 10-11 after adding elevation (ELEV) and curvature (CURV).....128

LIST OF FIGURES

CHAPTER 2

- Figure 1.** Snag distribution based on diameter.....61
- Figure 2.** Aggregation approach for converting the 20m-pixel LiDAR-based products of understory shrubs presence/absence into continuous, 100m-pixel (i.e. 1 ha) values. Snags were treated similarly62
- Figure 3.** LiDAR-based distribution maps for the different snag diameter classes (top) and understory shrubs (bottom), including the 20m-pixel presence/absence product to the right, and the 1-ha density map to the left. The presence/absence maps include, between parentheses, the proportional cover of the two classes (i.e. present vs. absent). 63
- Figure 4.** Habitat suitability maps for the different avian species. The maps on the right are simplified, aggregated and recoded versions depicting areas with habitat suitability index (HSI) ≥ 0.6 (i.e. suitable habitats)64
- Figure 5.** Conditional density plots for the understory shrub distribution vs. the 3 (LiDAR-derived) predictor variables included in the final model.....65
- Figure 6.** Conditional density plots for the snag distribution vs. the (LiDAR-derived) Median Absolute Deviation of Heights (HMAD). The figure shows the example for the snag diameter class ≥ 25 cm, but a similar trend was observed for the other snag classes66

CHAPTER 3

- Figure 1.** Simplified land cover map and tree canopy for Idaho.....91
- Figure 2.** Distribution of the 37 avian species according to land cover (simplified to forest/non-forest) and tree canopy cover preferences (three classes). Five groups of species were identified based on similar species-habitat relationships. Potential values of maximum species richness for scenarios with and without tree canopy information are also shown. Overestimation range refers to areas (in terms of habitat associations) where the distributions of the species have been overestimated according to the knowledge of the species' natural history.....92
- Figure 3.** Pixel based, species richness maps for the 37 avian species, including from the original ID-GAP predicted habitats and from the new ones incorporating the NLCD_TCC. The comparison of species richness between these products is also presented93
- Figure 4.** Pixel based, species richness layers for all the birds in Idaho (n=238), before and after incorporating the NLCD_TCC. Subsets of areas dominated by closed

forest and mid-open forest are shown in the center. Patterns of species richness and distribution of protected lands are displayed in the bottom94

CHAPTER 4

Figure 1. Location of the study area119

Figure 2. Map of forest types for Guánica derived from LiDAR data120

Figure 3. Classification trees decisions and resultant accuracy statistics for different successional stages using LiDAR (top) or Landsat ETM+ (bottom) variables. Mid-sec. = Mid-secondary; Late-sec. = Late-secondary; (A) and (L) describe agricultural or logging-based past land uses respectively121

Figure 4. Forest types vs. remotely sensed predictor variables. Mesq. = Mesquite; S-dec. = Semi-deciduous; S-everg. = Semi-evergreen.122

Figure 5. Topographic detailed captured by the 30-meter pixel LiDAR and SRTM digital elevation models123

Figure 6. Successional classes vs. remotely sensed predictor variables. 1 = Late-secondary forest; 2 = Mid-secondary forest; 3 = Primary forest; (A) and (L) describe agricultural or logging-based past land uses respectively.....124

APPENDIX

CHAPTER 3

Appendix A.....95

INTRODUCTION

Remote sensing data are essential in assessing species distributions because they allow the linkage of species with habitats over broad spatial extents (Nagendra, 2001; Ferrier, 2002; Kerr & Ostrovsky, 2003; Turner et al., 2003; McDermid et al., 2005; Leyequien et al., 2007). However, current efforts often depend on passive remote sensing data that are typically unable to describe the three-dimensional (3-D) structure of vegetation (e.g. Landsat imagery), an important variable affecting the presence, abundance and habitat use of many wildlife species (MacArthur & MacArthur, 1961; Brokaw & Lent, 1999). This limitation can introduce uncertainty and errors in local assessments of wildlife habitat, therefore affecting efforts related to wildlife habitat conservation and management. Today, spatially explicit data about forest structure is a major need to advance biodiversity and wildlife habitat management efforts (Russell et al., 2007; Venier & Pearce, 2007). LiDAR remote sensing has therefore opened new opportunities for assessing wildlife habitats in forests due to its unique sensitivity to the 3-D structure of ecosystems (Lefsky et al., 2002; Vierling et al., 2008; Bergen et al., 2009). Evaluating the consequences of integrating LIDAR data into current remotely-sensed assessments of wildlife habitat distribution and conservation is therefore an important need, and constitutes the main goal of this thesis.

While recent studies have explored the use of LiDAR data for quantifying different aspects of forest structure and wildlife habitat (see reviews by Wulder et al., 2008 and Vierling et al., 2008), more research is needed to better understand the potential of this technology for improved assessments of species distribution, wildlife habitats and conservation (Vierling et al., 2008). In this study I tried to answer the following questions: (1) which type of information about ecosystem structure is needed to refine predictions of wildlife species distributions?; (2) what are the consequences of incorporating geospatial information about vegetation structure into local and regional assessments of wildlife species distribution, conservation, and biodiversity?; (3) can LiDAR data enable the distribution of key wildlife habitat features in forests such as snags (i.e. standing dead trees) and understory shrubs to be mapped?; and (4) to what degree can LiDAR data improve forest land cover maps –i.e. the most common remote

sensing product for quantifying wildlife habitat and biodiversity? During my doctoral work, but not presented in this thesis, I also teamed with others to review and discuss the potential of LiDAR data for characterizing and modeling wildlife habitats in general (Vierling et al., 2008), and to evaluate the capability of LiDAR data for separating forest successional stages -an important step towards improved quantification of ecosystem characteristics related to wildlife habitats and biodiversity, among other important functions (Falkowski et al., 2009).

The study areas of this research encompassed temperate and tropical sites including Idaho (in the Inland Northwest, USA, and dominated by conifer forests) and Puerto Rico (in the Caribbean, characterized by tropical broad-leaf forests), and for which a variety of datasets were already available (e.g. airborne LiDAR, forest inventory plots, species-habitat models). For the purpose of assessing species distribution and conservation by this study, I used a methodological framework that is consistent with the U.S. Geological Survey's Gap Analysis Program (GAP). GAP is the major governmental initiative assessing wildlife habitat distribution and conservation using remote sensing, and the main source of geospatial data for supporting biodiversity planning in the nation (<http://gapanalysis.nbi.gov>). In GAP Analysis, species distributions are predicted using a deductive approach, that is, by identifying species-habitat relationships based on what is known about the species natural history, and then by reflecting these relationships using GIS environmental layers (e.g. land cover data, see Scott et al., 1993). Inductive approaches, on the other hand, base their inferences from the statistical relationships between species occurrence records and environmental layers, and are also widely used. The ultimate goal in GAP is to map vegetation types and to predict the distribution of wildlife species, and to evaluate their representation within the network of protected lands, so that "gaps" in conservation can be identified (Scott et al., 1993). While both inductive and deductive approaches have pros and cons (Scott et al., 2002a), the findings of this study are relevant for remote sensing, wildlife habitat assessments, and GAP conservation assessment, regardless of the modeling approach being used.

This thesis includes four chapters. In Chapter 1 I evaluated which type of information about ecosystem structure is needed to refine GAP predictions of species

distributions. I focused on avian and mammal species present in Idaho and for which the absence of ecosystem structure data was a recognized limitation in a previous Gap effort (Scott et al., 2002b). This chapter identifies priorities in the development of LiDAR-derived products of habitat structure, critical for refining wildlife habitat models in the Inland Northwest, USA. This chapter has been published in the *Gap Analysis Bulletin* (Martinuzzi et al., 2009a).

In Chapter 2 I evaluated the use of LiDAR data for mapping the distribution of snags and understory shrubs, as well as habitat suitability for avian species that depend on those features. I focused on a mixed-conifer forest in northern Idaho. This chapter shows new value of LiDAR data in characterizing key forest structure components for assessing wildlife and biodiversity in temperate conifer forests, and has been published in *Remote Sensing of Environment* (Martinuzzi et al., 2009b).

In Chapter 3 I evaluated the impacts of incorporating information about vegetation structure into State-wide assessments of wildlife species distribution and conservation. Basically, I compared the outcomes of a GAP analysis that incorporated percent tree canopy cover (a measure of horizontal vegetation structure) vs. a GAP analysis that does not (i.e. the classic GAPs). Using the avian species and the entire state of Idaho as a case study, this chapter highlights the value of area wide vegetation structure data for refined habitat assessments, and shows -among other findings- that the inclusion of vegetation structure can have major consequences in the results of GAP. This chapter has been published in the *Journal of Applied Remote Sensing* (Martinuzzi et al., 2009c).

Finally, in Chapter 4 I evaluated the utility of LiDAR data to improve current, Landsat-based classifications of forest types and forest successional stages. I focused on a complex tropical forest in the dry region of Puerto Rico. This chapter shows the value of LiDAR data for refining current forest land cover maps, and highlights the potential of LiDAR for improving our understanding of the biophysical and human dimensions of tropical dry forests – one of the world’s most threatened habitats. This chapter will be submitted to *Biotropica*.

References

- Bergen KM, Goetz SJ, Dubayah RO, Henebry GM, Hunsaker CT, Imhoff ML, Nelson RF, Parker GG, & Radeloff VV. (2009). Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications LiDAR and radar spaceborne missions. *Journal of Geophysical Research*, 114, G00E06:13.
- Brokaw NV & Lent RA. (1999). Vertical Structure. In: M. Hunter (Ed.), *Maintaining Biodiversity in Forest Ecosystems* (pp. 373-399). Cambridge University Press, Cambridge.
- Falkowski MJ, Evans JS, Martinuzzi S, Gessler PE, & Hudak AT. (2009). Characterizing Forest Succession with Lidar Data: an Evaluation for the Inland Northwest USA. *Remote Sensing of Environment*, 113, 946-956
- Ferrier S. (2002). Mapping Spatial Patterns in Biodiversity for Regional Conservation Planning: Where to from here? *Systematic Biology*, 51, 331-363.
- Kerr JT, & Ostrovsky M. (2003). From space to species: ecological applications for remote sensing. *Trends in Ecology and Evolution*, 18, 399-305.
- Lefsky MA, Cohen WB, Harding DJ, & Parker GG. (2002). Lidar remote sensing for forest ecosystem studies. *BioScience*, 52, 19-30
- Leyequien E, Verrelst C, Slot M, Schaepman-Strub G, Heitkönig IMA, & Skidmore A. (2007). Capturing the fugitive: Applying remote sensing to terrestrial animal distribution and diversity. *International Journal of Applied Earth Observation and Geoinformation*, 9, 1-20.
- MacArthur RH & MacArthur JW. (1961). On birds species diversity. *Ecology*, 42, 594-598.
- Martinuzzi S, Vierling L, Gould W, & Vierling K. (2009a). Improving the characterization and mapping of wildlife habitats with LiDAR data: measurement priorities for the Inland Northwest, USA. In: J. Maxwell et al., (Editors). *Gap Analysis Bulletin*, 16, USGS/BRD/Gap Analysis Program, Moscow, ID, USA.
- Martinuzzi S, Vierling L, Gould W, Falkowski M, Evans J, Hudak A, & Vierling, K. (2009b). Mapping snags and understory shrubs for a LiDAR-based assessment of wildlife habitat suitability. *Remote Sensing of Environment*, 113, 2533-2546.

- Martinuzzi, S., L. Vierling, K. Vierling, W. Gould, & A. Hudak. (2009c). Incorporating remotely sensed tree canopy cover data into broad scale assessments of wildlife habitat distribution and conservation. *Journal of Applied Remote Sensing*, 3, 033568.
- McDermid GJ, Franklin SE, & LeDrew EF. (2005). Remote sensing for large-area habitat mapping. *Progress in Physical Geography*, 29, 449-474.
- Nagendra H. (2001). Using remote sensing to assess biodiversity. *International Journal of Remote Sensing*, 22, 2377-2400.
- Russell RE, Saab VA, & Dudley J. (2007). Habitat suitability models for cavity-nesting birds in a postfire landscape. *The Journal of Wildlife Management*, 71, 2600-2611
- Scott JM, Davis F, Csuti B, Noss R, Butterfield B., et al. (1993). Gap Analysis – A geographic approach to protection of biological diversity. *Wildlife Monographs*, 123, 1-41.
- Scott JM, Heglund P, Morrison M, Huafler J, Raphael M, Wall W, & Samson F. (Eds) (2002a). *Predicting species occurrences: issues of accuracy and scale*. Island Press, Washington, DC, USA.
- Scott JM, Peterson CR, Karl JW, Strand E, Svancara LK, & Wright NM. (2002b). A Gap Analysis of Idaho: Final Report. Idaho Cooperative Fish and Wildlife Research Unit. Moscow, ID, USA
- Turner W, Spector S, Gardiner N, Fladeland M, Sterling E, & Steininger M. (2003). Remote sensing for biodiversity science and conservation. *Trends in Ecology and Evolution*, 18, 306-313.
- Venier LA & Pearce JL. (2007). Boreal forest landbirds in relation to forest composition, structure, and landscape: implications for forest management. *Canadian Journal of Forest Research*, 37, 1214-1226
- Vierling KT, Vierling LA, Martinuzzi S, Gould W, & Clawges R. (2008). Lidar: Shedding new light on habitat modeling. *Frontiers in Ecology and the Environment*, 6, 90-98
- Wulder MA, Bater CW, Coops NC, Hilker T, & White JC. (2008). The role of LiDAR in sustainable forest management. *Forest Chronicle*, 84, 807-826.

CHAPTER 1

IMPROVING THE CHARACTERIZATION AND MAPPING OF WILDLIFE HABITATS WITH LIDAR DATA: MEASUREMENT PRIORITIES FOR THE INLAND NORTHWEST, USA

Note: This chapter has been published as: Martinuzzi S, Vierling L, Gould W, & Vierling K. (2009). Improving the characterization and mapping of wildlife habitats with LiDAR data: measurement priorities for the Inland Northwest, USA. In: J. Maxwell et al., (Editors). *Gap Analysis Bulletin*, 16. USGS/BRD/Gap Analysis Program, Moscow, ID, USA.

Introduction

The development of region- and nation-wide predictive assessments of wildlife species distribution and habitat availability is a major component of the USGS Gap Analysis Program (GAP), which provides critical information for conserving biodiversity in the United States (Scott et al., 1993). Despite continuous advances in predictive modeling tools, the lack of detailed and accurate geospatial data is still a recognized, major challenge to improve species distribution modeling (Guisan and Zimmermann, 2000). Current predictions, for example, are based on environmental geospatial data that do not reflect the three-dimensional characteristics of vegetation (Gottschalk et al., 2005; McDermid et al., 2005; Leyequien et al., 2007), an important variable for determining the distribution and abundance of wildlife species (MacArthur and MacArthur, 1961; Brokaw and Lent, 1999). Modeling species distribution using environmental data that do not adequately represent important species-environment relationships can result in predictions that contain some level of uncertainty and error (Fielding and Bell, 1997; Beutel et al., 1999; Guisan and Zimmermann, 2000), affecting species conservation and biodiversity assessments such as those made through the GAP.

Lidar (light detection and ranging) is a relatively new source of geospatial data that, contrary to most available remote sensing technologies, provides fine-grained information about the 3-D physical structure of terrestrial and aquatic ecosystems (Lefsky et al., 2002), opening a novel spectrum of possibilities for characterizing wildlife habitats with remote sensing (Vierling et al., 2008). In forested environments, for example, lidar data have been useful for quantifying vegetation structure in terms of biomass (e.g. basal area and tree diameter), percent canopy cover, tree height, tree density, for separating forest successional stages and to characterize subcanopy topography (e.g. Nelson et al., 1988; Harding et al., 2001; Drake et al., 2002; Hofton et al., 2002; Hudak et al., 2006). Recent studies evaluating the utility of lidar for mapping understory shrubs and snag density yield also positive results (Goodwing, 2006; Bater, 2008). While lidar data have recently been utilized to investigate local-scale wildlife habitat quality as it relates to avian (e.g. Hinsley et al., 2002, 2006; Hill et al., 2004; Broughton et al., 2006; Goetz et al., 2007; Clawges et al., 2008) and fish (Jones, 2006; McKean et al., 2008) biology, application of lidar data to broad scale species distribution prediction is still in the exploratory stage (see Vierling et al., 2008).

Lidar data acquisitions are typically localized efforts conducted over small areas, and therefore these local efforts have not been ideal for the scales at which GAP work (e.g. state, region, country). However, an increasing number of states currently have or plan to have soon full lidar coverage (e.g. Florida, Iowa, Louisiana, Pennsylvania, North Carolina, Ohio, and Texas). Moreover, as a result of increasing demands from State and Federal agencies, academia, and private industry, the US government is currently evaluating the feasibility and strategy for a national acquisition of high resolution, high accuracy lidar data for all 50 states. This effort is known as the “National Lidar Initiative” (NLI) and it is organized by the USGS (Stoker et al. 2007). According to the Center for Lidar Information Coordination and Knowledge <<http://lidar.cr.usgs.gov/>>, the NLI “is currently in the early stages of determining viability, developing what this dataset should look like, what kinds of information contained in a lidar signal are most important for the U.S. people, and what each stakeholders’ roles and responsibilities could be”.

The objective of this study was to evaluate which habitat structure variables are needed to refine GAP species distribution predictions, in order to identify priorities in developing lidar-derived products. This study was focused on avian and mammal species inhabiting the Inland Northwest, USA. In this region, previous efforts to predict species distribution with traditional remote sensing data (e.g. Landsat) indicated that the distribution of many wildlife species has been likely overestimated due to the incapability of incorporating information (i.e. constraints) about vegetation structure (Scott et al. 2002). For example, species that are known to occur in closed forests have been predicted to occur in all forests (closed and open) due to the lack of geospatial data about percentage of tree canopy cover. Information from this report has direct implications for further ecological applications of lidar data, including from the NLI, and could have long-term ramifications for improving GAP species distribution predictions and land cover characterization.

Methods

First, we identified the mammal and avian species whose predicted habitat distribution has been overestimated, according to Scott et al. (2002). Scott et al. (2002) also provide information about the type of habitat variables needed to improve the predicted distribution of various species. We refined and expanded the habitat information using published material from habitat suitability models, such as those developed by the US Fish and Wildlife Service. For example, Scott et al. (2002) indicated that the predicted distribution of the pileated woodpecker was likely overestimated due to the lack of geospatial data about the presence of snags, which is a large determinant of the species habitat distribution. According to the habitat suitability model for the pileated woodpecker, not only the size and density of snags, but also the percentage of tree canopy cover, are important variables for predicting the distribution of the species (Schroeder, 1982). We then combined all the information in a table (one for the avian species and other for the mammals) that included, in the columns, the species whose

habitat distribution has been overestimated, and in the rows, the potential habitat variables needed to refine the predictive species distribution models.

We included information (i.e. habitat variables) about vegetation structure as well as topography. Scott et al. (2002) indicated that the original topographic data / digital elevation model was not adequate to characterize relevant habitat features for certain species. Lidar, on the other hand, is the best available technology for topographic mapping. In addition, the table lists seven species whose predicted distribution performed well according to Scott et al. (2002), but that may benefit from Lidar data due to the high affinity of the species to structural characteristics of vegetation. Examples of these species are the downy woodpecker and hairy woodpecker, whose presence depends on the availability of snags, among other factors.

Results and Conclusion

We identified a total of eleven variables of habitat structure potentially suitable for refining GAP predictions of species distribution. These variables included, for forests, 1) % of tree canopy cover, 2) some measure of forest stand biomass, such as the mean tree diameter, basal area, or age, 3) diameter and density of snags, 4) height of overstory trees, 5) diversity of the tree canopy (i.e. number of canopy strata), 6) tree density, and 7) % of understory shrub cover. For rangelands, the important variables were the height and % of shrub cover, as well as the height of the grasses. Finally, in terms of topography, important variables included rock outcrops (i.e. identification of rocky areas), and morphological measures of streams, creeks, and canyons (see Table).

The list included a total of 86 species, including 66 avian species and 20 mammal species, equivalent to almost 30% and 20% of all the avian and mammal species present in Idaho. In addition, 10 of the 86 species are of greatest conservation need according to the Idaho Fish and Game. We believe the list of species presented in this study may represent a conservative lower-bound of the actual overall number of species whose

predictive distribution models would benefit from the inclusion of lidar-derived data, since the structural habitat preferences of many vertebrate species are either unknown or often not reported.

According to the total number of species associated with each habitat variable (reported at the end of the Table), the results of this study indicated that the most needed variables are (in order of importance): 1) % tree canopy cover, 2) % shrub canopy cover (including understory shrubs), 3) some measure of stand biomass (mean tree diameter / basal area / age), 4) shrub height, and 5) size and density of snags. While lidar has been used to successfully quantify tree canopy cover and biomass in different forest types, little is known about the capabilities of this new technology for mapping the distribution of snags, and for measuring the characteristics of the shrub layer (whereas as part of the forest understory or in rangelands) (Goodwing, 2006; Bater, 2008). More research on these topics would serve to better evaluate the potential of lidar data to characterize wildlife habitats and support predictions of species distribution. In addition, in order to facilitate ecological and conservation applications of broad-scale lidar data such as those from the NLI, further studies should evaluate the type of information about the structural characteristics of habitats needed to model wildlife species distribution and habitat availability in other regions and across different taxa. For example, while information about vegetation structure is important for birds and mammals, information about microtopography appears to be critical for improving assessments of reptile habitats (C. Peterson, personal communication). An additional benefit of lidar data is that it allows the development of products and maps at a high spatial resolution, suitable not only for vegetation assessments in upland areas but also in riparian zones, which are important habitat features for wildlife species but are particularly challenging to map with traditional (i.e. 30m pixel) remote sensing technologies (Goetz 2006).

The impending acquisition of a US-wide lidar dataset has the potential to provide new and relevant geospatial data, suitable for supporting and refining GAP predictions of species distribution and further species conservation assessments for the United States. In order to take maximum benefit from current and future lidar data for GAP related purposes, further studies should evaluate the performance of species distribution models

with and without lidar data, and its consequences for GAP assessments of wildlife species distribution and conservation. Finally, we recommend that GAP continue to work in cooperation with a variety of governmental, private and non-governmental organizations to achieve nationwide improvements in remotely-sensed habitat mapping.

Acknowledgements

This work was made possible through funding by the USGS Gap Analysis Program and the USDA Forest Service International Institute of Tropical Forestry (IITF). Two anonymous reviewers provided valuable comments that improved the quality of this paper. Work at IITF is done in collaboration with the University of Puerto Rico.

References

- Bater, C.W. 2008. Assessing Indicators of Forest Sustainability Using LiDAR Remote Sensing (Thesis). University of British Columbia, Vancouver, Canada (97 pp.)
- Beutel, T.S., R.J.S. Beeton, G.S. Baxter. 1999. Building better wildlife-habitat models. *Ecography* 22:219-223.
- Brokaw, N.V.L., and R.A. Lent. 1999. Vertical Structure. Pages 373-399 in M.L. Hunter, editor. *Maintaining Biodiversity in forest ecosystems*. Cambridge, UK, Cambridge University Press.
- Broughton RK, S.A. Hinsley, P.E. Bellamy, et al. 2006. Marsh tit *Poecile palustris* territories in a British broad-leaved wood. *Ibis* 148:744–52.
- Clawges, R., K.T. Vierling, L.A. Vierling, E.M. Rowell. 2008. The use of lidar remote sensing to estimate avian species diversity, density, and occurrence in a pine/aspens forest. *Remote Sensing of Environment* 112:2064-2073.
- Drake, J. R.O. Dubayah, D.A. Clark, R.G. Knox, B. Blair, M. Hofton, R.L. Chazdon, J.F. Weishampel, and S.D. Prince. 2002. Estimation of forest structural characteristics using large-footprint lidar. *Remote Sensing of Environment* 79:305-319.

- Fielding, A.H., and J.F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24: 38-49.
- Goetz, S.J. 2006. Remote sensing of riparian buffers: past progress and future prospects. *Journal of the American Water Resources Association* 42:133-143.
- Goetz, S., D. Steinberg, R. Dubayah, and B. Blair B. 2007. Laser remote sensing of canopy habitat heterogeneity as a predictor of bird species richness in an eastern temperate forest, USA. *Remote Sensing of Environment* 108:254-263.
- Goodwing, N.R. 2006. Assessing Understorey Structural Characteristics in Eucalyptus Forests: an investigation of LiDAR techniques (Thesis). University of New South Wales, Sydney NSW Australia (206 pp.)
- Gottschalk, T.K., F. Huettmann, and M. Ehlers. 2005. Thirty years of analyzing and modeling avian habitat relationships using satellite imagery data: a review. *International Journal of remote Sensing* 26:2631-2656.
- Guisan, A. and N.E. Zimmermann. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* 135:147-186.
- Harding, D.J., M.A. Lefkys, G.G. Parker, J.B. Blair. 2001. Laser altimeter canopy height profiles: methods and validations for closed-canopy, broadleaf forests. *Remote Sensing of Environment* 76:283-297.
- Hill, R.A., S.A. Hinsley, D.L.A. Gaveau, and P.E. Bellamy. 2004. Predicting habitat quality for great tits (*Parus major*) with airborne laser scanning data. *International Journal of Remote Sensing* 25:4851-55.
- Hinsley, S.A., R.A. Hill, D.L.A. Gaveau, and P.E. Bellamy. 2002. Quantifying woodland structure and habitat quality for birds using airborne laser scanning. *Functional Ecology* 16:851-57.
- Hinsley, S.A., R.A. Hill, P.E. Bellamy, and H. Balzter. 2006. The application of lidar in woodland bird ecology: climate, canopy structure, and habitat quality. *Photogrammetric Engineering and Remote Sensing* 72:1399-1406.
- Hofton, M.A., L.E. Rocchio, J.B. Blair, and R. Dubayah. 2002. Validation of vegetation canopy lidar sub-canopy topography measurements for a dense tropical forest. *Journal of Geodynamics* 34:491-502.

- Hudak, A.T., N.L. Crookston, J.S. Evans, M.J. Falkowski, A.M.S. Smith, P.E. Gessler, and P. Morgan. 2006. Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return lidar and multispectral satellite data. *Canadian Journal of Remote Sensing* 32:126-138.
- Jones, J.L. 2006. Side channel mapping and fish habitat suitability analysis using lidar topography and orthophotography. *Photogrammetric Engineering and Remote Sensing* 72:1202-06.
- Lefsky M.A., W.B. Cohen, G.G. Parker, and D.J. Harding. 2002. Lidar remote sensing for ecosystem studies. *BioScience* 52:19-30.
- Leyequien, E., C. Verrelst, M. Slot, G. Schaepman-Strub, I.M A. Heitkönig, A. Skidmore. 2007. Capturing the fugitive: Applying remote sensing to terrestrial animal distribution and diversity. *International Journal of Applied Earth Observation and Geoinformation* 9:1-20.
- MacArthur, R.H., and J.W. MacArthur. 1961. On birds species diversity. *Ecology* 42:594-598.
- McDermid, G.J., S.E. Franklin, and E.F. LeDrew. 2005. Remote sensing for large-area habitat mapping. *Progress in Physical Geography* 29:449-474.
- McKean, J.A., D.J. Isaak, and C.W. Wright. 2008. Geomorphic controls on salmon nesting patterns described by a new, narrow-beam terrestrial-aquatic lidar. *Frontiers in Ecology and the Environment* 6:125-130.
- Nelson, R.F., Krabill, W.B., and Tonelli, J. 1988. Estimating forest biomass and volume using airborne laser data. *Remote Sensing of Environment* 24:247-267.
- Schroeder, R.L. 1982. Habitat suitability index models: Pileated woodpecker. US Dept. Int., Fish and Wildlife Service. FWS/OBS-82/10.39 (15 pp.)
- Scott, J.M., F. Davis, B. Csuti, R. Noss, B. Butterfield, et al. 1993. Gap Analysis – A geographic approach to protection of biological diversity. *Wildlife Monographs* 123:1-41.
- Scott, J.M., C.R. Peterson, J.W. Karl, E. Strand, L.K. Svancara, and N.M. Wright. 2002. A Gap Analysis of Idaho: Final Report. Idaho Cooperative Fish and Wildlife Research Unit. Moscow, ID, USA.

- Stoker, J., J. Parrish, D. Gisclair, D. Harding, R. Haugerud, M. Flood, H. Andersen, K. et al. 2007. Report of the First National Lidar Initiative Meeting, February 14-16, 2007, Reston, Va. US Department of Interior, USGS Report Series OF 2007-1189.
- Vierling, K.T., L.A. Vierling, W. Gould, S. Martinuzzi, R. Clawges. 2008. Lidar: Shedding new light on habitat characterization and modeling. *Frontiers in Ecology and the Environment* 2: 90-98.

chickadee	<i>rufescens</i>						
Chipping sparrow	<i>Spizella passerina</i>	X					
Clark's nutcracker	<i>Nucifraga columbiana</i>	X					
Common goldeneye	<i>Bucephala clangula</i>		X				
Common nighthawk	<i>Chordeiles minor</i>	X					
Common poorwill	<i>Phalaenoptilus nuttallii</i>	X					
Common raven	<i>Corvus corax</i>	X					
Cordilleran flycatcher	<i>Empidonax occidentalis</i>	X					
Downy woodpecker*	<i>Picoides pubescens</i>		X	X			
Dusky flycatcher	<i>Empidonax oberholseri</i>	X				X	
Ferruginous hawk ±	<i>Buteo regalis</i>					X	X X
Flammulated owl ±	<i>Otus flammeolus</i>	X					
Fox sparrow	<i>Passerella iliaca</i>	X					
Golden eagle	<i>Aquila chrysaetos</i>	X					
Great gray owl	<i>Strix nebulosa</i>	X					
Greater sage grouse ±	<i>Centrocercus urophasianus</i>	X					
Hairy woodpecker*	<i>Picoides villosus</i>	X	X	X			
Hammond's flycatcher	<i>Empidonax hammondii</i>					X	
Lark bunting	<i>Calamospiza melanocorys</i>						X
Lark sparrow	<i>Chondestes grammacus</i>	X					
Lazuli bunting	<i>Passerina</i>	X				X	

		<i>amoena</i>			
Lesser scaup ±*	<i>Aythya affinis</i>			X	X
Lewis' woodpecker ±	<i>Melanerpes lewis</i>	X	X	X	
Lincoln's sparrow	<i>Melospiza lincolnii</i>			X	
Loggerhead shrike	<i>Lanius ludovicianus</i>	X			
Long-eared owl	<i>Asio otus</i>	X			
Macgillivray's warbler	<i>Oporornis tolmiei</i>			X	
Mountain bluebird	<i>Sialia currucoides</i>			X	
Nashville warbler	<i>Vermivora ruficapilla</i>			X	
Northern flicker	<i>Colaptes auratus</i>	X			
Northern goshawk	<i>Accipiter gentiles</i>	X			
Northern pygmy-owl	<i>Glaucidium gnoma</i>	X			
Northern saw- whet owl	<i>Aegolius academicus</i>	X			
Olive-sided flycatcher	<i>Contopus cooperi</i>	X		X	
Orange- crowned warbler	<i>Vermivora celata</i>			X	
Oregon (Dark- eyed) junco	<i>Junco hyemalis</i>	X			
Peregrine falcon ±	<i>Falco peregrinus anatum</i>	X			
Pileated woodpecker	<i>Dryocopus pileatus</i>	X	X		
Red-breasted nuthatch	<i>Sitta Canadensis</i>	X			

squirrel												
Hoary bat	<i>Lasiurus cinereus</i>						X					
Hoary marmot	<i>Marmota caligata</i>											X
Long-legged myotis	<i>Myotis volans</i>						X					
Long-tailed vole	<i>Microtus longicaudus</i>	X										
Mule deer	<i>Odocoileus hemionus</i>	X										
Northern flying squirrel	<i>Glaucomys sabrinus</i>					X						
Pronghorn	<i>Antilocapra americana</i>							X		X		
Red-tailed chipmunk	<i>Tamias ruficaudus</i>	X										
Rock squirrel ±	<i>Spermophilus variegates</i>											X
Southern red-backed vole	<i>Clethrionomys gapperi</i>	X			X							
White-tailed jack rabbit	<i>Lepus townsendii</i>	X										
	Total	13	5	2	0	1	0	3	2	0	3	0
	GRAND TOTAL	57	11	7	1	1	1	23	8	5	4	3

± Species of greatest conservation need in Idaho

*Species whose predicted distribution performed well according to Scott et al. (2002), but which may benefit from lidar data

CHAPTER 2

MAPPING SNAGS AND UNDERSTORY SHRUBS FOR A LIDAR-BASED ASSESSMENT OF WILDLIFE HABITAT SUITABILITY

Note: This chapter has been published as: Martinuzzi S, Vierling L, Gould W, Falkowski M, Evans J, Hudak A, & Vierling, K. 2009. Mapping snags and understory shrubs for a LiDAR-based assessment of wildlife habitat suitability. *Remote Sensing of Environment*, 113, 2533-2546.

Abstract

The lack of maps depicting forest three-dimensional structure, particularly as pertaining to snags and understory shrub species distribution, is a major limitation for managing wildlife habitat in forests. Developing new techniques to remotely map snags and understory shrubs is therefore an important need. To address this, we first evaluated the use of LiDAR data for mapping presence/absence of understory shrub species and different snag diameter classes important for birds (i.e. ≥ 15 cm, ≥ 25 cm and ≥ 30 cm) in a 30,000 ha mixed-conifer forest in Northern Idaho (USA). We used forest inventory plots, LiDAR-derived metrics, and the Random Forest algorithm to achieve classification accuracies of 83% for the understory shrubs and 86% to 88% for the different snag diameter classes. Second, we evaluated the use of LiDAR data for mapping wildlife habitat suitability using four avian species (one flycatcher and three woodpeckers) as case studies. For this, we integrated LiDAR-derived products of forest structure with available models of habitat suitability to derive a variety of species-habitat associations (and therefore habitat suitability patterns) across the study area. We found that the value of LiDAR resided in the ability to quantify 1) ecological variables that are known to influence the distribution of understory vegetation and snags, such as canopy cover, topography, and forest succession, and 2) direct structural metrics that indicate or suggest the presence of shrubs and snags, such as the percent of vegetation returns in the lower strata of the canopy (for the shrubs) and the vertical heterogeneity of the forest canopy

(for the snags). When applied to wildlife habitat assessment, these new LiDAR-based maps refined habitat predictions in ways not previously attainable using other remote sensing technologies. This study highlights new value of LiDAR in characterizing key forest structure components important for wildlife, and warrants further applications to other forested environments and wildlife species.

Introduction

The lack of spatially explicit data about forest three-dimensional structure is a major challenge for managing biodiversity and wildlife habitat (Russell et al., 2007; Venier and Pearce, 2007). Such information is important because characteristics associated with the structure of forests (e.g. height of the trees, presence or absence of understory, canopy closure, tree diameter, abundance and size of dead trees, etc.) are important factors explaining 1) the presence of many wildlife species, 2) the functional use of the habitat (e.g. nesting, foraging, cover, roosting), and 3) the overall diversity of wildlife species in forests (MacArthur and MacArthur, 1961; Davis, 1983; Brokaw and Lent, 1999). During the last two decades, passive remote sensing data have been used to characterize successfully different aspects of forested habitats over broad areas, but have been typically unable to describe three-dimensional (3-D) structural characteristics (see reviews by Kerr and Ostrovsky, 2003; Wulder and Franklin, 2003; McDermid et al., 2005). As a result, it is necessary to develop novel ways to characterize forest structure, with a special emphasis on those aspects that are relevant to wildlife habitat and biodiversity.

LiDAR remote sensing can be used to measure directly the 3-D structure of terrestrial and aquatic ecosystems across broad spatial extents (Lefsky et al., 2002). LiDAR data, in conjunction with various sources of ancillary data, have been used to quantify successfully different aspects of forest 3-D structure, such as biomass, canopy cover and height, canopy height profiles, successional stages, as well as subcanopy topography (Nelson et al., 1988; Harding et al., 2001; Drake et al., 2002; Hofton et al.,

2002; Hudak et al., 2006, Clawges et al., 2007; Hudak et al., 2008a; Falkowski et al., 2009). These data have been recently incorporated into assessments of biodiversity (Goetz et al., 2007; Clawges et al., 2008) and wildlife habitat modeling (see Vierling et al., 2008 for a review). However, the mapping of certain habitat characteristics requires more research. For instance, little is known about the capability of LiDAR data for mapping the distribution of snags (i.e. standing dead trees) and understory shrub species, two critical components of wildlife habitat in forests (Thomas et al., 1979; Davis, 1983; Hagar, 2007) and indicators of forest biodiversity and ecosystem health (Sampson and Adams, 1994; Noss, 1999, Kerns and Ohmann, 2004).

This study advances the application of LiDAR remote sensing for mapping forest structure and wildlife habitat. Our objective was to evaluate the use of LiDAR data to map 1) the distribution of understory shrubs and snags, and 2) habitat suitability patterns for different wildlife species known to be dependent upon these habitat resources. This study was focused on Moscow Mountain, a mixed-conifer forest located in the Inland Northwest (US) that has previously served as a suitable testbed for numerous LiDAR applications (Hudak et al., 2006, 2008a,b; Evans and Hudak, 2007; Falkowski et al., 2009).

Background and rationale

Mapping the distribution of snags and understory shrub species across the landscape presents major challenges. Recent studies using LiDAR data have been able to characterize height and/or cover of the understory vegetation, where understory is represented by all the woody vegetation in the strata (i.e. shrubs and trees), or suppressed trees only (e.g. Riaño et al., 2003; Maltamo et al., 2005; Goodwin 2006; Skowronski et al., 2007; Hill and Broughton 2009). This work was done through the use of canopy height thresholds, cluster analysis and visual interpretation. The studies have shown, however, that assessments of understory vegetation with LIDAR are typically less accurate under dense tree canopies (e.g. Maltamo et al., 2005; Goodwin 2006; Skowronski et al., 2007; Su and Bork 2007), where the proportion of laser pulses reaching the lower forest strata decreases. Maps of understory shrub distribution should

be reliable under different forest density conditions so that spatially consistent ecological inferences can be made. In this sense, Hill and Broughton (2009) showed that it is possible to characterize understory vegetation in closed forests, by integrating leaf-on and leaf-off LiDAR data. This approach, however, requires 1) a forest dominated by deciduous trees, and 2) the availability of multiple LiDAR acquisitions over the same area. In addition, animal use of different understory vegetation components does vary. In the coniferous forests of the Pacific Northwest (US), for example, distinguishing deciduous shrubs from conifer saplings is vital for evaluating certain types of wildlife habitats, as these components have different ecological function (see Hagar, 2007). A recent study conducted in an Aspen parkland in Canada, however, found no relationship ($p > 0.05$) between the structure of the understory (true) shrub community and the LiDAR reference data (Su and Bork 2007). With regard to snags, Bater (2007) was able to relate the structural heterogeneity of forest stands from LiDAR with proportions of trees in different stages of decay; in a conifer-dominated coastal forest of British Columbia, Canada. The study indicated that more research is required to test this approach in other forest environments. In addition, it is important in many wildlife habitat applications to understand not only the spatial distribution of snags, but also their size (e.g. larger animal species typically use larger snag diameters than smaller species). In this sense, a previous effort predicting the volume of standing dead wood material from LiDAR derived canopy metrics achieved poor results (RMSE 79%) (Pesonen et al., 2008).

A variety of environmental factors can influence the presence of snags and understory shrubs in forests, and therefore have the potential to serve as predictor variables in a distribution modeling approach. Studies evaluating the structure and composition of understory vegetation found that overstory canopy structure, topography and land use can all influence the presence of understory shrub cover in forests (Kilina et al., 1996; McKenzie and Halpern, 1999; Van Pelt and Franklin, 2000; Kerns and Ohmann, 2004; Bartemucci et al., 2006; Gracia et al., 2007), with overstory density being, frequently, the most important variable. Understory vegetation is consistently

denser in open forests, where more light can reach the ground; but it is more variable and less predictable in closed forests (Bartemucci et al., 2006).

The abundance and size of snags, on the other hand, is a result of the combined processes of forest succession, natural snag dynamics, forest management practices, and episodic disturbance events (Cline et al., 1980; Ohmann et al., 1994; Flanagan et al., 2002; Korol et al., 2002; Kennedy et al. 2008). Older forest stands typically support larger snags than do younger stands. In mountainous regions, topographic positions (i.e. slopes and aspects) exposed to more severe weather conditions tend to support higher abundance of snags (Flanagan et al., 2002). At the same time, managed forest stands have typically fewer larger snags than non-managed stands (Korol et al., 2002; Kennedy et al., 2008). Episodic disturbance events such as drought, snow, ice, fire or insect outbreaks can also increase the number of snags locally (Morrison and Raphael, 1993). A previous study modeling snag density with Landsat and geoclimatic data showed modest results, with only half of the predictions falling within a 15% deviation from the field validation values (Frescino et al., 2001). On the other hand, Bater (2007) found that the coefficient of variation of the LiDAR height data was a strong predictor of the proportion of trees in different stages of decay at the stand level ($r= 0.85$, $p < 0.001$, $RMSE = 4.9\%$). We are unaware of efforts to model the presence of snags of different sizes.

LiDAR data can be utilized to derive a variety of environmental factors known to explain the presence of understory vegetation and snags, including canopy structure (e.g. Hudak et al., 2008a,b), forest successional stage (e.g. Falkowski et al., 2009), and topography (e.g. Hudak et al., 2008a,b). Coupled with the fact that laser pulses can also interact directly with understory vegetation and dead trees, the use of LiDAR data should provide a way to advance the mapping of understory shrub and snag distributions in forested environments. In addition, while previous efforts assessing wildlife habitat with LiDAR have been focused from an inductive perspective, that is, by allowing the canopy metrics explain the variation in some type of field animal data (such as abundance, reproductive success, and richness) (e.g. Hinsley et al., 2002, 2008; Broughton et al., 2006; Goetz et al., 2007; Clawges et al., 2008; Graf et al., 2009), few studies have assessed wildlife habitats from a deductive perspective, that is, through the mapping of

known, key species habitat features (Hyde et al., 2006; Nelson et al., 2006; Swatantran et al., 2008).

Habitat suitability models or indices (a.k.a. HSIs) are common tools used by researchers and managers with the objective of assessing the potential of an area to support the resource, shelter and reproductive needs for given wildlife species (Turner et al., 2001; Edenius and Mikusiński, 2006). These models quantify species-habitat relationships based on empirical data and literature review on limiting resources, with values that range between 0.0 (unsuitable habitat) and 1.0 (optimum habitat). For the United States, a large number of habitat suitability models is available through the US Fish and Wildlife Service (USFWS) and individual efforts. Edenius and Mikusiński (2006) provide a comprehensive description of HSI worldwide sources and applications.

The spatial output of HSIs is a map depicting habitat suitability values across the landscape, for the target species. In this sense, Nelson et al. (2006) used LiDAR to identify forest patches with more than 20 m in height, which are known to be suitable for the endangered Delmarva fox squirrel (*Sciurus niger cinereus*). This assessment, however, recognized the lack of spatial data about understory vegetation, which is a complementary variable explaining the distribution of the species (Nelson et al., 2006). In California, Hyde et al. (2005, 2006) used LiDAR data to map forest biomass, canopy cover and height at the landscape scale, with the expectation that these products will be helpful to assess habitat suitability for the California spotted owls. Finally, Swatantran et al. (2008) combined forest structural data from LiDAR with maps of stressed and dead vegetation from a hyperspectral sensor, to map potential habitats for the Ivory-billed woodpecker (*Campephilus principalis*). Here, we use known information about species habitat preferences to map habitat suitability. Our study includes 1) multiple species from a different wildlife group (i.e. avian), and 2) the use and development of additional habitat variables of forest structure relative to previous studies.

Methods

Study area

Moscow Mountain comprises about 30,000-ha of managed, mixed temperate coniferous forest in Northern Idaho (Latitude 46°44'N, Longitude 116°58'W) (Falkowski et al., 2005). The area is topographically complex. Common tree species include ponderosa pine (*Pinus ponderosa*), Douglas fir (*Pseudotsuga menziesii*), grand fir (*Abies grandis*), western red cedar (*Thuja plicata*), and western larch (*Larix occidentalis*). Shrub species include Ocean Spray (*Holodiscus discolor*), Ninebark (*Physocarpus malvaceus*), Common snowberry (*Symphoricarpos albus*), Spiraea (*Spiraea betulifolia*), Huckleberry (*Vaccinium membranaceum*), and Mountain Maple (*Acer galbrum*) (Falkowski et al., 2005). Forest species composition varies with temperate/moisture gradient (Cooper et al., 1991). Private industrial forest companies manage most of the area for timber, but a large tract of experimental forest is also owned and managed by the University of Idaho for research purposes. The city of Troy, ID manages a watershed. Private landowners manage many land parcels, and there is a small tract of old growth forest protected as a county park. All these factors contribute to the structural and compositional complexity found in the Moscow Mountain forests. Approximately 83% of the study area is covered by forest in different stages of succession (Falkowski et al., 2009). Young and mature forests cover 65% of the total area; stand initiation (i.e. growing space reoccupied by seedlings, saplings, or shrubs following stand replacing disturbance) represents 10%; understory reinitiation (i.e. older cohort of trees being replaced by new individuals) represent 7%, and old growth forest 1%. The remaining 17% corresponds to non-forest, open areas of grasses or weeds (Falkowski et al., 2009).

Target wildlife species

To apply our work to wildlife habitat, we selected four bird species that inhabit Moscow Mountain (Scott et al., 2002), including the dusky flycatcher (*Empidonax oberholseri*), hairy woodpecker (*Picoides villosus*), Lewis's woodpecker (*Melanerpes lewis*), and downy woodpecker (*Picoides pubescens*), and made use of the published habitat

suitability models available for these species (i.e. Schroeder, 1982; Souza, 1982, 1987; Roloff 2001). The habitat requirements of these species comprise a broad range of forest structural variables including but not limited to snags and understory shrubs, making these species ideal for evaluating the potential of mapping habitat suitability from LiDAR data (Table 1). Furthermore, woodpecker species have been found to be indicators of overall forest bird diversity (Virkkala, 2006).

Definition of understory shrub and snag classes

We considered understory shrubs to be present if they covered more than 25% in a 20m by 20m pixel. This definition was based on land cover mapping standards and the species HSIs. A 25% threshold cover per pixel was established by the Multi-Resolution Land Characteristics Consortium to define the class “shrubland” in the 1992 (US) National Land Cover map. At the same time, the habitat suitability for the species that use understory was low or zero when shrub cover was less than 25%. For the purposes of our study, the understory shrub class is comprised of true shrub species only, and did not include saplings that can be common in the understory. The reason for this resides in the ecological function that non-coniferous vegetation has in Pacific Northwest conifer forests, as described by Hagar (2007). In this region, non-coniferous vegetation determines the abundance and distribution of many vertebrates, providing the foundation for food webs through direct and indirect food resources (i.e., broad leaf forage, fruits, flowers, and insects) that are not provided by conifers (Hagar, 2007). This is the case of the Lewis’s woodpecker and the dusky flycatcher, which use the understory shrub layer as a food source of insects (see Table 1). The dusky flycatcher uses the understory shrub layer also for nesting, and studies in Idaho have found that nesting occurs exclusively in non-coniferous plants (i.e. in true shrubs) (Kroll and Haufler, 2006).

We focused our attention on the snag diameters that are used by the woodpecker species of this study, based on their specific HSIs. According to these models, the snag diameters (at breast height, or DBH) are ≥ 15 cm for the Downy woodpecker (Schroeder, 1982), ≥ 25 cm for the Hairy woodpecker (Souza, 1987), and ≥ 30 cm for the Lewis’s

woodpecker (Souza, 1982) (Table 1). We use the term “classes”, “ranges”, or “categories” indistinctively to refer to these snag diameters.

Field data acquisition and interpretation

We utilized forest inventory plots and LiDAR data that have been acquired by previous efforts to characterize various aspects of forest structure in the region (Falkowski et al., 2005, 2009; Hudak et al., 2006, 2008a,b; Evans and Hudak, 2007). These plots contain a variety of snag and understory shrub densities, and thus were suitable for this study. The use of standard forest inventory plots should facilitate the application of the findings of this study to other areas, as well as the evaluation of limitations of such datasets for wildlife habitat assessments.

Eighty-three, 405 m² fixed-radius (11.35 m radius) forest inventory plots were located across the study area in 2003 by Falkowski et al. (2005), using a stratified random sampling protocol designed to capture the full range of canopy structure conditions and forest species composition. Information in each plot included the number and diameter (i.e., DBH) of dead trees, and the percentage of (true) shrub cover, among other standard forest inventory data. All trees (live or dead) with diameter > 2.7 cm were measured. Within each plot, visual estimates of true shrub cover were obtained in 4 subplots (4m by 4m in size), using a reference schema of 12 canopy cover classes that ranges from 0% to 95-100% (see Falkowski et al., 2005). The percentage of shrub cover for each plot was obtained by averaging the estimates of the smaller subplots. More information about the field data used in this study can be obtained in Falkowski et al. (2005).

Understory shrubs were present (i.e. >25% cover) in 48 of the 83 plots. The median shrub cover for those 48 plots was 53% (Table 2). The height of the shrubs observed in the individual subplots (N = 83 x 4 =332) was typically below 2 m (80% of the cases). In addition, there were 177 snags in the sampled population, with a diameter ranging from 12.7 cm to 97.0 cm. and with small snags greatly outnumbering larger snags (Fig. 1). Within those, there were 151 snags with DBH ≥ 15cm, 73 snags with DBH ≥ 25cm, and 46 snags with DBH ≥ 30 cm. Snags were common as they appeared in

about 55% of the plots. Eighty five percent of the snags sampled were smaller than 40 cm in diameter. Comparable snag diameters and a skewed class distribution have been found in other managed conifer forests (Ganey, 1999; Spiering and Knight, 2005). The median snag density of the Moscow Mountain plots was 1 snag per plot. Expressed at the ha-scale, this is equivalent to 25 snags/ha, which is close to the 32 snags/ha found in a comparable area (Spiering and Knight, 2005). For those plots in which snags were present, the snags ≥ 15 cm and ≥ 25 cm appeared with a median density of 2 snags per plot, and the snags ≥ 30 cm with a median density of 1 snag per plot. In addition, the presence of snags differed depending upon the successional stage of the plot (Table 3).

LiDAR data acquisition and preprocessing

We used LiDAR-derived metrics developed by previous studies, which have proved useful for mapping and predicting different attributes of forest structure (see Hudak et al., 2006, 2008a,b; Falkowski et al., 2009). Discrete, multiple return LiDAR data (1.95 m nominal post spacing) was acquired by Horizons, Inc., in the summer of 2003, using an ALS40 system operating at a wavelength of 1064 nm and flown at approximately 2500 m elevation. LiDAR data were first separated into ground and non-ground returns using the Multi-scale Curvature Classification algorithm by Evans and Hudak (2007). Thirty four LiDAR-based metrics, consisting of 19 canopy height metrics and 15 topographic metrics, were then calculated at the plot scale (Hudak et al., 2008a; Falkowski et al., 2009) (Table 4). This was done by clipping the raw LiDAR data using the plot extent. LiDAR metrics were also calculated for the entire area at a spatial grid resolution of 20 m, which corresponded to the dimensions of the field plots (Hudak et al., 2008a). We used these metrics as predictor variables for mapping snags and understory shrubs. We also utilized auxiliary, LiDAR-derived products developed in previous studies. This included a map of basal area (BA) by Hudak et al. (2008a) (accuracy = 98 %), and a map of six forest successional stages by Falkowski et al. (2009) (accuracy = 95 %).

Classification tool

For mapping snag and understory shrub distribution we used the Random Forest (RF) algorithm (Breiman, 2001), a novel and powerful extension of classification tree techniques that has been shown to produce excellent results in classifications of remotely sensed and ecological data (Lawrence et al., 2006; Prasad et al., 2006; Cutler et al., 2007). Random Forest can handle complex interactions among predictor variables without making distributional assumptions and without overfitting (Breiman, 2001; Lawrence et al., 2006; Cutler et al., 2007). The RF algorithm develops classification rules by estimating a large number of trees (100s to > 1,000s; i.e., a forest), in which each classification tree is based on a random subset of the training data, and each classification tree split is based on a random subset of the predictor variables (Breiman, 2001). After the iterations, the predictions from the individual classification trees are combined using the rule of majority votes. Classification accuracies that results from this approach perform very well compared to other classifiers (Liaw and Wiener, 2002). The RF algorithm provides a reliable internal estimate of classification accuracy using the portion of the data that is randomly withheld as each classification tree is developed (i.e., the out-of-bag sample [OOB], approximately 37% of the training data), which makes it unnecessary to have a separate accuracy assessment (Breiman, 2001; Prasad et al., 2006; Lawrence et al., 2006). In addition, the RF algorithm provides information about the importance of each predictor variable, by quantifying changes in classification error when the OOB data for that variable is altered. Hudak et al. (2008a,b) and Falkowski et al. (2009) found the RF algorithm to be practical for analyzing the field plot and LiDAR data used in this study.

We used the RF package (Liaw and Wiener, 2002) in R (www.r-project.org; R Development Core Team, 2005). We added a model selection step using the model selection algorithm varSelRF (Díaz-Uriarte and Alvarez, 2006), which is a RF-based tool available also in R. The varSelRF algorithm iteratively eliminates the least important variables (with importance as measured from RF), resulting in a model with the smallest possible number of variables and whose error rate is within one standard error of the minimum error rate of all forests (Díaz-Uriarte, 2008). Our predictor variables (i.e. the

LiDAR-derived metrics) were continuous, therefore avoiding potential bias in variable selection that can result from combining discrete and continuous predictors (see Strobl et al., 2007). Our training data involved a random component, avoiding potential bias in accuracy assessments involving cross validation (Huang et al., 2003).

Understory shrub and snag mapping with LiDAR – Modeling approach and evaluation

We modeled the distribution (i.e. presence/absence) of understory shrubs and snags with RF, by applying the corresponding best model to the entire region. In addition, we modeled the distribution of snags using a segmentation-based approach. Data segmentation fragments the data into smaller, more homogeneous regions based on some ecological, spectral, or geographic attribute. This approach typically increases the quality of the final classification. We evaluated the consequences of segmenting the snag data based on forest succession, because succession is an important ecological variable influencing the presence and size of snags in forests, which we observed to be occurring in our study area as the abundance of snags differed among the different successional classes (see Table 3). As a result, we segmented the study area into three regions with distinctive snag abundances, including 1) an area composed by the Open and Stand Initiation categories (OA&SI), without snags; 2) an area composed by the Young Multistory and Understory Reinitiation (YMS&UR) categories, with snags present but less than those observed in 3) the Mature Multistory (MMS) category. Finally, the MMS class was combined with the Old Growth forest (OF) (less than 1% of the study area) to form the third segmentation region (MMS&OF). As a result, the three areas defined were: OA&SI (without snags), and YMS&UR and MMS&OF (with snags present in different proportions).

Segmentation of our data reduced the amount of training data available for classification and the area to be classified (i.e. the target area), but maintained the overall relationship between training data / target area, and therefore the representation of the field plots. For example, the area-wide ratio of training data / target area for Moscow Mountain was 0.83 (i.e. 83 plots / 100% of the study area), for the MMS&OF area the ratio was 0.68 (i.e. 26 plots / 38% of Moscow Mountain) and for the YMS&UR area the

ratio was 1.14 (i.e. 40 plots / 35% of Moscow Mountain). Fu et al. (2005) showed that the results from bootstrap cross-validations are reliable with small sample sizes (as small as 16). In summary, we modeled the distribution of the three different snag classes (i.e. ≥ 15 cm, ≥ 25 cm and ≥ 30 cm) using two approaches, one that included developing a single predictive model/classification for the entire area, and another that resulted from a combination of different predictive models within three different areas (i.e. OA&SI, YMS&UR, and MMS&OF). Since snags were absent in the OA&SI, we did not model their distribution in this region. Finally, we compared the results of the final snag classifications with and without including the segmentation approach.

Previous to any presence/absence classification (for either shrubs or snags) with RF, we ensured that the input data were balanced. Studies have shown that severe imbalanced data sets (i.e., when the presence of absence classes constitute a very small minority of the data distribution) can pose significant drawbacks in the performance attainable by most machine learning classification systems, including RF (see Chen et al., 2004; Sun et al., 2007). The ratio between the number of samples for the minority class and the number of samples for the majority class constitutes the minority/majority ratio, ranging in values between 1 and >0 . Unfortunately, there is not a universal ratio between minority and majority classes defining what constitutes an imbalanced vs. balanced data set. In practice, however, imbalanced data in presence/absence classifications typically include cases in which the minority class (whether presence or absence) represents 10% or less of the data, equivalent to a minority/majority class ratio ≤ 0.11 (see Chen et al., 2004; Sun, 2007). In another study with RF, Ruiz-Gazen and Villa (2002) used a conservative ≤ 0.2 class-ratio threshold to define the presence of imbalanced classes. In our data set, the minority/majority class ratios were much higher (i.e. closer to 1.00) than those reported by previous studies, indicating that the data were balanced and suitable for classification with RF (Table 5).

We ran the VarSelRF algorithm several times for each classification, including 50 runs for those that were conducted in the entire area and 20 runs for those that were conducted in the smaller, segmented portions. Running the varSelRF several times allowed us to evaluate the potential presence of different candidate solutions. Each time,

we incorporated the varSelRF solution model into the RF algorithm to evaluate the resulting misclassification error. If different candidate solutions were present, we selected the final model based on the criteria of smallest total and within class errors and smaller number of variables. The final predictive distribution models were applied across the region using the AsciiGridPredict command in the yaImpute package (Crookston and Finley, 2008) available in the R software package. We ran the RF algorithm with 5,000 bootstrap replicates in order to stabilize individual class error. Redundant (i.e. multi-collinear) predictor variables were removed in an early stage of this study (see Falkowski et al., 2009). Finally, and in order to verify that the results from our study were not conditioned by the variable selection method used, we compared the models identified by varSelRF with those from another RF-based algorithm, recently developed in ecological applications (Murphy et al., 2009). We found that the models selected by the VarSelRF were consistent with the models identified by the other routine.

We assessed the accuracy of the final classifications (i.e. snag and understory shrub presence/absence) using the confusion matrices and errors generated by RF. From these, we calculated the overall accuracy, user and producer accuracies, commission and omission errors, and the kappa statistic (Congalton and Green, 1999). Although widely used, some studies have criticized the use of the kappa statistic for assessing binary classifications, due to its sensitivity to prevalence (i.e., the proportions of Presences in the data) (see Alluoché et al., 2006 for a review). In order to verify the suitability of the kappa statistic, we compared it to the true skill statistic (TSS; Alluoché et al., 2006), which is a novel variation of kappa that corrects for potential biases introduced by prevalence. We found insignificant differences between the kappa and TSS values (i.e. maximum difference of 0.03), therefore supporting the use of the kappa statistic in this study.

Habitat Suitability Mapping

The first step for mapping wildlife habitat suitability involved the aggregation of the LiDAR-derived layers of vegetation structure (i.e. the habitat variables) to 1 ha pixel, because the hectare is the spatial unit of application of the HSIs used in this study (see

Schroeder, 1982; Souza, 1982, 1987; Roloff, 2001; Kroll and Haufler, 2006). The presence/absence maps of understory shrub and snag distributions were developed at a spatial resolution of 20 m, similar to the field plots. We aggregated these maps to the 1-ha pixel by multiplying the proportion of (20m-pixel) shrub (or snag) presences found within the 1ha-pixel, by the median plot shrub cover (or snag density) derived from the field data (see Figure 2). This approach allowed us not only to aggregate the data to the proper spatial scale for HSI mapping (i.e. 1-ha grain size), but also to transform the data from a presence/absence binary format to a continuous approximation of shrub cover and snag density at the landscape scale. This aggregation approach, however, imposed a maximum threshold of shrub cover and snag density per ha that we could distinguish. In this sense, if all the 20m-plots located within a 1ha-pixel were predicted as shrub presences, then the percentage of shrub cover for that hectare would be 53% ($(25/25)*52.9$; where 25/25 is the proportion of 20m-pixels with presences and 52.9 is the median plot shrub cover). As a result, we were able to distinguish continuous shrub cover categories below that threshold (53%) but not above that. Similarly, we were able to distinguish continuous snag densities below 50 snags per ha (for classes ≥ 15 cm and ≥ 25 cm diameter), and below 25 snags per ha (for the class ≥ 30 cm diameter), but not above. This issue, however, did not affect the HSI modeling. This is because the habitat suitability values above those shrub and snag densities were constant, making the distinction of more classes unnecessary.

We also aggregated to 1ha grain size the LiDAR derived canopy density metric and the map of basal area by Hudak et al. (2008a). In addition, we generated two auxiliary habitat layers reflecting information about the mean diameter of the overstory trees, a variable included in the cover and reproduction components of the HSI for the hairy woodpecker. In this HSI model, the mean diameter (DBH) of the overstory trees was used as a surrogate of forest succession (see Souza, 1987). We made use of the map of forest succession by Falkowski et al. (2009), and recoded the different classes into different habitat suitability values based on 1) the mean overstory DBH values observed in the different classes in the field plots, and 2) the relationships between mean DBH and

habitat suitability established by Souza (1987). This new auxiliary layer reflects the species' higher preference for mature forests than younger forests.

Once all the habitat geospatial variables were at 1-ha pixel resolution, we applied the HSI formulae for the different avian species. Specifically, the HSI map for the Dusky flycatcher was constructed from the tree canopy cover and understory shrub cover layers; the HSI map for the Hairy woodpecker was constructed from the layer of DBH, tree canopy cover and snags ≥ 25 cm diameter; the HSI map for the Lewis's woodpecker was constructed from the layers of understory shrub cover, snags ≥ 30 cm diameter, and tree canopy cover; and the HSI map for the Downy woodpecker was constructed with the layers of basal area and snags ≥ 15 cm diameter (see Table 1). The original HSI for the Dusky flycatcher included an additional variable, which is the cover of understory vegetation with height less than 1 m (Roloff, 2001). This variable, however, receives only half of the weight of the other two variables included in the HSI (see Roloff, 2001), and therefore was not considered in this study. Finally, each final map of habitat suitability included a measure of overall accuracy based on the accuracy of the individual layers used to build the models, as an estimate of error propagation.

Results

Understory shrub distribution mapping

The understory shrub presence/absence prediction yielded overall and individual class accuracies of 83%. The model included three predictor variables (i.e. LiDAR metrics), including two from canopy (STRATUM0 and STRATUM2) and one from topography (SCOSA) (Table 6). The metric STRATUM0 is the proportion of ground returns (i.e., height = 0 m); the metric STRATUM2 is the proportion of vegetation returns between 1 and 2.5 meters in height, and the metric SCOSA (Stage 1976) describes the percent slope times the cosine of aspect transformation.

Snag distribution mapping

The accuracy of the snag classifications without segmenting the data yielded acceptable overall accuracies, ranging between 72% and 80% (Table 7). However, the Present category had low accuracies, especially for the snag classes $\geq 25\text{cm}$ and $\geq 30\text{cm}$. The kappa for classifications without segmentation ranged between 0.43 and 0.59. The inclusion of a segmentation approach resulted in a net increase in the quality of the classifications. It increased the overall accuracy (ranging now between 86% and 88%), the kappa values (now ≥ 0.7), and the accuracy of both the presence and absence classes. At the same time, it decreased the commission and omission errors (see Table 7). In the classifications without segmentation the accuracy of the individual classes ranged between 58% and 89%, with most below 80%. After the segmentation, the accuracy of the individual classes ranged between 73% and 95%, with most above 80%. The number of predictor variables included in the models ranged between 2 and 6, and included typically a combination of canopy height and topographic metrics. The most common variable was the Median Absolute Deviation of Height (HMAD). Among the topographic variables, landform (i.e. BOLSTAD metrics) and distance to streams (i.e. FLDIST metric) appeared several times.

Distribution maps of understory shrubs and snags

The 20m-pixel presence/absence maps revealed that less than half of Moscow Mountain has shrubs present in the understory. For the snags, the extent of the different diameter classes decreased rapidly with increasing snag diameters, from 45% of the study area for the snags $\geq 15\text{cm}$ to 30% for the snags $\geq 30\text{cm}$ (Fig. 3). At the hectare scale, the map of understory shrub density revealed that less than 10% of the total area had more than 50% of understory shrub cover per ha. Areas with shrub cover between 25% and 50% per ha represented about 40% of the study area, and the remaining 50% of Moscow Mountain was dominated by very low or no shrub cover. The density of snags per hectare rapidly decreased with increasing snag diameters (Fig. 3).

Habitat Suitability Modeling

The spatial representation of the HSI models revealed the presence of different patterns of habitat suitability and habitat availability for the four different avian species (Fig. 4). For the Lewis's woodpecker, a species that uses snags and understory shrubs, approximately half of Moscow Mountain was unsuitable (HSI=0.0). Approximately 15% of the area could be deemed "suitable" (considered as areas with $HSI \geq 0.6$ [Prosser and Brooks 1998]) and there were no optimum habitats (i.e. $HSI = 1.0$). For the dusky flycatcher, a species associated with understory shrubs, the amount of suitable habitat was also low, with no optimum areas. For the hairy woodpecker and the downy woodpecker, two relatively common species, few areas were determined to be unsuitable (i.e. $HSI = 0.0$), and about one third of Moscow Mountain habitat was suitable ($HSI \geq 0.6$). However, the suitable habitat for these species occurred in different portions of Moscow Mountain, reflecting the different habitat requirements (see Table 1). Some areas, although small, were classified as optimum habitats for these two avian species.

The final accuracy of the HSI maps (after adding the errors of the input layers) ranged between 79% and 91%. Specifically, the HSI map accuracy was 79% for the hairy woodpecker, 90% for the Lewis's woodpecker, 92% for the dusky flycatcher, and 92% for the downy woodpecker.

Discussion and conclusions

The lack of spatial data about snags and understory shrub distribution is a recognized limitation for managing wildlife habitat in forests (Russell et al., 2007; Venier and Pearce, 2007). We found that LiDAR data provided valuable information for mapping the distribution of snags, understory shrubs, and wildlife habitat suitability in a mixed-conifer forest, representing an important step in the characterization of forest structure and wildlife habitat with remote sensing.

Metrics of canopy height and topography derived from LiDAR allowed us to map the distribution of understory shrubs with an overall and individual class accuracy of more than 80%. Only three metrics were needed, including two from canopy (STRATUM0 and STRATUM2), and one from topography (SCOSA). STRATUM0 (the percentage of ground returns) is inversely proportional to canopy density (percentage of non-ground returns ≥ 1 m in height; $r^2 = 0.73$). Previous research indicated that forests with open canopies tend to support more shrubs than those with closed canopies (Kilina et al., 1996; Bartemucci et al., 2006). Analysis of conditional density function plots (CDF) for the shrub field data vs. the percent of ground returns from LiDAR revealed that the previous finding (i.e. more shrubs under open canopies) is true for both young and mature forests (Fig. 5 top). In young and mature forest plots, which reported less than 40% ground returns, the presence of shrubs (as measured by the proportion of shrub presences) decreased with decreasing amounts of ground LiDAR returns. In other words, forest tree canopies that intercepted fewer LiDAR pulses tended to have more shrubs than those that intercepted more pulses, a result agreeing with the ecological findings of Kilina et al. (1996) and Bartemucci et al. (2006). Information about the percentage of ground returns might also capture variations in shrub cover due to management practices, which is another factor influencing the distribution of shrubs in forested landscapes (Kerns and Ohmann, 2004). In plots with high values of ground returns ($\geq 40\%$), and consequently with very low or absent tree canopy cover, shrubs were primarily absent, contrary to the general expectation (see Fig 5 top). This region of the CDF plot includes, among others, the 9 open area plots, of which 8 have no shrubs, and most of the stand initiation plots, half of which have no shrubs either. Management practices typically prevent the development of shrubs in areas with trees recently planted or to maintain the open areas in grasslands. Kerns and Ohmann (2004) found similar responses in the coastal forests of Oregon, with open forests supporting lower than expected shrub cover, due to forest management.

Topography is another variable known to influence the presence of shrubs in temperate forests (Gracia et al., 2007). The variable SCOSA reflects topographic positions based on slope and aspect simultaneously. The CDF plot revealed that northern

aspects and steeper slopes (i.e. high SCOSA values) supported fewer shrubs than southern aspects and gentler slopes (lower SCOSA values) (Fig. 5 center). In the region of our study, northern aspects are colder and steep slopes are drier, thus less suitable for the development of broad-leaf understory shrubs.

Finally, the variable STRATUM2 corresponds to the percent of vegetation returns between 1m and 2.5m height, a range where interaction of the laser pulses with the shrub layer can be expected. According to the field data, the shrub cover was typically less than 2m in height. The CDF revealed that most of the field plots (72 of 83) in the study area have less than 20% of the vegetation returns in the STRATUM2 layer, and within these plots, those with more returns in the STRATUM2 layer tend to have more shrubs (as measured by the proportion of shrub presences) (Fig. 5 bottom). This supports the idea that shrubs are contributing to many of the returns found in this layer. However, there were cases with no true shrubs in the understory but still high a proportion of returns within the STRATUM2 layer, indicating the presence of other components in the understory (e.g. saplings and lower branches of small trees). In the CDF plot, for example, the peak in the absences observed in STRATUM2 values of around 30% was caused by plots of stand initiation and understory reinitiation, which have low tree cover with high understory cover composed only by conifers.

One of the major challenges for characterizing understory vegetation with LiDAR data is that increased forest cover reduces the chances of detecting understory returns (Goodwin, 2006, 2007), which complicates applications in areas containing dense canopies. In addition, not all understory “shrubby” vegetation is equally relevant for wildlife species, particularly in Pacific Northwest coniferous forests (Hagar, 2007). The use of LiDAR data allowed us to map the distribution of understory shrub species by quantifying not only vegetation returns from the lower strata of the canopy forest (i.e. where understory shrubs occur), but also ecological variables (e.g. tree canopy cover and topography) that are known to influence the distribution and abundance of understory vegetation in closed and open forests. We found that the errors in our understory shrub map were distributed in similar proportions under open and closed canopies (as reflected

by the canopy density metric, using a threshold of 50% to separate low vs. high), suggesting that the model was useful under both open and dense tree canopies.

The results of this study indicated that the median absolute deviation of height LiDAR returns (HMAD) is an important variable for predicting the distribution of different diameter classes of snags, as it was the most common variable selected in the models. Bater (2007) found that a similar LiDAR-derived measure of canopy variation, the log-transformed coefficient of variation of heights, was a significant predictor of the proportion of trees in different stages of decay. Clark et al. (2004) and Bater (2007) suggested that canopies became more structurally complex (or variable) partly because of the presence of snags. Our CDF plot revealed that the presence of snags increased with increasing canopy complexity as reflected by HMAD (Fig. 6). The findings of this study reinforce the notion that LiDAR-derived measures of variation in canopy height are valuable for characterizing the distribution of snags and trees in different stages of decay, whether in terms of overall abundance (Bater 2007), or in terms of abundance of different diameter classes (this study). Pesonen et al. (2008) found that the log-transformed coefficient of variation of heights from LiDAR was also a significant predictor of downed woody debris. In addition, while the study of Bater (2007) was conducted in a flat area, our study was conducted in complex topography. Topography is a common factor influencing the abundance of snags and dead woody material in mountainous areas (Flanagan et al. 2002; Kennedy et al., 2008), and it appeared to be important in our study area as most of the models included LiDAR-derived topographic variables. For example, we observed a higher proportion of samples with snags in areas with more exposed, convex terrain (positive BOLSTAD values) than under more protected, concave terrain (negative BOLSTAD values) (data not shown). Similar to the shrubs, the use of LiDAR metrics allowed us to quantify structural variables that are known to indirectly indicate the presence of snags, as well as environmental variables that are known to influence their presence and distribution in forests. Finally, we found that the incorporation of information about forest succession (derived also from LiDAR; Falkowski et al., 2009) improved the accuracy of the predictive distribution for the different snag diameter classes. The age of the stand can be a natural indicator of the potential diameter of the

snags found in the forests (typically, the older the forests, the larger the snags). We found that the segmentation based on succession produced areas with different HMAD values (e.g. HMAD for MMS = 9.2 ± 3.5 ; HMAD for YMS&UR = 3.3 ± 3.4), effectively reducing variation in relevant data, for a better classification. The accuracy of the different snag diameter classes was slightly higher in the MMS&OMS successional area (88%) than in the YMS&UR area (77% to 82%).

In summary, for snags and understory shrub mapping, the value of LiDAR data resided in the ability to quantify 1) structural metrics that are known to directly or indirectly indicate the presence of understory shrubs and snags, such as the percent of vegetation returns in the lower strata of the canopy (for the shrubs) and the vertical heterogeneity of the forest canopy (for the snags), and 2) ecological variables that are known to influence the distribution and abundance of understory vegetation and snags in temperate mountainous forests (e.g. canopy cover, topography, forest succession).

The RF algorithm (Breiman, 2001) played an important role in these findings, as it allowed us to identify those relevant predictor variables for understory shrub and snags mapping, and integrate them in a predictive mapping approach. Similar to Cutler et al. (2007) and Falkowski et al. (2009), we found that the variables identified by RF agreed with the expectations based on the literature, making good intuitive sense in how the variables relate to the ecological processes governing snag and understory shrub distribution, and highlighting the value of the RF algorithm for ecological modeling using remote sensing data.

For wildlife habitat suitability assessment, the value of LiDAR data resided in its ability to derive a variety of habitat variables related to forest 3-D structure, which are known to be important for wildlife species, but have been difficult or impossible to derive from other remote sensing technologies (see Vierling et al. 2008). In this sense, we were able to map habitat suitability for avian species that depend on a broad variety of forest structural conditions (including those related to understory vegetation, snag size and density, tree canopy cover, basal area, etc.). These findings are important for advancing the management of biodiversity and wildlife habitat in forests (Russell et al., 2007;

Venier and Pearce, 2007), biodiversity applications of remote sensing (Turner et al. 2003), and species distribution modeling (Guisan and Zimmermann, 2000). For instance, the lack of maps of understory shrub and snag distribution has been identified as a major limitation for managing wildlife habitat in forests (Russell et al., 2007; Venier and Pearce, 2007). In our case, we mapped understory shrubs from an initial presence/absence approach. While the availability of a simple presence/absence layer can make a difference in assessing wildlife habitat suitability (see for example the Giant Panda's case in Linderman et al. 2005), further efforts should evaluate the capabilities of LiDAR data to derive continuous estimates of understory shrub cover at the plot level and below any type of overstory condition. Goodwin (2006, 2007) suggested that increasing the plot size might serve to detect more returns from the understory. In a recent study, Korpela (2008) found that calibrated LiDAR intensity data is sensitive to understory, ground-surface composition. Because true shrubs and saplings found in the understory of Inland Northwest forests are compositionally and functionally different (i.e. non-coniferous vs. coniferous), there is potential in the use of calibrated intensity data for understory characterization (but see Su and Bork, 2007). Because the intensity data were not calibrated, we chose not to include intensity information in our study.

In terms of snags, it is important to expand the LiDAR-based mapping approach to other diameter classes. We focused on common snags used by some species of woodpecker, but larger snag classes (e.g. ≥ 50 cm DBH) are also of critical interest for wildlife and biodiversity assessment (Davis 1983). However, especially in highly managed/harvested forests, these snag classes can be rare (see Figure 1), and thus, working with them may require other sampling approaches (see for example Bate et al., 2002) and/or dealing with heavily imbalanced data that can present challenges for classification (Chen et al., 2004). The use of high-spatial resolution, color infrared aerial photos has proved useful for mapping the distribution of large snags in a forest by photo interpretation (Bütler and Schlaepfer, 2004). The integration of high-density LiDAR data and high spatial and spectral resolution imagery might facilitate the mapping of more snag classes, as well as the identification of patches dominated by dead trees (see Swatantran et al., 2008).

LiDAR data have proved useful for assessing wildlife habitat in forests (see review by Vierling et al., 2008), including in our study. However, these efforts encompass a small total number of species and have been conducted at a relatively small spatial scale. A similar spatial situation can be found in analyses of biodiversity/species richness with LiDAR (see Goetz et al., 2007 and Clawges et al., 2008). As a result, expanding the applications to other areas and to other organisms with different habitat requirements is highly desired. In addition, it is also important to evaluate which type of LiDAR information (i.e., rough data, metrics and/or variables) are needed to support wildlife habitat suitability and biodiversity assessments. For instance, Bater (2008) evaluated which of the known indicators of forest biodiversity in Canada can be derived from LiDAR, and Martinuzzi et al. (2009) did the same for habitat variables that are needed for refining predictions of species distribution by the US Gap Analysis Program. Expanding the applications of LiDAR remote sensing for wildlife habitat and biodiversity assessments should be feasible considering the increasing availability of LiDAR data. Furthermore, these efforts are particularly relevant for evaluating the potential of future large-scale LiDAR acquisitions, such as those related to the US National LiDAR Initiative (Stoker et al., 2008) or the NASA's planned DESDynI (Deformation, Ecosystem Structure and Dynamics of Ice) mission (<http://desdyni.jpl.nasa.gov/>).

Acknowledgements

This work was made possible through funding by the USGS Gap Analysis Program and the USDA Forest Service International Institute of Tropical Forestry (IITF). The LiDAR acquisition, processing, and analysis were supported by the Agenda 2020 program, a research partnership between the USFS Rocky Mountain Research Station, Bennett Lumber Products, Inc., and Potlatch Forest Holdings, Inc., and the University of Idaho. We thank G. Roloff for facilitating the HSI model for the dusky flycatcher, R. Nelson for commenting on an earlier version of this manuscript, N. Crookston and R. Díaz-Uriarte for fruitful discussions about Random Forest and varSelRF, and three anonymous reviewers. Work at IITF is done in collaboration with the University of Puerto Rico.

References

- Alluoché, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, 43, 1223-1232
- Bartemucci, P., Messier, C, and Canham, C.D. (2006). Overstory influences on light attenuation patterns and understory plant community diversity and composition in southern boreal forests of Quebec. *Canadian Journal of Forest Research*, 36, 2065-2079
- Bate, L.S., Garton, E.O., & Wisdom, M.J. (2002). Sampling methods for snags and large trees important for wildlife. USDA Forest Service General Technical Report PSW-GTR-181
- Bater, C.W., Coops, N.C., Gergel, S.E., & Goodwin, N.R. (2007). Towards the estimation of tree structural class in northwest coastal forest using LiDAR remote sensing. Proceedings of the ISPR Workshop “Laser Scanning 2007 and SilviLaser 2007”, Espoo, September 12-14, 2007, Finland. ISSN 1682-1777, Volume XXXVI, P3/W52, pages 38-43
- Bolstad, P.V. & Lillesand, T.M. (1992). Improved classification of forest vegetation in Northern Wisconsin through rule-based combination of soils, terrain, and Landsat TM data. *Forest Science*, 38, 5-20
- Bater, C.W. (2008). Assessing Indicators of Forest Sustainability Using LiDAR Remote Sensing (Thesis). University of British Columbia, Vancouver, Canada (97 pp.)
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5-32
- Brokaw, N.V., & Lent, R.A. (1999). Vertical Structure. In: M. Hunter (Ed.), *Maintaining Biodiversity in Forest Ecosystems* (pp. 373-399). Cambridge University Press, Cambridge
- Broughton, R.K., Hinsley, S.A., Bellamy, P.E., Hill, R., Rothery, P. (2006). Marsh Tit *Poecile palustris* territories in a British broad-leaved wood. *Ibis*, 148, 744-752.
- Bütler, R., & Schlaepfer, R. (2004). Spruce snag quantification by coupling colour infrared aerial photos and a GIS. *Forest Ecology and Management*, 195, 325-339

- Chen, *Larix occidentalis* C., Liaw, A., Breiman, L. (2004), Using random forest to learn imbalanced data. Technical Report 666, Statistics Department, University of California at Berkeley
- Clark, D.B., Castro, C.S., Alvarado, L.D.A., & Read, J.M. (2004). Quantifying mortality of tropical rain forest trees using high-spatial-resolution satellite data. *Ecology Letters*, 7, 52-59
- Clawges, R., Vierling, L.A., Calhoun, M., Toomey, M.P. (2007). Use of a ground-based scanning lidar for estimation of biophysical properties of western larch (. *International Journal of Remote Sensing*, 28, 4331-4344
- Clawges, R., Vierling, K., Vierling, L., Rowell, E. (2008). The use of airborne lidar to assess avian species diversity, density, and occurrence in a pine/aspen forest. *Remote Sensing of Environment*, 112, 2064-2073
- Cline, S.P., Berg, A.B., Wight, H.M. (1980). Snag characteristics and dynamics in Douglas-fir forests, western Oregon. *Journal of Wildlife Management*, 44, 773-786
- Congalton, R. & Green, K. (1999). *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. CRC/Lewis Press, Boca Raton, FL. 137 p.
- Cooper, S.V., Neiman, K.E., Roberts, D.W. (1991). Forest habitat types of Northern Idaho: a second approximation. USDA Forest Service IRS, GTR-INT-236, 143 pp.
- Crookston, N.L., & Finley, A.O. (2008). yaImpute: An R package for KNN imputation. *Journal of Statistical Software*, 23(10), 16p.
- Cutler, R.D., Edwards, T.C. Jr., Beard, K. H., Cutler, A., Hess. K.T., Gibson, J., & Lawler, J.J. (2007). Random Forests for classification in ecology. *Ecology*, 88, 2783-2792
- Davis, J.W. (1983). Snags are for wildlife. In: J.W. Davis, G.A. Goodwin, & R.A. Ockenfels (Eds.). *Proceedings of the Symposium on Snag Habitat Management* (pp. 4-9). USDA Forest Service General Technical Report RM-99, Flagstaff, AZ
- Díaz-Uriarte, R., and Alvarez, S. (2006). Gene selection and classification of microarray data using random forest. *BMC Bioinformatics*, 7, 3.
- Díaz-Uriarte, 2008. The varSelRF Package. 23 pp. Available at: <http://cran.r-project.org/web/packages/varSelRF/varSelRF.pdf>

- Drake, J. Dubayah, R.O., Clark, D.A., Knox, R.G., Blair, B., Hofton, M., Chazdon, R.L., Weishampel, J.F., & Prince, S.D. (2002). Estimation of forest structural characteristics using large-footprint lidar. *Remote Sensing of Environment*, 79, 305-319
- Edenius, L., & Mikusiński, G. (2006). Utility of habitat suitability models as biodiversity assessment tools in forest management. *Scandinavian Journal of Forest Research*, 21, 62-72
- Evans, I.S. (1972). General geomorphometry, derivatives of altitude, and descriptive statistics. In Chorley, R.J., *Spatial Analysis in Geomorphology*, Now York, Karper & Raw pp. 17-90
- Evans, J.S. & Hudak, A.T. (2007). A multiscale curvature algorithm for classifying discrete return LiDAR in forested environments. *IEE Transactions on Geoscience and Remote Sensing*, 45, 1029-1038
- Falkowski, M.J., Gessler, P.E., Morgan, P., Hudak, A.T. and Smith, A.M.S. (2005). Evaluating ASTER satellite imagery and gradient modeling for mapping and characterizing wildland fire fuels. *Forest Ecology and Management*, 217, 129-146
- Falkowski, M.J., Evans, J.S., Martinuzzi, S, Gessler, P.E., & Hudak, A.T. (2009). Characterizing Forest Succession with Lidar Data: an Evaluation for the Inland Northwest USA. *Remote Sensing of Environment*, 113, 946-956
- Flanagan, P.T., Morgan, P., & Everett, R.L. (2002). Snag Recruitment in Subalpine Forest of the North Cascades, Washington State. USDA Forest Service General Technical Report PSW-DTR-181
- Frescino, T.S., Edwards, T.C., & Moisen, G.G. (2001). Spatially Explicit Forest Structural Attributes Using Generalized Additive Models. *Journal of Vegetation Science*, 12, 15-26
- Fu, W.J., Raymond, J.C., & Wand, S. 2005. Estimating misclassification error with small samples via bootstrap cross-validation. *Bioinformatics*, 21, 1979-1986
- Ganey, J.L. (1999). Snag density and composition of snag populations in two National Forests in northern Arizona. *Forest Ecology and Management* 117, 169-178

- Goetz, S., Steinberg, D., Dubayah, R., Blair, B. (2007). Laser remote sensing of canopy habitat heterogeneity as a predictor of bird species richness in an eastern temperate forest, USA. *Remote Sensing of Environment*, 108, 254-263
- Goodwin, N.R. (2006). *Assessing Understorey Structural Characteristics in Eucalyptus Forests: an investigation of LiDAR techniques* (Thesis). University of New South Wales, Sydney NSW Australia (206 pp.)
- Goodwin, N.R., Coops, N.C., Bater, C, Gergel, S.E. (2007). Assessment of sub-canopy structure in a complex coniferous forest. *Proceedings of the ISPR Workshop "Laser Scanning 2007 and SilviLaser 2007"*, Espoo, September 12-14, 2007, Finland. ISSN 1682-1777, Volume XXXVI, P3/W52, pages 169-172
- Gracia, M., Montané, F., Piqué, J., Retana, J. (2007). Overstorey structure and topographic gradients determining diversity and abundance of understory shrub species in temperate forests in central Pyrenees (Spain). *Forest Ecology and Management*, 242, 391-397
- Graf, R.F., Mathys, L., & Bollmann, K. (2009). Habitat assessment for forest dwelling species using LiDAR remote sensing: Capercaillie in the Alps. *Forest Ecology and Management*, 257, 160-167.
- Guisan, A., & Zimmerman, N.E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135, 147-186
- Hagar, J.C. (2007). Wildlife species associated with non-coniferous vegetation in Pacific Northwest forests: A review. *Forest Ecology and Management*, 246, 108-122.
- Harding, D.J., Lefksy, M.A., Parker, G.G., Blair, J.B. (2001). Laser altimeter canopy height profiles: methods and validations for closed-canopy, broadleaf forests. *Remote Sensing of Environment*, 76, 283-297
- Hill, R.A., Broughton, R.K. (2009). Mapping understorey from leaf-on and leaf-off airborne LiDAR data of deciduous woodland. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64, 223-233.
- Hinsley, S.A., Hill, R.A., Gaveau, D.L.A., Bellampy, P.E. (2002). Quantifying woodland structure and habitat quality for birds using airborne laser scanning. *Functional Ecology*, 16, 851-857

- Hinsley, S.A., Hill, R.A., Bellamy, P.E., Harrison, N.M., Speakman, J.R., Wilson, A.K., Ferns, P.N. (2008). Effects of structural and functional habitat gaps on breeding woodland birds: working harder for less. *Landscape Ecology*, 23, 615-626
- Hofton, M.A., Rocchio, L.E., Blair, J.B., & Dubayah, R. (2002). Validation of vegetation canopy lidar sub-canopy topography measurements for a dense tropical forest. *Journal of Geodynamics*, 34, 491-502
- Huang, C., Homer, C., & Yang, L. (2003). Regional forest land cover characterization using Landsat type data. In M. Wulder & M. Franklin (Eds.) *Methods and Applications for Remote Sensing of Forests: Concepts and Case Studies* (pp. 389-410). Kluwer, Academic Publishers.
- Hudak, A.T., Crookston, N.L., Evans, J.S., Falkowski, M.J. Smith, A.M.S., Gessler, P.E., Morgan, P. (2006). Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return lidar and multispectral satellite data. *Canadian Journal of Remote Sensing*, 32, 126-138
- Hudak, A.T., Crookston, N.L., Evans, J.S., Hall, D.E., & Falkowski, M.J. (2008a). Nearest neighbor imputation modeling of species-level, plot-scale structural attributes from LiDAR data. *Remote Sensing of Environment*, 112, 2232-2245
- Hudak, A.T., Evans, J.S., Crookston, N.L., Falkowski, M.J., Steigers, B., Tylor, R. and Hemmingway, H. (2008b). Aggregating pixel-level basal area predictions derived from LiDAR data to industrial forest stands in Idaho. In: R.N. Havis and N.L. Crookston (Comps). *Third Forest Vegetation Simulator Conference 2007*. USDA Forest Service, Rocky Mountain Research Station; Proceedings RMRS-P-54, pp. 133-146 (refereed)
- Hyde, P., Dubayah, R., Peterson, B., Blair, J.B., Hofton, M., Hunsaker, C., Knox, R., & Walker, W. (2005). Mapping forest structure for wildlife habitat analysis using waveform lidar: Validation of montane ecosystems. *Remote Sensing of Environment*, 96, 427-437
- Hyde, P., Dubayah, R., Walker, W., Blair, J.B., Hofton, M., Hunsaker, C. (2006). Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSar, ETM+, Quickbird) synergy. *Remote Sensing of Environment*, 102, 63-73

- Kennedy, R.S.H, Spies, T.A., & Gregory, M.J. (2008). Relationships of dead wood patterns with biophysical characteristics and ownership according to scale in Coastal Oregon, USA. *Landscape Ecology*, 23, 55-68.
- Kerns B.K, & Ohmann, J.L. (2004). Evaluation and prediction of shrub cover in coastal Oregon forests (USA). *Ecological Indicators*, 4, 83-98
- Kerr, J.T., & Ostrovsky, M. (2003). From space to species: ecological applications for remote sensing. *Trends in Ecology and Evolution*, 18, 299-305
- Kilina, AK, Chen, H.Y.H., Wang, Q., & Montigny, L. (1996). Forest canopies and their influence of understory vegetation in early seral stands on west Vancouver island. *Northwest Science*, 70, 193-200
- Korpela, I.S. (2008). Mapping of understory lichens with airborne discrete-return LiDAR data. *Remote Sensing of Environment*, 112, 3891-3897
- Korol, J.J., Hemstrom, M.A., Hann, W.J. & Gravenmier, R.A. (2002). Snags and Down Wood in the Interior Columbia Basin Ecosystem Management Project. USDA Forest Service General Technical Report PSW-GTR-181
- Kroll, A.J., & Haufle, J.B. (2006). Development and evaluation of habitat suitability models at multiple spatial scale: A case study with the dusky flycatcher. *Forest Ecology and Management*, 229, 161-169
- Lawrence, R.L., Wood, S.D., & Sheley, R.L. (2006). Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (RandomForest). *Remote Sensing of Environment*, 100, 356-362
- Lefsky, M.A., Cohen, W.B., Harding, D.J., & Parker, G.G. (2002). Lidar remote sensing for forest ecosystem studies. *BioScience*, 52, 19-30
- Liaw, A., & Wiener, M. (2002). Classification and regression by Random-Forest. *R News*, 2, 18-22
- Linderman, M., Beaur, S., An, L., Tan, Y.C., Ouyang, Z.Y., Liu, H.G. (2005). The effects of understory bamboo or broad-scale estimated of giant panda habitat. *Biological Conservation*, 121, 383-390
- MacArthur, R.H., & MacArthur J.W. (1961). On Bird Species Diversity. *Ecology*, 42, 594-598

- Maltamo, M., Packalén, P., Yu, X., Eerikäinen, K., Hyyppä, J., Pitkänen, J. (2005). Identifying and quantifying structural characteristics of heterogeneous boreal forests using laser scanner data. *Forest Ecology and Management*, 216, 41-50
- Martinuzzi, S., Vierling, L., Gould, W., & Vierling, K. (2009). Improving the characterization and mapping of wildlife habitats with LiDAR data: measurement priorities for the Inland Northwest, USA. In: J. Maxwell et al., (Editors). *Gap Analysis Bulletin No. 16*. USGS/BRD/Gap Analysis Program, Moscow, ID, USA.
- McDermid, G.J. Franklin, S.E., LeDrew, E.F. (2005). Remote sensing for large-area habitat mapping. *Progress in Physical Geography*, 29, 449-474
- McKenzie, D. & Halpern, C.B. (1999). Modeling the distributions of shrub species in Pacific Northwest forests. *Forest Ecology and Management*, 114, 293-307
- Moore, I.D., Gessler, P.E., Nielsen, G.A., Petersen, G.A. (1993). Terrain attributes: estimation methods and scale effects. In A.J. Jakeman, M.B. Beck, & M. McAleer (Eds) *Modeling Change in Environmental Systems*, London: Wiley, pp. 189-214
- Morrison, M.L., & Raphael, M.G. (1993). Modeling the dynamics of Snags. *Ecological Applications*, 3, 322-330
- Murphy, M.A., Evans, J.S., Storfer. (2009). Quantifying Bufo bores connectivity in Yellowstone National Park with landscape genetics. *Ecology*, in press.
- Nelson, R.F., Krabill, W.B., & Tonelli, J. (1988). Estimating forest biomass and volume using airborne laser data. *Remote Sensing of Environment*, 24, 247-267
- Nelson, R., Keller, C., & Ratnaswamy, M. (2005). Locating and estimating the extent of delmarva fox squirrel habitat using an airborne LiDAR profiler. *Remote Sensing of Environment*, 96, 292-301
- Noss, R.F. (1999). Assessing and monitoring forest biodiversity: A suggested framework and indicators. *Forest Ecology and Management*, 115, 135-146
- Ohmann, J.L., McComb, W.C., Zunrawi, A. (1994). Snag abundance for primary cavity-nesting birds on nonfederal forest lands in Oregon and Washington. *Wildlife Society Bulletin*, 22, 607-620
- Pesonen, A., Maltamo, M., Eerikäinen, & Packalen, P. (2008). Airborne laser scanning-based prediction of coarse woody debris volumes in a conservation area. *Forest Ecology and Management*, 255, 3288-3296

- Prasad, A.M., Iverson, L.R., & Liaw, A. (2006). Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems*, 9, 181-199
- Prosser, D.J., & Brooks, R.P. (1998). A verified habitat suitability index for the Louisiana Waterthrush. *Journal of Field Ecology*, 69, 288-298
- R Development Core Team. (2005). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Riaño, D., Meier, E., Allgöwer, B., Chuvieco, E., Ustin, S.L. (2003). Modelling airborne laser scanning data for the spatial generation of critical forest parameters in fire behaviour modelling. *Remote Sensing of Environment*, 86, 177-186
- Riley, S.J., DeGloria, S.D., Elliot, R. (1999). A terrain ruggedness index that quantifies topographic heterogeneity. *Intermountain Journal of Science*, 5, 1-4
- Roloff, G.F. (2001). Habitat suitability model for dusky flycatcher (*Empidonax oberholseri*) in the intermountain West. Timberland Resources, Boise Cascade Corporation, Boise, ID, 6pp.
- Ruiz-Gazen, A., & Villa. (2007). Storms prediction: Logistic regression vs. random forest for unbalanced data. *Case Studies in Business, Industry and Government Statistics 1*: 91-101
- Russell, R.E., Saab, V.A., Dudley, J. (2007). Habitat suitability models for cavity-nesting birds in a postfire landscape. *The Journal of Wildlife Management*, 71, 2600-2611
- Sampson, R.N. & Adams, D.L. (1994). *Assessing forest ecosystem health in the inland west* (461 pp.). Food Products Press, New York
- Schroeder, R.L. (1982). Habitat suitability index models: Downy woodpecker. US Department of the Interior, Fish and Wildlife Service, FWS/OBS/82/10.38. 10pp.
- Scott, J.M., Peterson, C.R., Karl, J.W., Strand, E., Svancara, L.V., & Wright, N.M. (2002). *A GAP Analysis of Idaho: Final Report*. Idaho Cooperative Fish and Wildlife Research Unit. Moscow, ID, USA
- Skowronski, N., Clark, K., Nelson, R., Hom, J., Patterson, M. (2007). Remotely sensed measurements of forest structure and fuel loads in the Pinelands of New Jersey. *Remote Sensing of Environment*, 108, 123-129.

- Souza, P.J. (1982). Habitat suitability index models: Lewis' woodpecker. US Department of the Interior, Fish and Wildlife Service, FWS/OBS/82/10.32. 14pp.
- Souza, P.J. (1987). Habitat suitability index models: hairy woodpecker. US Department of the Interior, Fish and Wildlife Service, Biol. Rep. 82(10.146) 19pp.
- Spiering, D.J., & Knight, R.L. (2005). Snag density and use by cavity-nesting birds in managed stands of the Black Hills National Forest. *Forest Ecology and Management*, 214, 40-52
- Stage, A.R. (1976). An expression of the effects of aspect, slope, and habitat type on tree growth. *Forest Science*, 22, 457-460
- Stoker, J., Harding, D., & Parrish, J. (2008). The need for a national lidar dataset. *Photogrammetric Engineering and Remote Sensing*, 74, 1066-1068
- Strobl, C., Boulesteix, A., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, 8, 25
- Su, J., Bork, E.W. (2007). Characterization of diverse plant communities in Aspen Parkland rangeland using LiDAR data. *Applied Vegetation Science*, 10, 407-416.
- Sun, Y., Kamel, M.S., Wong, A.K.C., Wang, Y. (2007). Cost-sensitive boosting for classification of imbalanced data. *Pattern Recognition*, 40, 3358-3378
- Swatantran, A., Dubayah, R., Hofton, M., Blair, J.B., Handley, L. (2008). Mapping potential Ivory billed woodpecker habitat using LiDAR and hyperspectral data fusion. *Eos Trans. AGU* 89(53), 2008 Fall Meeting Suppl., Abstract B32A-04.
- Tarboton, D.G. (1997). A new method for the determination of flow directions and contributing areas in grid digital elevation models. *Water Resources Research*, 33, 309-319
- Thomas J.W, Anderson, R.G., Maser, C. & Bull, E.L. (1979). Snags. In J.W. Thomas (Ed.). *Wildlife habitats in managed forests: the Blue Mountains of Oregon* Washington (pp. 60-77). US Forest Service Agriculture Handbook 553, Wildlife management Institute, Washington, D.C., USA
- Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E. & Steininger, M. (2003). Remote sensing for biodiversity and conservation. *Trends in Ecology and Evolution*, 18, 306-314

- Turner, M.G., Gardner, R.H., & O'Neill. (2001). *Landscape Ecology in Theory and Practice*. New York: Springer.
- Van Pelt, R. & Franklin, J.F. (2000). Influence of canopy structure on the understory environment in tall, old-growth, coniferous forest. *Canadian Journal of Forest Research*, 30, 1231-1245
- Venier, L.A., & Pearce, J.L. (2007). Boreal forest landbirds in relation to forest composition, structure, and landscape: implications for forest management. *Canadian Journal of Forest Research*, 37, 1214-1226
- Vierling, K.T., Vierling, L.A., Martinuzzi, S., Gould, W., Clawges, R. (2008). Lidar: Shedding new light on habitat modeling. *Frontiers in Ecology and the Environment*, 6, 90-98
- Virkkala, R. (2006). Why study woodpeckers? The significance of woodpeckers in forest ecosystems. *Annales Zoologici Fennici*, 43, 82-85
- Wulder, M. and Franklin, S. (2003). *Remote Sensing of Forest Environments: Concepts and Case Studies* (519 pp.). Kluwer Academic Publishers, Dordrecht / Boston / London
- Zucknick, M., Richardson, S., & Stronach, E.A. (2008). Comparing the Characteristics of Gene Expression profiles Derived by Univariate and Multivariate Classification Methods. *Statistical Applications in Genetics and Molecular Biology*, 7, 7

Table 1. Target avian species and habitat characteristics, based on the corresponding habitat suitability models (i.e. Schroeder, 1982; Souza, 1982, 1987; Roloff 2001).

Species name	Habitat variables and corresponding life requisite	Optimum habitat (i.e. HSI = 1.0)
Dusky flycatcher	Percent tree canopy cover (nesting and foraging) Understory shrub cover (nesting and foraging)	Open forested conditions with a well-developed understory shrub cover
Hairy woodpecker	Number of snags ≥ 25 cm diameter per ha (nesting) Mean diameter of overstory trees (cover and nesting) Percent tree canopy cover (cover)	Mature forest stands with moderate tree canopy cover and at least 5 snags ≥ 25 cm diameter per ha.
Lewis's woodpecker	Percent tree canopy cover (summer food) Understory shrub cover (summer food) Number of snags ≥ 30 cm diameter per ha (nesting)	Open forested conditions with a well-developed understory shrub cover and at least 2.5 snags ≥ 30 cm diameter per ha.
Downy woodpecker	Basal area (food) Number of snags ≥ 15 cm diameter per ha (nesting)	Forest stands with low basal area and more than 13 snags ≥ 15 cm diameter per ha.

Table 2. Understory shrub cover characteristics.

Understory shrub cover	# of plots	(% cover, when present)				
		Mean	St. dev.	Min.	Max.	Median
Present (i.e. > 25% cover)	48	54.0	18.7	25.8	100.0	52.5
Absent (i.e. ≤ 25% cover)	35					

Table 3. Snag distribution within forest successional stages. Inventory data was not available for old-growth forest (only GPS locations)

Successional class	Area (ha)	Plots (%)	Plots (#)	Proportion of plots with snags of different diameter classes (in cm)				
				≥15	≥25	≥30	≥35	≥50
Open	5426	17	9	0.00	0.00	0.00	0.00	0.00
Stand Initiation	3165	10	8	0.00	0.00	0.00	0.00	0.00
Understory Reinitiation	2073	7	6	0.50	0.50	0.33	0.17	0.17
Young Multistory	8727	28	34	0.53	0.41	0.32	0.12	0.06
Mature Multistory	11673	37	26	0.81	0.62	0.50	0.35	0.15
Old Multistory	393	1	NA	NA	NA	NA	NA	NA
Total	31458	100	83					

Table 4. LiDAR-derived metrics of canopy height (top group) and topography (bottom group). The list does not include multi-collinear variables, as they were initially identified and removed using QR-Decomposition (Becker et. al., 1988) in Falkowski et al. (2009). Variables selected for understory shrub and snag map predictions are identified with an “X”. Since snags were mapped using two different approaches (i.e. with and without image segmentation), an additional “X” was added when the variable was selected in both methods.

Metric name	Metric description	Understory shrubs	Snags ≥ 15 cm diameter	Snags ≥ 25 cm diameter	Snags ≥ 30 cm diameter
HMIN	Minimum Height				
HMAX	Maximum Height			X	X
HSKEW	Skewness of Heights				
HKURT	Kurtosis of Heights			X	
HMEAN	Mean Height		X		
HMAD	Median Absolute Deviation of Heights $1.4826 * \text{median}(\text{abs}(x - \text{median}(x)))$		XX	XX	XX
HIQR	Heights interquartile range		X		
H05PCT	Heights 5th Percentile				
H10PCT	Heights 10th Percentile				
H25PCT	Heights 25th Percentile				
HMEDIAN	Median Height				X
CANOPY	Canopy Cover/density (Vegetation Returns/Total Returns * 100)				
STRATUM0	Percentage of Ground Returns = 0 m	X			
STRATUM1	Percentage of Non-Ground Returns > 0 m and ≤ 1 m				
STRATUM2	Percentage of Vegetation Returns > 1 m and ≤ 2.5 m	X			
STRATUM3	Percentage of Vegetation Returns > 2.5 m and ≤ 10 m				
STRATUM4	Percentage of Vegetation Returns > 10 m and ≤ 20 m				X
STRATUM5	Percentage of Vegetation Returns > 20 m and ≤ 30 m			X	
STRATUM6	Percentage of Vegetation Returns > 30 m				
ELEV	Elevation (meters)				
SLP	Slope/Area (Wetness Indicator)				
SLPPCT	Slope in %				
ASPDEG	Aspect in degrees				
CURV	Curvature (Tarboton 1997)			X	
EASTNESS	UTM Easting (meters)			X	X
NORTHNESS	UTM Northing (meters)				
ERR	Elevation to relief ratio (Evans 1972)			X	
SRR	Slope to relief ratio (Evans 1972)				
FLDIST	Flow Distance to Streams (Tarboton 1997)		X	X	
INVCTI	Compound Topographic Index (Moore et al., 1993)			X	X
BOLSTAD	Bolstad's Landform index (Bolstad and Lillesand 1992)			X	XX
SSINA	Percent slope* $\sin(\text{aspect})$ transformation (Stage 1976)				
SCOSA	Percent slope* $\cos(\text{aspect})$ transformation (Stage 1976)	X			
TRI	Topographic ruggedness index (Riley et al., 1999)		X		

Table 5. Minority/majority ratios for the different presence/absence classifications. The number of samples per class (P denotes presence and A denotes absence) used to derive the ratios is shown between parentheses.

	<u>Entire Area (n=83)</u>	<u>YMS&UR (n=40)</u>	<u>MMS&OMS (n=26)</u>
Snags $\geq 15\text{cm}$	0.98 (41A/42P)	0.91 (19A/21P)	0.24 (5A/21P)
Snags $\geq 25\text{cm}$	0.66 (33P/50A)	0.74 (17P/23A)	0.63 (10A/16P)
Snags $\geq 30\text{cm}$	0.50 (26P/57A)	0.48 (13P/27A)	1.00 (13P/13A)
Understory shrubs	0.73 (35A/48P)		

Table 6. Accuracy statistics for the model of understory shrub presence/absence.

Undersory shrub

		<i>Actual Data</i>			Producer's	User's	Omission	Commission
		Present	Absent	Sum	accuracy	accuracy	error	error
<i>Predicted Data</i>	Present	40	6	46	83%	87%	0.17	0.13
	Absent	8	29	37	83%	78%	0.17	0.22
	Sum	48	35	83				

Overall accuracy = 83%; kappa = 0.66; TSS = 0.66

Predictor variables: STRATUM0, SCOSA, STRATUM2

Table 7. Accuracy statistics for the different models of snags presence/absence, including with and without segmentation.

Snags ≥ 15 cm diameter

Model 1 (without forest successional data)

	Class	Actual Data		Sum	Producer's accuracy	User's accuracy	Omission error	Commission error
		Present	Absent					
Predicted Data	Present	34	9	43	81%	79%	0.19	0.21
	Absent	8	32	40	78%	80%	0.22	0.20
	Sum	42	41	83				

Overall accuracy = 80%; kappa = 0.59; TSS = 0.59

Predictor variables: HMAD, HIQR

Model 2 (with forest successional data)

	Class	Actual Data		Sum	Producer's accuracy	User's accuracy	Omission error	Commission error
		Present	Absent					
Predicted Data	Present	37	7	44	88%	84%	0.12	0.16
	Absent	5	34	39	83%	87%	0.17	0.13
	Sum	42	41	83				

Overall accuracy = 86%; kappa = 0.71; TSS = 0.71

Predictor variables: (YMS&UR) HMAD, FLDIST; (MMS&OMS) HMEAN, TRI

Snags ≥ 25 cm diameter

Model 1 (without forest successional data)

	Class	Actual Data		Sum	Producer's accuracy	User's accuracy	Omission error	Commission error
		Present	Absent					
Predicted Data	Present	23	13	36	70%	64%	0.30	0.36
	Absent	10	37	47	74%	79%	0.26	0.21
	Sum	33	50	83				

Overall accuracy = 72%; kappa = 0.43; TSS = 0.44

Predictor variables: HMAD, HMAX

Model 2 (with forest successional data)

	Class	Actual Data		Sum	Producer's accuracy	User's accuracy	Omission error	Commission error
		Present	Absent					
Predicted Data	Present	26	5	31	79%	84%	0.21	0.16
	Absent	7	45	52	90%	87%	0.10	0.13
	Sum	33	50	83				

Overall accuracy = 86%; kappa = 0.70; TSS = 0.69

Predictor variables: (YMS&UR) EASTNESS, FLDIST, HMAD; (MMS&OMS) INVCTI, BOLSTAD, STRATUM5, HKURT, ERR, CURVATURE

Snags ≥ 30 cm diameter

Model 1 (without forest successional data)

	Class	Actual Data		Sum	Producer's accuracy	User's accuracy	Omission error	Commission error
		Present	Absent					
Predicted Data	Present	15	6	21	58%	71%	0.42	0.29
	Absent	11	51	62	89%	82%	0.11	0.18
	Sum	26	57	83				

Overall accuracy = 80%; kappa = 0.50; TSS = 0.47

Predictor variables: BOLSTAD, HMAD, HMAX, HMEDIAN

Model 2 (with forest successional data)

	Class	Actual Data		Sum	Producer's accuracy	User's accuracy	Omission error	Commission error
		Present	Absent					
Predicted Data	Present	19	3	22	73%	86%	0.27	0.14
	Absent	7	54	61	95%	89%	0.05	0.11
	Sum	26	57	83				

Overall accuracy = 88%; kappa = 0.71; TSS = 0.68

Predictor variables: (YMS&UR) STRATUM4, HMAD, EASTNESS; (MMS&OMS) BOLSTAD, INVCTI, EASTNESS

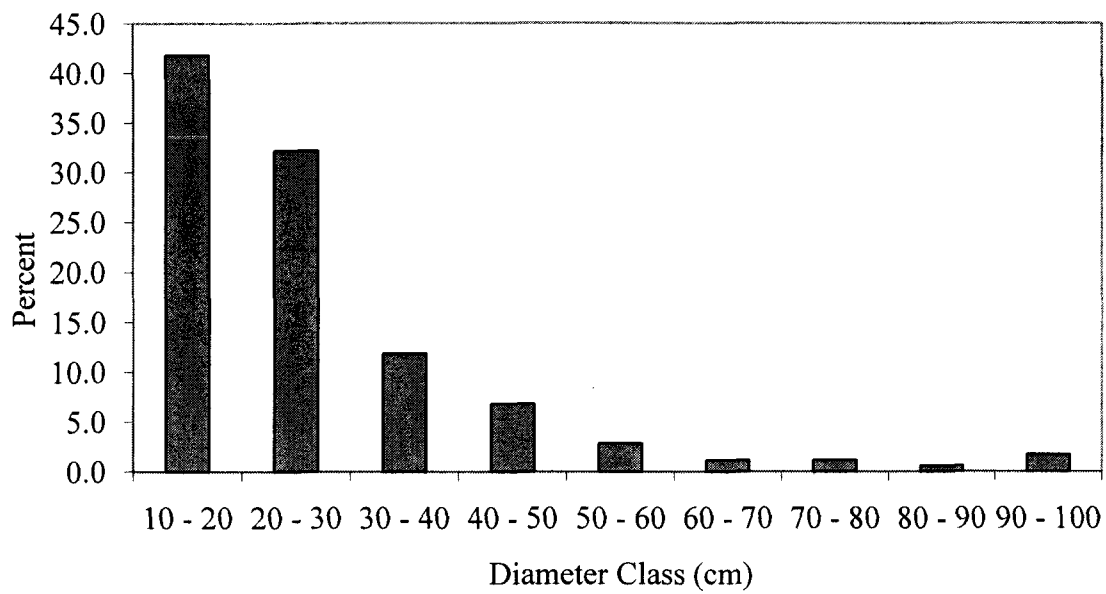


Figure 1. Snag distribution based on diameter.

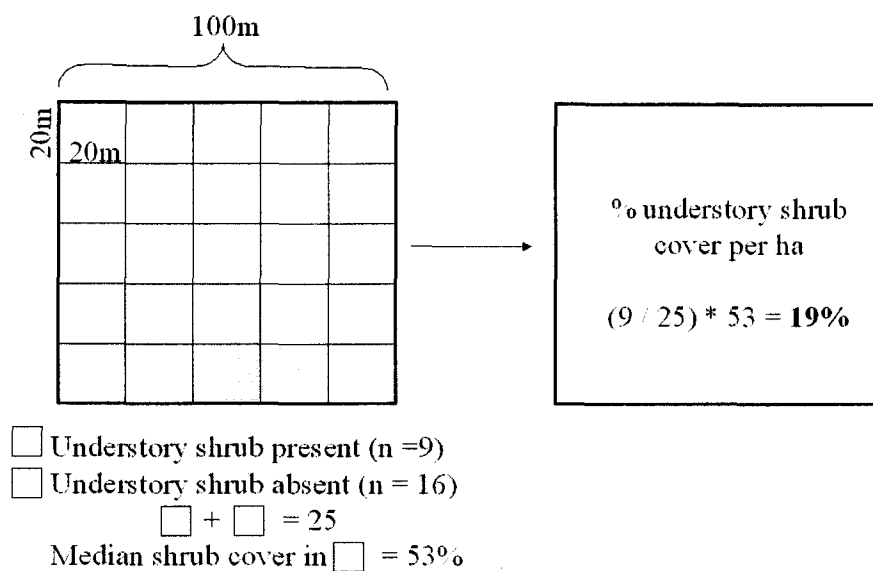


Figure 2. Aggregation approach for converting the 20m-pixel LiDAR-based products of understory shrubs presence/absence into continuous, 100m-pixel (i.e. 1 ha) values. Snags were treated similarly.

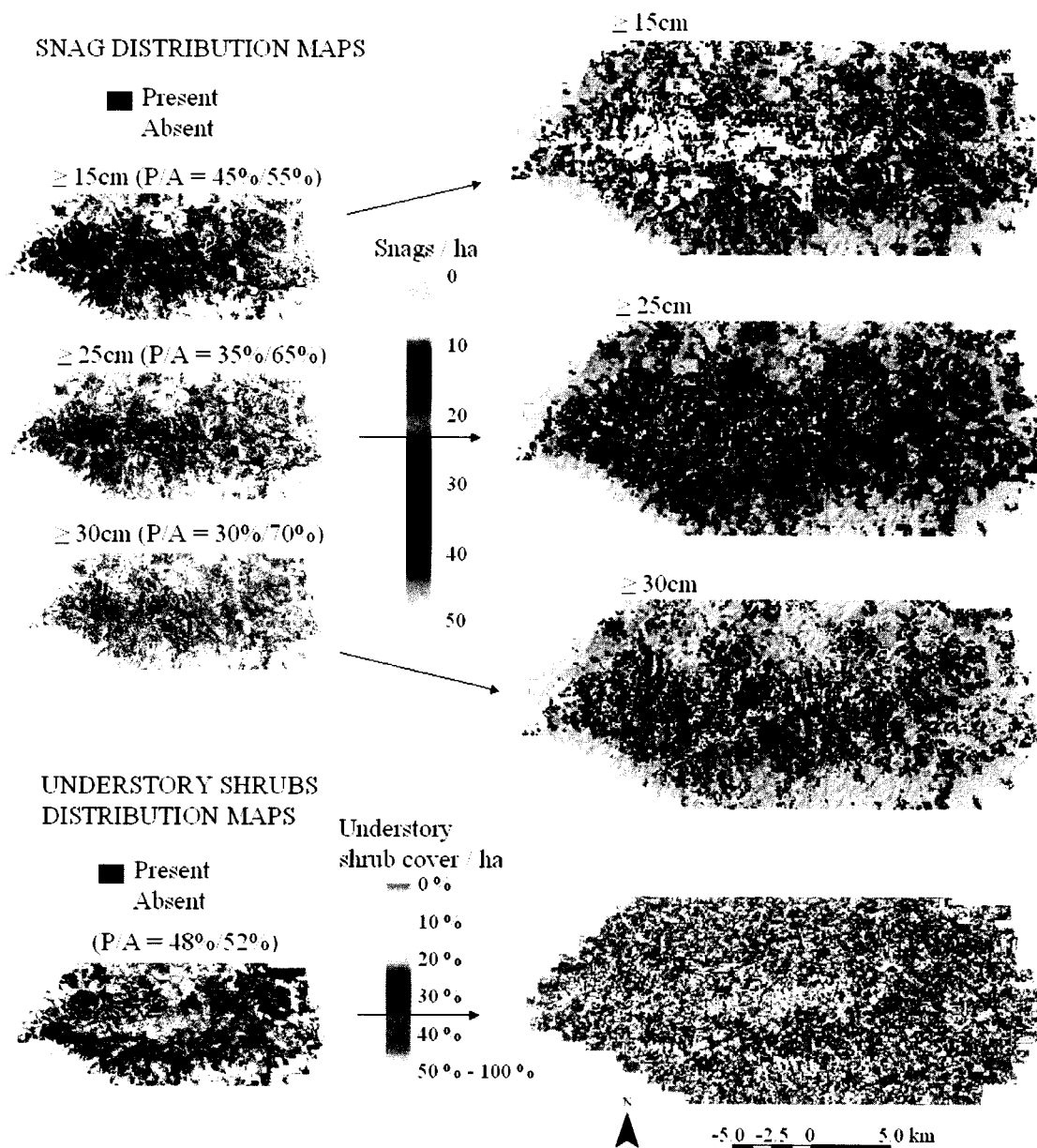


Figure 3. LiDAR-based distribution maps for the different snag diameter classes (top) and understory shrubs (bottom), including the 20m-pixel presence/absence product to the right, and the 1-ha density map to the left. The presence/absence maps include, between parentheses, the proportional cover of the two classes (i.e. present vs. absent)

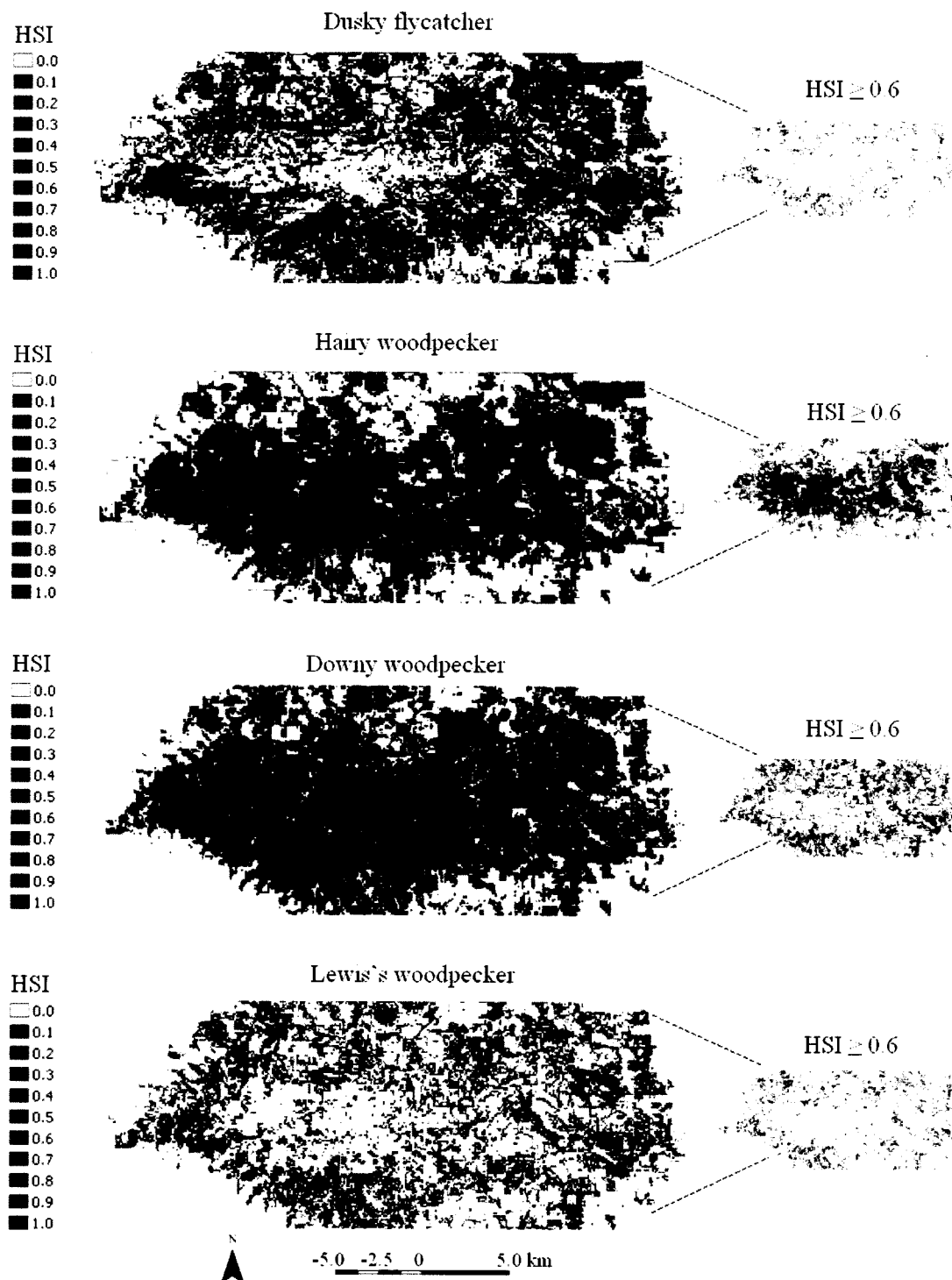


Figure 4. Habitat suitability maps for the different avian species. The maps on the right are simplified, aggregated and recoded versions depicting areas with habitat suitability index (HSI) ≥ 0.6 (i.e. suitable habitats).

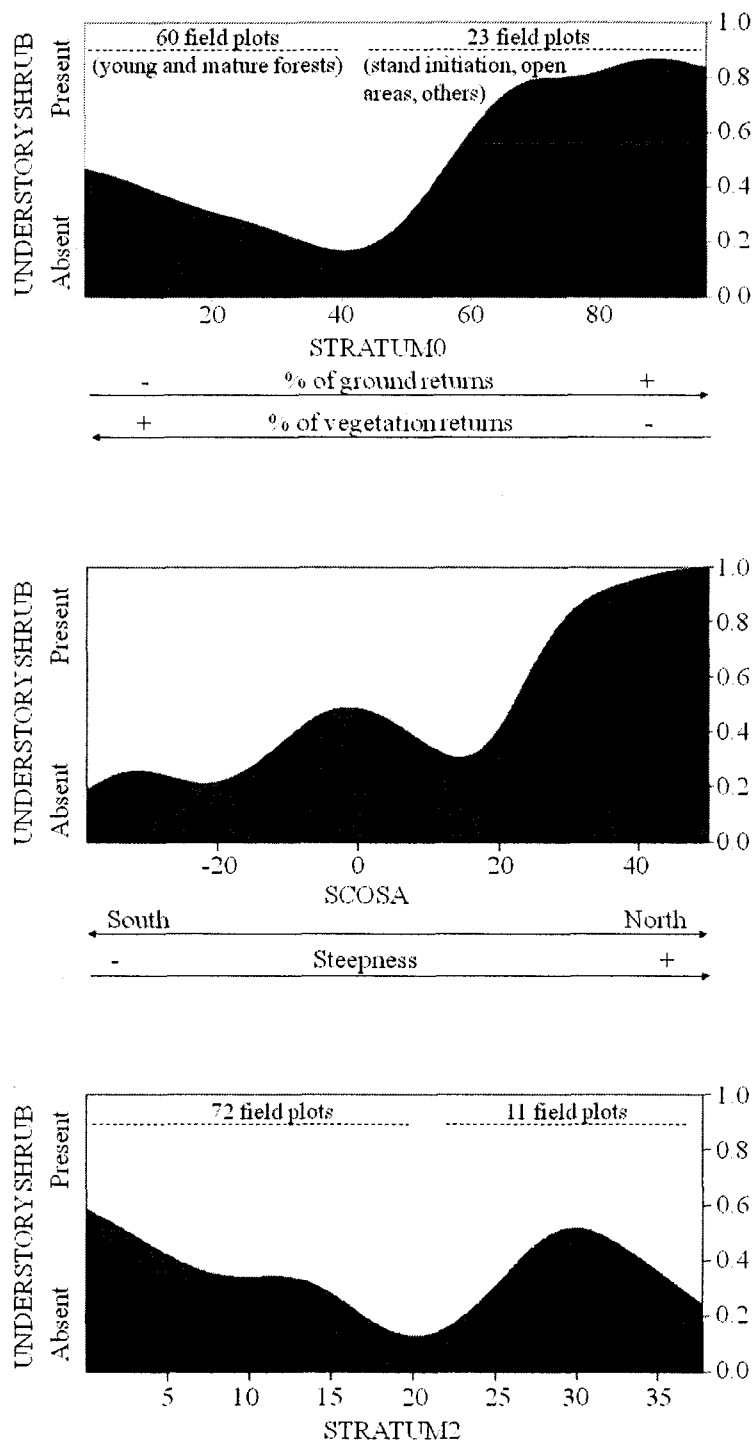


Figure 5. Conditional density plots for the understory shrub distribution vs. the 3 (LiDAR-derived) predictor variables included in the final model.

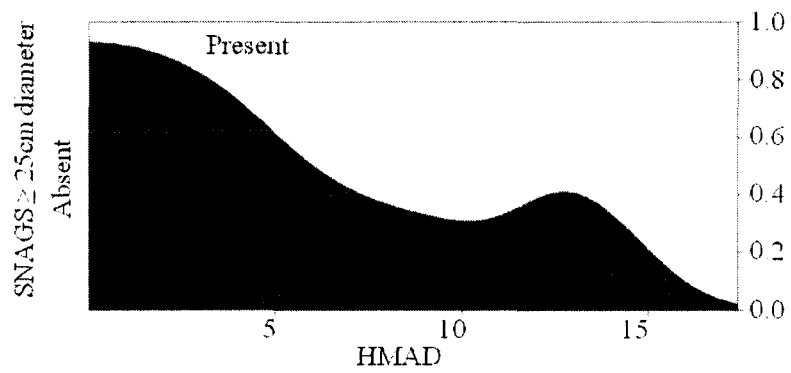


Figure 6. Conditional density plots for the snag distribution vs. the (LiDAR-derived) Median Absolute Deviation of Heights (HMAD). The figure shows the example for the snag diameter class ≥ 25 cm, but a similar trend was observed for the other snag classes.

CHAPTER 3

INCORPORATING REMOTELY SENSED TREE CANOPY COVER DATA INTO BROAD SCALE ASSESSMENTS OF WILDLIFE HABITAT DISTRIBUTION AND CONSERVATION

Note: This chapter has been published as: Martinuzzi, S., L. Vierling, K. Vierling, W. Gould, A. Hudak. 2009. Incorporating remotely sensed tree canopy cover data into broad scale assessments of wildlife habitat distribution and conservation. *Journal of Applied Remote Sensing*, 3, 033568.

Abstract

Remote sensing provides critical information for broad scale assessments of wildlife habitat distribution and conservation. However, such efforts have been typically unable to incorporate information about vegetation structure, a variable important for explaining the distribution of many wildlife species. We evaluated the consequences of incorporating remotely sensed information about horizontal vegetation structure into current assessments of wildlife habitat distribution and conservation. For this, we integrated the new Landsat-derived percent tree canopy cover product (from the NLCD) into the US GAP Analysis database, using avian species and the finished Idaho GAP Analysis as a case study. We found: (1) a 15-68% decrease in the extent of the predicted habitat for avian species associated with specific tree canopy conditions, (2) a marked decrease in the species richness values predicted at the Landsat pixel scale, but not at a coarser (i.e. GAP hexagon) scale, (3) a modified distribution of biodiversity hotspots, and (4) surprising results in conservation assessment: despite the strong changes in the species predicted habitats, their distribution in relation to the reserves network remained the same. This study highlights the value of area wide vegetation structure data for refined biodiversity and conservation analyses. We discuss further opportunities and limitations for the use of the NLCD data in wildlife habitat studies.

Keywords: species distribution model, National Land Cover Database, avian habitat, GAP, horizontal vegetation structure, wildlife conservation.

Introduction

Maps describing the distribution of wildlife species are of great importance for biodiversity and conservation assessments. Because remote sensing provides the only means for measuring a range of habitat characteristics across broad scales, scientists commonly use remote sensing data to model species distribution [1-4]. Specifically, because vegetation characteristics hold great predictive potential for the distribution of wildlife species [5-7], satellite based land cover and vegetation maps are actively used in the modeling process [8-10]. An example of these efforts is the Gap Analysis Program (GAP) in the United States, a major governmental initiative to model the distribution of wildlife species with remote sensing and associated geospatial datasets, with the main purpose of assessing species conservation for the country [11-13]. Furthermore, the GAP approach has been applied worldwide [14,15].

Although land cover maps are considered adequate to derive species distribution models [11,16] they may not adequately represent the relevant vegetation characteristics for many species' habitats. For example, ecologists have long understood that the presence of certain bird and mammal species can be highly dependent on particular conditions of forest structure, such as those related to tree canopy cover [17,18]. However, land cover and vegetation maps typically do not characterize forest structure. If geospatial data used to support habitat models are not adequate to represent the relevant species-environment relationships, the final distribution maps may not match the observed or expected distributions [19-21], affecting subsequent conservation or biodiversity assessments generated from those maps. The lack of accurate, high spatial resolution biophysical data is considered a major limitation to producing more reliable predictions of species distribution [21]. For broad scale modeling efforts such as those

from GAP, detailed information about percent tree canopy cover has been recognized as a major need [22,23].

The recently completed tree canopy cover product of the 2001 National Land Cover Database (NLCD 2001 [24], herein after NLCD_TCC), provides new information about horizontal vegetation structure in the United States, and therefore may serve to fill this important need in wildlife habitat modeling. Originally developed to support land cover requirements for the country, the NLCD_TCC is a nationwide map containing information about the percentage of tree canopy cover at a Landsat spatial resolution (i.e. 30-meter pixel). Evaluating the consequences of incorporating forest structure information into broad scale predictions of species distribution is important given the significance that these maps have for supporting conservation and biodiversity assessments.

In this study, we integrated GAP and NLCD2001 data in order to (1) quantify differences in accuracy of GAP predictions of species distribution given the inclusion of tree canopy cover data (i.e. NLCD_TCC), (2) quantify differences in GAP estimates of species distributions and species richness patterns given the inclusion of tree canopy cover data, and (3) quantify differences in the GAP estimates about the species representation within the network of protected lands.

We addressed these questions using a case study comprised of data from the finished Idaho GAP Analysis (ID-GAP, [22]). Idaho contains a diverse array of ecosystems and environmental gradients, and is therefore a good test bed for understanding the general applicability of these questions. From a total of 238 species of birds that occur in Idaho, the ID-GAP identified 37 species that are known to occur under specific tree canopy cover conditions, equivalent to 1 in every 7 birds species present in the state. The authors indicated that the predicted distribution of these 37 species was likely overestimated because the models did not incorporate tree canopy cover constraints [22], but no formal evaluation was made. Our study is an attempt to evaluate the consequences for GAP assessments brought about by the inclusion of novel remote sensing data of vegetation structure. We worked from a GAP perspective because GAP

projects are developed across the United States (although the approach has been applied internationally), and because data from GAP are actively used in conservation and planning efforts. However, lessons from this study do not relate solely to GAP and/or the United States; they may also help to assess the value of remote sensing products for advancing biodiversity and conservation assessments regardless of geographic location worldwide.

Material and methods

GAP predictions of species distribution and the NLCD2001 tree canopy cover product

GAP predictive maps of species distribution are developed at a State or regional (i.e. multi-State) scale, using a two-step process [25]. First, the species' geographic range is determined by placing the known species occurrences (from GPS points from field surveys, recent museum records, and species lists) in geographic subunits (represented typically by the 635-km² hexagon grid from the Environmental Protection Agency Ecological Mapping and Assessment Program, or EPA-EMAP). About four hundred hexagons are needed, for example, to cover Idaho. Second, information about species-habitat associations (from the scientific literature) is used to identify the suite of (Landsat derived) land cover types, special habitat features (e.g. riparian areas, distance to water bodies, distance to roads) and/or other environmental variables that are suitable for the particular species. The predicted species distribution maps are obtained by intersecting the hexagon based range map with the fine scale habitat requirements. For more information please see the GAP web page <<http://GAPanalysis.nbi.gov>>).

The NLCD_TCC product [24] characterizes nationwide vegetation characteristics for the year 2001. The product was developed using Landsat 7 ETM+ satellite imagery in 66 different mapping zones. Within each zone, multiple digital orthophotoquads (DOQ's) were classified into either tree canopy or non-tree canopy areas at 1-m resolution, and these values were then aggregated to the 30 meter scale to determine the percentage of tree canopy [24]. By combining these DOQ derived training data with Landsat spectral

data and ancillary information, tree canopy cover predictions were developed using regression tree algorithms. Predictions were applied in areas corresponding to deciduous, coniferous, and mixed forests, woody wetlands, and developed open space. A cross-validation procedure reported an accuracy of about 85%. For more information please see the Multi-Resolution Land Characteristics (MRLC) Consortium website <<http://www.mrlc.gov>>.

Study area

Idaho encompasses about 216,000 km² in the northwestern United States. Forests represent about 78,000 km² (~40% of the state), are comprised mostly of coniferous species, and occur principally in the mountainous regions of the north and central part of the state (Fig. 1). The southern portion of Idaho is dominated by sagebrush and shrub-steppe vegetation (33% of the state), and grasslands and agricultural lands (24%). Riparian vegetation, wetlands, and urban areas cover less than 4% [22]. Protected lands (i.e. reserves) represent about 12% of the state. About 70% of the lands in Idaho are public with a majority under US Forest Service management. Excluding riparian areas, forests support the highest wildlife diversity [22]. Forests are subject to a variety of anthropogenic and natural processes that can influence structure and function, such as such as those related to timber extraction, wildfires, blowdowns and landslides.

Data

The data used in this study are of public domain. Data from the ID-GAP were obtained from the GAP server <<http://www.GAPanalysis.nbi.gov>>, including: (1) species geographic range maps, (2) predicted species distribution models and maps, (3) land cover classification map (developed from Landsat imagery from 1996-1998) (see Fig. 1), (4) map of protected areas, (5) ID-GAP Final Report, and (6) metadata. The MRLC Consortium's portal <<http://www.mrlc.gov>> provided the NLCD_TCC coverage for Idaho (zones 01 and 03) (see Fig. 1). We used ArcGIS V9.2 (1999-2006 ESRI Inc.) and ERDAS IMAGINE V9.1 (Leica Geosystems) to process the data.

Approach

The overall approach of this study can be summarized in 3 major steps: (1) identification of the species' tree canopy cover preference, (2) prediction of the species distribution with the new biophysical data (e.g. the NLCD_TCC), and (3) evaluation of species conservation and biodiversity patterns. Steps one and two focused on the 37 bird species identified by the ID-GAP as depending on specific conditions on tree canopy cover (and whose habitats have been probably overestimated due to the lack of such data layers). Step three was conducted at two different levels: one that considered the 37 species, and another that included the entire pool of bird species in Idaho (n=238).

First, we identified the tree canopy preferences for the 37 species of birds using a classification system suggested by the ID-GAP [22], which includes 3 categories: low tree canopy cover ($\leq 40\%$), medium tree canopy cover ($>40\%$ and $\leq 70\%$), and high tree canopy cover ($>70\%$). The ID-GAP provides tree canopy preferences for 20 of the 37 species. We provided information for the other 17 species using (1) internet based scientific reviews, such as The Birds of North America Online <http://bna.birds.cornell.edu/BNA/>, the Point Reyes Bird Observatory (PRBO) Conservation Science <http://www.prbo.org>, and reports from the US Forest Service Timber Management and Wildlife Interactions Project and Fire Effects Information System database <http://www.fs.fed.us/>; (2) recent studies (e.g. [26,27]), and (3) expert opinion.

Second, we refined the ID-GAP species distribution models by adding the NLCD_TCC data. This is equivalent to subtracting from the original ID-GAP species distribution maps those areas that did not meet the species' habitat preferences in terms of tree canopy cover. As a result, we developed 37 new species distribution maps. We assumed no changes in vegetation between 1998 (the year of the ID-GAP data) and 2001 (the year of the NLCD data).

Third, we compared the original (i.e. ID-GAP) and the refined (i.e. with the NLCD_TCC) predicted distribution maps. We evaluated the changes in terms of total

extent of the predicted habitat as well as in the proportion of the predicted habitat that occur within the network of protected lands. We created maps of species richness for the 37 species before and after the NLCD_TCC, as well as for the entire pool of species (n=238). For the entire species pool, we used two different spatial scales of analysis: GAP hexagon (635 km²) and Landsat pixel (900 m²). We compared the new maps of species richness with the original from the ID-GAP in terms of number of species and regional distribution of biodiversity patterns.

Accuracy assessment of the new maps of bird species distribution

We followed the GAP protocol for accuracy assessment [11], using the independent reference data provided by the ID-GAP. GAP uses reference information from locations where high confidence lists of species occurrences have been compiled [25]. Species lists are used because GAP projects develop maps for hundreds of species and over millions of hectares, which makes it impossible to conduct a thorough, field based accuracy assessment of each species map using randomly sampled locations [25]. With this, GAP provides a measure of overall agreement between the predictions and the set of known species locations, and a measure of omission error (failure to predict a species that was present). However, GAP assessments do not provide an estimate of commission errors (prediction of species occurrence in unoccupied area), which is an inherent limitation of GAP [11,25]. In species distribution assessments, commission is more difficult to measure than omission due to the challenges associated with the true and apparent absences in the reference data [28-30]. Although [31] suggested that commission errors can be considered risk-averse for GAP-related purposes, information about both commission and omission errors is ultimately important for species distribution maps used in conservation assessment and planning [19,32]. Finally, if five or fewer reference sites are available for assessing the accuracy of a given species, the accuracy assessment for that species is considered not reliable [25].

The independent reference data (i.e. species list) from the ID-GAP encompasses 62 sites. We calculated the % of correct predictions (CP%) and the % of omissions (OM%) for each of the new 37 species maps, and compared the predictions' accuracy

(i.e. CP% and OM%) before and after the inclusion of tree canopy cover data. Evaluating omission error is important for this study because incorporating tree canopy constraints in the original ID-GAP species-habitat models will likely reduce the extent of the predicted habitats in different amounts.

In addition, we were able to evaluate commission errors. The ID-GAP indicated that the initial distribution of the 37 species was likely overestimated because the models did not incorporate tree canopy cover data [22], but no formal evaluation of the commission error was conducted because of the GAP limitations previously mentioned. We evaluated the magnitude of the initial commission errors by quantifying the changes in the extent of the predicted habitats after adding the tree canopy data. If omission errors are not added after including the tree canopy constraints, any reduction in the predicted habitat will be a consequence of a decrease in original overestimations, and thus, in commission errors. This estimate of commission error is not a result of an accuracy assessment using independent data, but rather is a measure of improvement that arises from interpreting the outputs from the original species-habitat models with the new, more precise ones.

Results

Observed species-habitat associations based on NLCD_TCC and expected distribution patterns

Five major groups of species emerged after evaluating the species-habitat relationships with respect to land over type (simplified to forest/non-forest) and tree canopy cover (i.e. NLCD_TCC) for the 37 avian species (Fig.2). The groups covered a wide range of habitat characteristics; from groups of species that occur both in non-forests and open forests (i.e. forest cover <40%, group 1), to groups of species that occur only in closed forests (i.e. >70% tree canopy cover, group 5). The number of species in each group was variable, with more species in groups associated with forest and non-forest lands (groups 1 and 2) than in groups exclusive from forests (groups 3, 4, and 5). These groupings

provided insights about potential patterns of species richness for these 37 species: including (1) open forest pixels were expected to support more species than closed forest pixels, and (2) incorporating tree canopy cover data might produce small changes in open forest pixels, but relatively larger changes in closed forest pixels (see Fig.2).

Predicted species distribution incorporating NLCD_TCC data

The extent of the predicted species distributions decreased markedly after incorporating the tree canopy cover data. For thirty of the thirty-seven bird species, the new predicted habitat was 15% to 68% smaller than the original predicted by the ID-GAP (Table 1). For species associated with both non-forested and forested lands (i.e. groups 1 and 2), the decrease in predicted habitat became larger as forest affiliation increased. For example, the smallest reductions in habitat ($\leq 5\%$) occurred in six avian species that typically occur in grasslands and/or shrublands but use some forests marginally (e.g. Lark sparrow; Table 1).

Generalist species such as the Common raven, Golden eagle, or Brown headed cowbird, which are known to occur in almost any type of land cover, reported intermediate decreases in habitat size (15% to 30%), while species that utilize forested areas more frequently (yet occasionally use some non-forest lands; e.g. Cedar waxwing) reported the highest changes in habitat size (decreasing between 30% and 68% from the original estimates). Finally, for those species that occur exclusively in forests (groups 3, 4, 5), the changes in the predicted habitat differed depending on the tree canopy preferences. The predicted habitat for species which are thought to occur in forests with tree canopy density $<70\%$ or $>70\%$ showed a similar decrease in habitat (between 40% and 50%), while those species that occur in forests with tree canopy density $>40\%$ showed a smaller habitat reduction (between 15% and 25%).

Model evaluation

The accuracy assessment of the 37 new predictive species distribution models revealed that the incorporation of the tree canopy cover constraints did not result in the addition of omission errors. This was true for all of the species. As a result, neither the percent of

correct predictions (CP%) nor the percentage of omission (OM%) changed after incorporating the tree canopy cover data (Table 2). Because model refinement resulted in habitat reduction without the incorporation of omission errors, the observed changes can be attributed to a decrease in previous commission errors/overestimations. Most of the species models were assessed with 10 to 40 reference sites. Only 3 species were below the ideal minimum of 5 sites (*sensu* Jennings, 2000), and thus, their accuracy assessment might be unreliable.

Species representation within the network of protected lands

The extent of the species' predicted habitat within protected lands decreased markedly after incorporating the tree canopy cover constraints (Table 1). For most species, these reductions were equivalent to 20% and 60% of the original area. Few species from groups 1 and 2 showed changes smaller than 5%. However, when evaluating the percentage of the predicted habitat within protected lands, we found practically no differences between the original ID-GAP estimates and the new ones incorporating tree canopy cover data ($t = 1.48$, $p = 0.15$) (Table 3). In this sense, the changes in the estimates of the species representation within the network of protected lands (*i.e.* before and after the NLCD_TCC data) did not surpass 5% (see last column in Table 1).

Patterns of species richness after adding NLCD_TCC data

The species richness values at the pixel scale changed after incorporating tree canopy cover information (Fig. 3). In the map of species richness created with the original predictions for the 37 species (*i.e.* from the ID-GAP), all the forested pixels appeared to support a high, and relatively constant, number of species (between 25 and 30). In the map that incorporated vegetation structure (*i.e.* NLCD_TCC), the number of species per pixel was considerably lower (between 8 and 22 for most of the forested pixels) (Fig. 3). The difference between the two maps revealed that the number of species decreased in practically all the forested pixels after incorporating tree canopy constraints, with the largest reductions occurring in areas corresponding to closed forests (the north and center

part of the state). Less severe reductions occurred in areas dominated by open forests (e.g. the south-central portion of Idaho).

The pixel based values of species richness including all the birds in Idaho (n=238) also changed after incorporating the 37 new models (Fig. 4 top). Forested pixels showed a decrease in the number of bird species, in proportions that ranged mostly between 5% and 35%. In the original GAP map, forests were dominated by species richness values between 55 and 80 in the northern and central region, and by slightly lower values in the south. The new map (i.e. after the NLCD_TCC) exhibited lower values of species richness in forests, mostly ranging between 40 and 60. The changes were higher in areas dominated by closed forests than in areas dominated by moderate density or open forests (see Fig. 4 center). When evaluating the location of species richness hotspots in the forests of Idaho in relationship with the protected lands, the most evident change after adding the tree canopy data was that the richest and largest hotspot shifted towards non protected lands, a pattern that differed from the original GAP outputs.

At the hexagon scale, the species composition remained unchanged after adding the tree canopy cover data. After refining the predicted distributions of the 37 species with the NLCD_TCC there was still some habitat available for all of the original species listed in the hexagons. Although the predicted habitat per species decreased within the hexagon, this change never resulted in an absence of habitat. The changes observed were a function of the species' preferences and of the characteristics of the dominant vegetation. Hexagons in areas dominated by closed forests experienced higher habitat reductions for species associated with non-forests and open forests (groups 1, 2, and 3) and smaller reductions in the habitat for species associated with denser forests (groups 4 and 5), while the opposite was observed in areas of open vegetation (see Fig. 4). Only 6 of the 404 hexagons showed some decrease in the species composition, however these were not the typical 635-km² GAP-hexagons, but rather smaller fractions of those hexagons located along the state border (data not shown).

Discussion and conclusions

Remote sensing data provide vital information for mapping the distribution of wildlife species, which is a common requisite for assessing species conservation and biodiversity patterns. However, broad scale assessments such as the US GAP Analysis have been conducted using species distribution models that do not incorporate information about vegetation structure, an important variables explaining the distribution of many birds and mammals [5,6,17,26]. In this sense, geospatial layers reflecting tree canopy closure has been recognized as a major data need for improving GAP assessments [22,23]. In this study, we evaluated the consequences for broad scale species distribution and conservation assessments brought about by the inclusion of novel remote sensing data about vegetation structure. We integrated the new tree canopy product from the NLCD2001 [24] into GAP species-habitat models, using the state of Idaho as a case study.

The incorporation of the NLCD_TCC into the GAP habitat assessment protocol resulted in: (1) remarkable changes in the predicted distribution of many avian species, (2) changes in the values of avian species richness at a certain scale, (3) a modified distribution of pixel based biodiversity hotspots in forested areas, and (4) surprising results in conservation assessment. We improved the predicted distribution models of 37 avian species with the NLCD_TCC data, allowing us to represent more precise species-habitat relationships, and reducing previous habitat overestimations without incorporating omission errors. As a result, the assessments of species distribution and conservation based on these refined predictions differed from the original ID-GAP ones (Table 3). Some assessments, however, were not sensitive to model refinement. For example, the most representative GAP measure of conservation, that is, the percentage of the species habitat occurring within the network of protected lands, did not change despite remarkable decreases in the predicted habitat after the NLCD_TCC was added (Table 3).

At the scale of this study (216,000 km², with 78,000 km² of forests) the addition of geospatial data of vegetation structure represented the difference between a significant habitat overestimation and a more accurate prediction for certain bird species. Modeling

the distribution of wildlife species that depend on specific conditions of tree canopy closure in the absence of such data resulted in large overestimation errors. Similar consequences have been observed in other studies with different variables of forest structure. For example, [33] noted that giant panda (*Ailuropoda melanoleuca*) distributions were significantly overestimated without the inclusion of information about understory vegetation. Similarly, the potential distributions of the endangered Delmarva fox squirrel (*Sciurus niger cinereus*) decreased considerably when adding constraints about forest canopy height [34]. Reductions in habitat size may have consequences for assessing habitat connectivity [33] or for evaluating the species' conservation status based on available habitat. In this sense, our study showed significant habitat reductions for two species listed with the Idaho Department of Fish and Game as in greatest conservation need: the Peregrine falcon (*Falco peregrinus anatum*) and the Flammulated owl (*Otus flammeolus*). While our analysis did not evaluate the size and configuration of the new predicted habitats, we speculate that because available habitat became more fragmented after incorporating the NLCD_TCC data, our estimates of habitat reductions are conservative and could be further refined using, for example, concepts of minimum patch size and patch isolation.

It was surprising to find that, after the large reductions in the predicted distributions for the 37 avian species that resulted from adding the NLCD_TCC data, the estimates of the percentage of the species' habitats occurring within protected lands remained practically the same. The reason for this resides in the similar proportion of tree canopy classes inside and outside protected lands, observed in the forests of Idaho. While some forest types were denser within protected lands (such as Mixed Subalpine Forest or Mesic Forest), others more open (such as Ponderosa pine or Douglas-fir/Lodgepole Pine) and others were relatively similar (such as Douglas-fir or Aspen), the combination of all the forest types canceled such structural differences; as a result, the proportion of low, medium or high tree canopy cover was the same for the forests located inside or outside protected lands (data not shown). The inclusion of NLCD_TCC data in the species habitat modeling process therefore reduced the predicted habitat in a similar fashion inside and outside the reserves, maintaining the original proportions (i.e. by the ID-GAP).

In absolute terms, however, the extent of the species predicted habitat (for both inside and outside reserves) decreased markedly after including the NLCD_TCC. Further study is warranted to evaluate these issues in other regions (e.g. broadleaf forests).

For the assessments of species richness, the modification of the predicted distributions of the 37 species produced discernable changes in the pixel based map habitat patterns for all the birds in Idaho (n=238). The number of bird species predicted to occur in forested pixels decreased as a result of adding tree canopy information. These changes were not homogeneously distributed, and in general, areas dominated by closed forests (mainly in the Rocky Mountains in the north and central part of the state) were more affected than areas dominated by open forests (the southern part). The reason for this is that the majority of the original 37 species are associated with open forests and not with closed forests, and thus, higher overestimations and changes were observed (see Fig. 4) and expected (see Fig. 2) in closed forests. Areas with open forests, including the southern region of the state, exhibited smaller changes. A previous study in an open forest area of Spain comparable to the southern part of Idaho (i.e. a mix of shrublands, sparse forests, some closed forests, and grasslands) found that incorporating tree canopy cover in predictive distribution models of avian species did not result in significant changes [16]. However, while model refinement decreased the number of bird species predicted at the pixel level, it did not alter the number of species predicted at the hexagon scale. Differences in responses and patterns of species richness are expected because these products represent information with different spatial resolution (hexagon vs. pixel) [35]. Between these two spatial scales, however, there will likely exist a new pixel size at which the consequences of model refinement with vegetation structure data are still imperceptible for species richness analyses, and which is worth to identify in further efforts.

Because the reserve network solution from conservation planning efforts is highly sensitive to the quality of the environmental layers and species distribution maps [32,36], further research should evaluate the impacts that novel remote sensing derived data have on the outputs of reserve design analyses. In our study, for example, the new map of bird species richness at the pixel scale revealed the presence of a large biodiversity hotspot

located outside of the protected lands (see Fig. 5), a pattern not evident in the original ID-GAP data due to the habitat overestimation problem that occurs in much of the forests. The importance of this hotspot for conservation actions may increase when considering that Idaho is one of the fastest growing states in terms of human population and land development.

The implications of this study, including the potential use of NLCD_TCC data in wildlife habitat assessment, reach beyond bird species. Certain mammals are also known to occur under specific tree canopy cover conditions; for instance the ID-GAP identified nine species whose habitats have been likely overestimated due the lack of tree canopy cover data [22]. The list includes species of major economic importance such as elk (*Cervus elaphus*) and mule deer (*Odocoileus hemionus*), and the fisher (*Martes pennanti*), which is a candidate to be listed under the Endangered Species Act. Considering the relevance of the information about the distribution of these species for supporting conservation and management decisions, continue evaluating the incorporation of tree canopy data emerges as an important task for Idaho and beyond.

There is great potential for the immediate application of the NLCD_TCC data for large scale biodiversity mapping and conservation assessments. For instance, the US GAP is developing a second generation of species distribution models for the North Western states, including Oregon, Washington, Idaho, Montana, Wyoming, and California (J. Aycrigg, personal communication). In addition, agencies such as the US Fish and Game are focusing efforts to improve the predicted habitat maps for their species of interest, such as elk and mule deer. Finally, the nationwide development of the State Wildlife Action Plans (SWAP), mandated by the US Congress, might provide an additional framework for species-specific applications. Global assessments looking for tree canopy data, on the other hand, can potentially benefit from the global-1km product developed by the Global Land Cover Facility [37].

The date of the NLCD_TCC data (2001) can be a limitation for analyses seeking to reflect the current (i.e. year 2009) landscape. However, for species that depend on some specific condition of tree canopy cover, the incorporation of the NLCD_TCC may

be more relevant than a new land cover that does not reflect vegetation horizontal structure, even if the structural layer is few years old. An additional limitation of the use of these data for wildlife habitat assessment is spatial extent. The product does not provide information about the percent tree canopy cover in areas dominated by grasslands or shrublands, which represent about 35% of the United States [24]. These areas can contain some tree cover ($< 20\%$), which represent important features for certain wildlife species. The lack of this type of information when predicting species distribution in rangelands has been identified as a potential cause of omission errors [22].

The avian modeling refinements made possible by incorporating tree canopy cover data in this study highlight the utility of wide area vegetation structure data products for improved species distribution and conservation assessments. Although our study focused on the horizontal component of forest structure, many species select habitat based on 3-dimensional forest canopy structure [5,6]. Information about understory conditions, location of old growth forests, and canopy height have been identified, among others, as important variables for improving the habitat predictions for some wildlife species [22,23,33,34]. While obtaining such information from satellite imagery has been difficult, the relatively new airborne lidar (light detection and ranging, or laser altimeter) data may be a potential answer to that problem [38]. The availability of such datasets over regions, nations, or continents will open myriad novel avenues for advancing biodiversity and conservation assessments.

Acknowledgments

This work was made possible through funding by the US GAP Analysis Program and the USDA Forest Service International Institute of Tropical Forestry (IITF). We thank L. Svancara at the ID-GAP for facilitating the species lists for the accuracy assessment, and A. Haines and H. Jageman for providing information about species-habitat associations for some species. Work at IITF is done in collaboration with the University of Puerto Rico.

References

- [1] H. Nagendra, "Using remote sensing to assess biodiversity," *Int. J. Remote Sens.* 22, 2377-2400 (2001) [doi: 10.1080/01431160117096].
- [2] W. Turner, S. Spector, N. Gardiner, M. Fladeland, E. Sterling, M. Steininger, "Remote sensing for biodiversity science and conservation," *Trends Ecol. Evol.* 18, 306-313 (2003) [doi:10.1016/S0169-5347(03)00070-3].
- [3] W. Gillespie, G. M. Foody, D. Rocchini, A. P. Giorgio, S. Saatchi, "Measuring and modelling biodiversity from space," *Prog. Phys. Geogr.* 32, 203-221 (2008) [doi:10.1177/0309133308093606].
- [4] N. Coops, M. A. Wulder, A. M. Pidgeon, V. C. Radeloff, "Bird diversity: a predictable function of satellite-derived estimates of seasonal variation in canopy light absorbance across the United States," *J. Biogeogr.* (in press) [doi:10.1111/j.1365-2699.2008.02053.x].
- [5] R. H. MacArthur, and J. W. MacArthur, "On bird species diversity," *Ecology* 42, 594-598 (1961) [doi:10.2307/1932254].
- [6] M. L. Cody, Ed., *Habitat selection in Birds*, Academic Press, Inc., San Diego (1985).
- [7] J. M. Scott, P. Heglund, M. Morrison, J. Huafler, M. Raphael, W. Wall, F. Samson, Eds., *Predicting species occurrences: issues of accuracy and scale*, Island Press, Washington DC (2002).
- [8] T. K. Gottschalk, F. Huettmann, M. Ehlers, "Thirty years of analyzing and modeling avian habitat relationships using satellite imagery data: a review," *Int. J. Remote Sens.* 26, 2631-2656 (2005) [doi:10.1080/01431160512331338041].
- [9] G. J. McDermid, S. E. Franklin, E. F. LeDrew, "Remote sensing for large-area habitat mapping," *Prog. Phys. Geogr.* 29, 449-474 (2005) [doi:10.1191/0309133305pp455ra].
- [10] E. Leyequien, C. Verrelst, M. Slot, G. Schaepman-Strub, I. M. A. Heitkönig, A. Skidmore, "Capturing the fugitive: Applying remote sensing to terrestrial animal distribution and diversity," *Int. J. Appl. Earth Obs. Geoinf.* 9, 1-20 (2007) [doi:10.1016/j.jag.2006.08.002].

- [11] J. M. Scott, F. Davis, B. Csuti, R. Noss, B. Butterfield, C. Groves, H. Anderson, S. Caicco, F. D'Erchia, T. C. Edwards Jr., J. Ulliman, R. G. Wright, "GAP Analysis – A geographic approach to protection of biological diversity," *Wildlife Monographs* 123, 1-41 (1993).
- [12] B. R. Noon, D. D. Murphy, S. R. Beissinger, M. L. Shaffer, D. Dellasala, "Conservation planning for the US national Forests: conducting comprehensive biodiversity assessments," *Bioscience* 53, 1217-1220 (2003) [doi:10.1641/0006-3568(2003)053[1217:CPFUNF]2.0.CO;2].
- [13] J. Maxwell, "Role of GAP data in State Wildlife Plan Development: Opportunities and Lessons Learned," in *GAP Analysis Bulletin No 14*, J. Maxwell, Ed., pp. 4-11, USGS/BRD/GAP Analysis Program, Moscow ID, USA (2006).
- [14] P. M. Fearnside, J. Ferraz, J., "A conservation GAP analysis of Brazil Amazon vegetation," *Conserv. Biol.* 9, 1134-1147 (1995) [doi:10.1046/j.1523-1739.1995.9051134.x].
- [15] A. S. L. Rodrigues, S. J. Andelman, M. I. Bakar, L. Boitani, T. M. Brooks, R. M. Cowling, L. D. C. Fish-pool, G. A. B. da Fonseca, K. J. Gaston, M. Hoffman, J. S. Long, P. A. Marquet, J. D. Pilgrim, P. L. Pressey, J. Schipper, W. Sechrest, S. N. Stuart, L. G. Underhill, R. W. Waller, M. E. J. Watts, X. Yan, "Global gap analysis: Towards a representative network of protected areas," in *Advances in Applied Biodiversity Science*, No. 5, Center for Applied Biodiversity Science at Conservation International, Washington DC, USA (2003).
- [16] J. Seoane, J. Bustamante, R. Díaz-Delgado, "Are existing vegetation maps adequate to predict bird distributions?," *Ecol. Modell.* 175, 137-149 (2004) [doi:10.1016/j.ecolmodel.2003.10.011].
- [17] M. F. Wilson, "Avian community organization and habitat structure," *Ecology* 55, 1017-1029 (1974) [doi:10.2307/1940352].
- [18] F. C. James, N. O. Wamer, "Relationships between temperate forest bird communities and vegetation structure," *Ecology* 63, 159-171 (1982) [doi:10.2307/1937041].

- [19] A. H. Fielding, J. F. Bell, "A review of methods for the assessment of prediction errors in conservation presence/absence models," *Environ. Conserv.* 24, 38-49 (1997) [doi:10.1017/S0376892997000088].
- [20] T. S. Beutel, R. J. S. Beeton, G. S. Baxter, "Building better wildlife-habitat models," *Ecography* 22, 219-223 (1999) [doi:10.1111/j.1600-0587.1999.tb00471.x].
- [21] A. Guisan, N. E. Zimmerman, "Predictive habitat distribution models in ecology," *Ecol. Modell.* 135, 147-186 (2000) [doi:10.1016/S0304-3800(00)00354-9].
- [22] J. M. Scott, C. R. Peterson, J. W. Karl, E. Strand, L. K. Svancara, N. M. Wright, "A GAP Analysis of Idaho," Final Report, Idaho Cooperative Fish and Wildlife Research Unit, Moscow, ID, USA (2002).
- [23] S. Martinuzzi, L. Vierling, W. Gould, K. Vierling, "Improving the characterization and mapping of wildlife habitats with LiDAR data: measurement priorities for the Inland Northwest, USA," in GAP Analysis Bulletin No 16, J. Maxwell, Ed., USGS/BRD/GAP Analysis Program, Moscow ID, USA (in press).
- [24] C. Homer, J. Dewitz, J. Fry, M. Coan, N. Hossain, C. Larson, N. Herold, A. McKerrow, J. N. VanDriel, J. Wickham, "Completion of the 2001 National Land Cover database for the Conterminous United States," *Photogramm. Eng. Remote Sens.* 73, 337-341 (2007).
- [25] M. D. Jennings, M.D., "GAP analysis: concepts, methods, and recent results," *Landscape Ecol.* 15, 5-20 (2000) [doi:10.1023/A:1008184408300].
- [26] A. J. Kroll, J. B. Haufler, "Development and evaluation of habitat suitability models at multiple spatial scales: A case study with the dusky flycatcher," *Forest Ecol. Manag.* 229, 161-169 (2006) [doi:10.1016/j.foreco.2006.03.026].
- [27] R. Sallabanks, J. B. Haufler, C. A. Mehl, "Influence of forest vegetation structure on avian community composition in west-central Idaho," *Wildlife Soc. B.* 34, 1079-1093 (2006) [doi:10.2193/0091-7648(2006)34[1079:IOFVSO]2.0.CO;2].
- [28] W. Krohn, W., "Predicted vertebrate distributions from GAP analysis: Considerations in the design of statewide accuracy assessments," in GAP Analysis: A landscape approach to biodiversity planning, J. M. Scott, T. H. Tear, F. W. Davis,

- Eds., pp. 147-162, American Society of Photogrammetry and Remote Sensing, Bethesda, Md., USA (1996).
- [29] J. W. Karl, N. M. Wright, P. J. Heglund, E. O. Garton, J. M. Scott, R. L. Hutto, "Sensitivity of species habitat relationships models performance to factors of scale," *Ecol. Appl.* 10, 1690-1705 (2000) [doi:10.1890/1051-0761(2000)010[1690:SOSHRM]2.0.CO;2].
- [30] S. M. Schaefer, W. B. Krohn, "Predicting vertebrate occurrences from species habitat associations: Improving the interpretation of commission error rates," In *Predicting species occurrences: issues of accuracy and scale*, J. M. Scott, P. Heglund, M. Morrison, J. Huafler, M. Raphael, W. Wall, F. Samson, Eds., pp. 419-427, Island Press, Washington DC, USA (2002).
- [31] T. C. Edwards, E. T. Deshler, D. Foster, G. G. Moisen, "Adequacy of wildlife habitat Relation Models for estimating Spatial Distributions of Terrestrial vertebrates," *Conserv. Biol.* 10, 263-270 (1996) [doi:10.1046/j.1523-1739.1996.10010263.x].
- [32] C. Rondinini, K. A. Wilson, L. Boitani, H. Grantham, H. P. Possingham, "Tradeoffs of different types of species occurrence data for use in systematic conservation planning," *Ecol. Lett.* 9, 1136-1145 (2006) [doi:10.1111/j.1461-0248.2006.00970.x].
- [33] M. Linderman, S. Beier, L. An, Y. C. Tan, Z. Y. Ouyang, H. G. Liu, "The effects of understory bamboo on broad-scale estimated of giant panda habitat", *Biol. Conserv.* 121, 383-390 (2005) [doi:10.1016/j.biocon.2004.05.011].
- [34] R. Nelson, C. Keller, M. Ratnaswamy, "Locating and estimating the extent of the Delmarva fox squirrel using airborne Lidar profiler," *Remote Sens. Environ.* 96, 292-301 (2005) [doi:10.1016/j.rse.2005.02.012].
- [35] A. H. Hurlbert, E. P. White, "Disparity between range map- and survey-based analysis of species richness: pattern, process and implications," *Ecol. Lett.* 8, 319-327 (2005) [doi:10.1111/j.1461-0248.2005.00726.x].
- [36] K. A. Wilson, M. I. Westphal, H. P. Possingham, J. Elith, "Sensitivity of conservation planning to different approaches to using predicted species distribution data," *Biol. Conserv.* 122, 99-112 (2005) [doi:10.1016/j.biocon.2004.07.004].

- [37] R. DeFries, M. Hansen, J. R. G. Townshend, A. C. Janetos, T. R. Loveland, "A new global 1km data set of percent tree cover derived from remote sensing," *Glob. Change Biol.* 6, 247-254 (2000) [doi:10.1046/j.1365-2486.2000.00296.x].
- [38] K. T. Vierling, L. A. Vierling, W. A. Gould, S. Martinuzzi, R. M. Clawges, "Lidar: shedding new light on habitat characterization and modeling," *Front. Ecol. Environ.* 2, 90-98 (2008) [doi:10.1890/070001].

Table 1. Predicted habitats before and after adding the NLCD_TCC. The table includes estimates of total habitat area and area predicted within protected lands (i.e. reserves). Area unit corresponds to thousands of km². The species are presented by groups (denoted by the letter “G”) similar to Fig. 2. The scientific names of the species is provided in Appendix A

	Predicted habitat area			Predicted habitat area within protected lands			% of habitat within protected lands		
	Before	After	% Diff.	Before	After	% Diff.	Before	After	Diff.
Species common name									
Lark sparrow	100	99	-0.6	6	6	-0.9	5.6	5.6	0.0
Loggerhead shrike	109	108	-0.6	6	6	-0.5	5.1	5.1	0.0
Common nighthawk	120	119	-1.0	8	8	-1.7	6.4	6.4	0.0
Common poorwill	86	83	-3.2	6	5	-1.6	6.4	6.5	-0.1
Golden eagle	206	152	-26.3	25	14	-45.8	12.2	8.9	3.2
G1 Common raven	212	154	-27.3	25	14	-45.7	11.9	8.9	3.0
Brown-headed cowbird	208	150	-27.9	24	13	-47.6	11.7	8.5	3.2
Long-eared owl	143	96	-32.5	21	10	-49.7	14.5	10.8	3.7
Lazuli bunting	158	100	-36.6	22	11	-52.4	14.0	10.5	3.5
Cedar waxwing	123	66	-46.3	18	6	-64.5	14.5	9.6	4.9
Broad-tailed hummingbird	62	27	-56.8	9	3	-60.9	13.8	12.5	1.3
Western tanager	84	27	-67.8	16	5	-69.9	19.6	18.3	1.2
Blue-gray gnatcatcher	31	31	-0.1	3	3	0.0	8.3	8.3	0.0
Brewer's blackbird	131	130	-0.9	9	9	-1.0	6.8	6.8	0.0
Black-capped chickadee	16	15	-8.3	1	1	-5.2	8.8	9.1	-0.3
Peregrine falcon	138	116	-15.8	19	14	-23.6	13.5	12.2	1.2
Red-tailed hawk	213	177	-16.8	25	18	-28.3	11.9	10.3	1.7
Northern flicker	209	173	-17.1	24	17	-29.4	11.6	9.9	1.7
Turkey vulture	188	155	-17.5	20	14	-31.0	10.7	9.0	1.8
G2 Blue grouse	114	79	-31.1	19	12	-37.9	16.6	14.9	1.6
Dusky flycatcher	96	61	-36.9	17	10	-41.8	17.8	16.4	1.4
Chipping sparrow	96	60	-37.1	18	10	-40.7	18.4	17.3	1.0
Oregon (Dark-eyed) junco	95	60	-37.1	18	10	-40.8	18.4	17.3	1.1
Black-headed grosbeak	75	46	-38.4	14	8	-42.7	18.3	17.0	1.3
Fox sparrow	90	54	-39.4	17	10	-42.4	18.7	17.8	0.9
Cassin's finch	88	54	-39.5	17	10	-41.2	19.6	19.1	0.5
Northern saw-whet owl	78	43	-44.4	16	9	-44.2	20.7	20.8	0.0
Great gray owl	76	42	-44.8	16	9	-44.3	21.0	21.2	-0.2
Flammulated owl	37	22	-40.9	6	3	-45.1	16.9	15.7	1.2
G3 Cassin's vireo	81	46	-43.5	16	8	-45.7	19.2	18.4	0.8
Clark's nutcracker	73	39	-46.7	15	8	-46.3	21.0	21.1	-0.2
Chestnut-backed chickadee	47	39	-15.2	10	8	-15.8	21.0	20.8	0.1
G4 Cordilleran flycatcher	74	58	-22.4	15	12	-22.5	20.6	20.6	0.0
Northern goshawk	80	61	-23.7	16	12	-22.6	19.4	19.7	-0.3
Red-breasted nuthatch	80	61	-23.9	16	12	-22.6	19.6	19.9	-0.3
G5 Pileated woodpecker	70	36	-48.6	15	8	-50.2	21.9	21.2	0.7
Northern pygmy-owl	78	38	-50.8	15	8	-50.7	19.8	19.9	-0.1

Table 2. Accuracy assessment of the initial (i.e. ID-GAP) and refined (i.e. ID-GAP + NLCD_TCC) predicted species distributions, including the number of sites used for evaluation, the % of correct predictions (CP%), and the % of omissions (OM%).

Species common name	Reference sites (#)	ID-GAP		ID-GAP + NLCD_TCC	
		CP%	OM%	CP%	OM%
Black-capped chickadee	25	100	0	100	0
Black-headed grosbeak	18	94	6	94	6
Blue grouse	18	100	0	100	0
Blue-gray gnatcatcher	1	0	100	0	100
Brewer's blackbird	32	97	3	97	3
Broad-tailed hummingbird	8	75	25	75	25
Brown-headed cowbird	31	100	0	100	0
Cassin's finch	14	100	0	100	0
Cassin's vireo	14	93	7	93	7
Cedar waxwing	17	100	0	100	0
Chestnut-backed chickadee	5	100	0	100	0
Chipping sparrow	26	92	8	92	8
Clark's nutcracker	15	93	7	93	7
Common nighthawk	25	96	4	96	4
Common poorwill	9	100	0	100	0
Common raven	29	100	0	100	0
Cordilleran flycatcher	5	100	0	100	0
Dusky flycatcher	13	100	0	100	0
Flammulated owl	4	100	0	100	0
Fox sparrow	10	100	0	100	0
Golden eagle	25	100	0	100	0
Great gray owl	5	100	0	100	0
Lark sparrow	11	100	0	100	0
Lazuli bunting	19	100	0	100	0
Loggerhead shrike	10	100	0	100	0
Long-eared owl	14	100	0	100	0
Northern flicker	40	100	0	100	0
Northern goshawk	15	100	0	100	0
Northern pygmy-owl	11	100	0	100	0
Northern saw-whet owl	14	100	0	100	0
Oregon (Dark-eyed) junco	21	100	0	100	0
Peregrine falcon	2	100	0	100	0
Pileated woodpecker	13	100	0	100	0
Red-breasted nuthatch	16	100	0	100	0
Red-tailed hawk	38	100	0	100	0
Turkey vulture	21	100	0	100	0
Western tanager	20	100	0	100	0

Table 3. Summary table of the consequences for GAP assessments brought about by the inclusion of broad scale remote sensing data about vegetation structure.

GAP estimates/assessments	Outcome after adding the NLCD TCC
Predicted Species Distribution Maps	
Area	Decreased
Omission error	No change
Commission error	Decreased
Overall accuracy	Increased
Species Richness	
Pixel-based richness	Decreased
Hexagon-based richness	No change
Species Conservation Assessment	
Predicted habitat within the network of protected lands (in km ²)	Decreased
Predicted habitat within the network of protected lands (in %)	No change

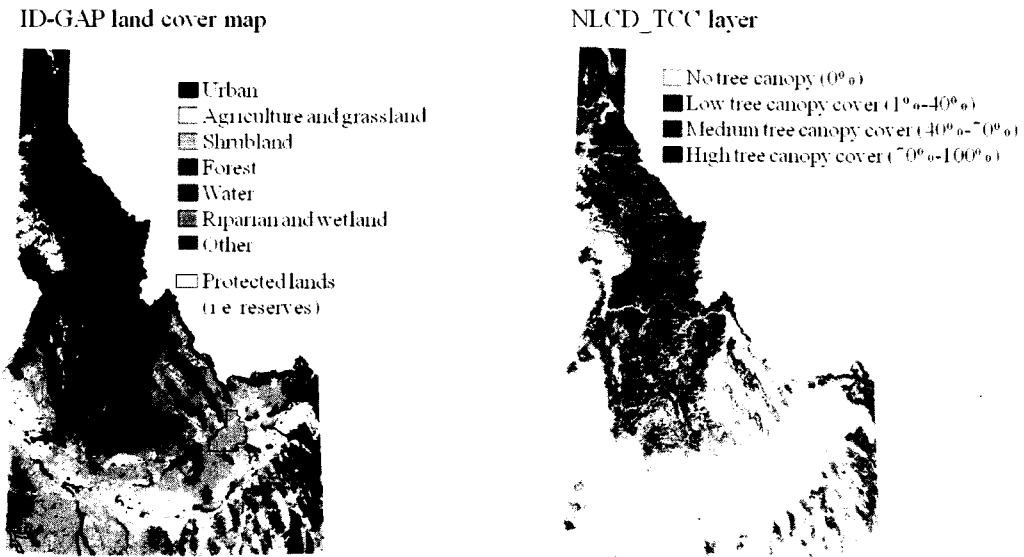


Figure 1. Simplified land cover map and tree canopy for Idaho.

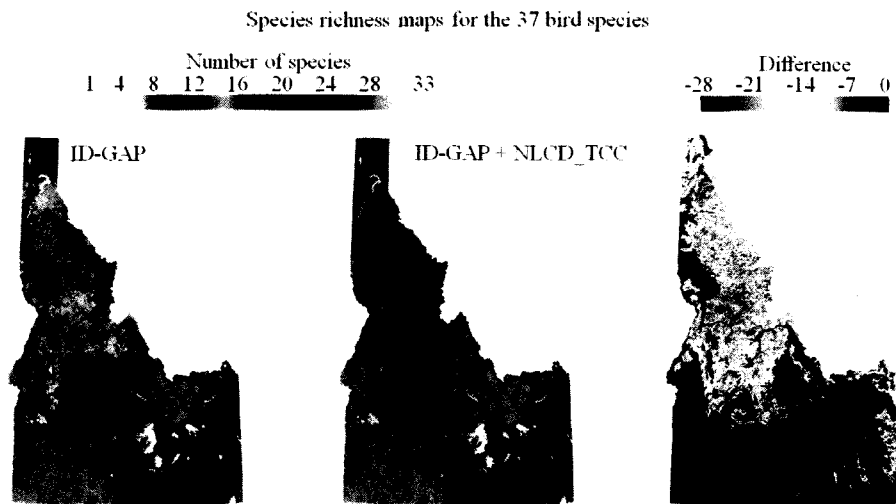


Figure 3. Pixel based, species richness maps for the 37 avian species, including from the original ID-GAP predicted habitats and from the new ones incorporating the NLCD_TCC. The comparison of species richness between these products is also presented.

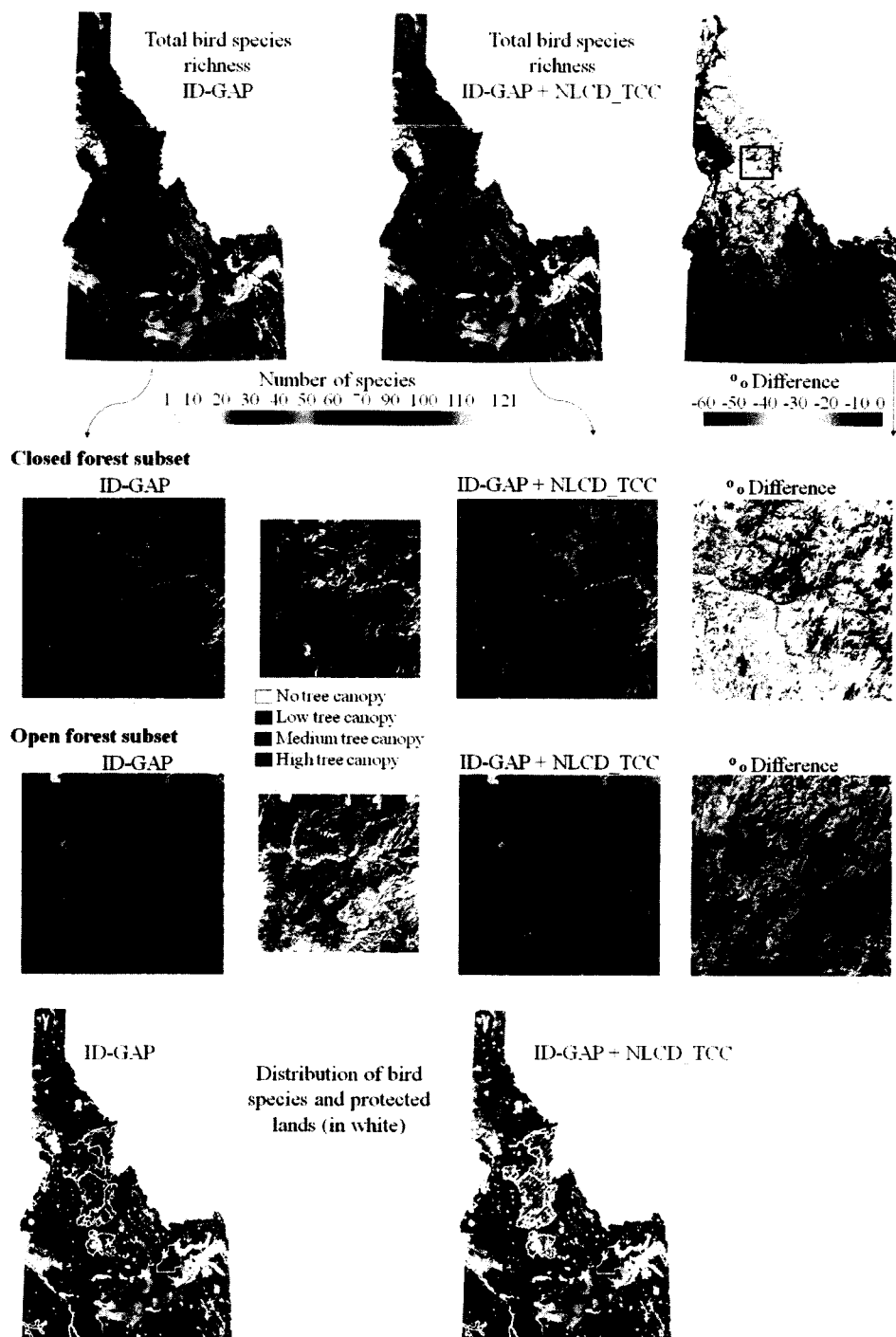


Figure 4. Pixel based, species richness layers for all the birds in Idaho ($n=238$), before and after incorporating the NLCD_TCC. Subsets of areas dominated by closed forest and mid-open forest are shown in the center. Patterns of species richness and distribution of protected lands are displayed in the bottom.

Appendix A

Species' scientific names in order of appearance in Table. G1: Lark sparrow (*Chondestes grammacus*); Loggerhead shrike (*Lanius ludovicianus*); Common nighthawk (*Chordeiles minor*); Common poorwill (*Phalaenoptilus nuttallii*); Golden eagle (*Aquila chrysaetos*); Common raven (*Corvus corax*); Brown-headed cowbird (*Molothrus ater*); Long-eared owl (*Asio otus*); Lazuli bunting (*Passerina amoena*); Cedar waxwing (*Bombycilla cedrorum*); Broad-tailed hummingbird (*Selasphorus platycercus*); Western tanager (*Piranga ludoviciana*). G2: Blue-gray gnatcatcher (*Polioptila caerulea*); Brewer's blackbird (*Euphagus cyanocephalus*); Black-capped chickadee (*Poecile atricapilla*); Peregrine falcon (*Falco peregrinus anatum*); Red-tailed hawk (*Buteo jamaicensis*); Northern flicker (*Colaptes auratus*); Turkey vulture (*Cathartes aura*); Blue grouse (*Dendragapus obscurus*); Dusky flycatcher (*Empidonax oberholseri*); Chipping sparrow (*Spizella passerina*); Oregon (Dark-eyed) junco (*Junco hyemalis*); Black-headed grosbeak (*Pheucticus melanocephalus*); Fox sparrow (*Passerella iliaca*); Cassin's finch (*Carpodacus cassinii*); Northern saw-whet owl (*Aegolius acadicus*); Great gray owl (*Strix nebulosa*). G3: Flammulated owl (*Otus flammeolus*); Cassin's vireo (*Vireo cassinii*); Clark's nutcracker (*Nucifraga columbiana*). G4: Chestnut-backed chickadee (*Poecile rufescens*); Cordilleran flycatcher (*Empidonax occidentalis*); Northern goshawk (*Accipiter gentiles*); Red-breasted nuthatch (*Sitta Canadensis*). G5: Pileated woodpecker (*Dryocopus pileatus*); Northern pygmy-owl (*Glaucidium gnoma*).

CHAPTER 4

CLASSIFYING FOREST TYPE AND SUCCESSION IN A TROPICAL DRY FOREST: A LiDAR PERSPECTIVE.

Abstract

Improved remote sensing technologies are needed to advance the research and management of tropical dry forests, one of the world's most threatened ecosystems. This study evaluated the use of airborne LiDAR data to (1) classify forest types and (2) identify successional stages in a mature tropical dry forest of Puerto Rico. We used classification tree techniques and compared the results with classifications made from commonly available remote sensing data, including Landsat satellite imagery and non-LiDAR topographic data from the Shuttle Radar Topography Mission (SRTM). We found that the accuracy of the LiDAR-based forest type classification was substantially higher than those from previously available remote sensing data (*i.e.* kappa value 0.90 and 0.63 respectively), with the best result obtained when combining LiDAR-derived data of canopy structure and topography. Surprisingly, adding Landsat spectral data did not improve the LiDAR-based classification. For the second objective, we observed that LiDAR-derived variables of vegetation structure were better predictors of forest successional stage (including the identification of primary and different secondary forests) than spectral information from Landsat. Finally, we found that the most important LiDAR-based predictors of forest types (including native- and exotic-dominated classes) and succession tend to follow previous ecological knowledge of canopy structural variations in these forests. Our study highlights the value of LiDAR data for supporting forest land cover mapping and land use legacy assessments in tropical dry forests, and warrants applications in other areas in order to further advance our knowledge of the biophysical and human dimensions of these important ecosystems.

Introduction

Remote sensing is an important tool for supporting research and management of tropical dry forests (Sánchez-Azofeifa et al., 2003), one of the world's most threatened ecosystems (Janzen, 1988). However, current uncertainties in remotely sensed assessments of forest types and succession hamper efforts to quantify important ecosystem characteristics such as those related to biodiversity and carbon storage (*e.g.* Sánchez-Azofeifa et al., 2005, 2009; Quesada et al., 2009). Improved technologies are therefore needed to advance our knowledge of the biophysical and human dimensions of these ecosystems (Sánchez-Azofeifa et al. 2005). In particular, hyperspectral and LiDAR remote sensing hold a special interest because they can add a great deal of dimensionality to the spectral and structural information available to make ecological inferences (Kalacska and Sánchez-Azofeifa, 2008; Sánchez-Azofeifa et al., 2009). The process of applying these technologies in tropical dry forests is just beginning within the scientific community. In this study we evaluated the utility of LiDAR relative to other commonly used remote sensing datasets to (1) classify forest types and (2) identify successional status in a tropical dry forest of Puerto Rico.

LiDAR (*i.e.* light detection and ranging) uses a laser to directly measure canopy height, subcanopy topography, and the vertical distribution of intercepted surfaces, opening new opportunities for assessing ecosystem structure and function with remote sensing data (Lefsky et al., 2002). LiDAR datasets are becoming increasingly common and they are part of the next generation of satellite and airborne-based systems for ecological observations (*e.g.* Carnegie Airborne Observatory [Asner et al., 2007], National Ecological Observatory Network [NEON; Keller et al., 2008], NASA DESDynI mission [<http://desdyni.jpl.nasa.gov>]).

Because different forest types support different groups of plants and animals, maps of forest types derived from remote sensing are commonly used to assess biodiversity (Turner et al., 2003). While the combination of Landsat satellite imagery with environmental GIS layers is the most common approach for classifying vegetation

types with remote sensing (Cohen and Goward, 2004), accurately delineating tropical dry forest types has been problematic. Frequent confusion has been reported, for example, between semi-deciduous forests and shrublands/woodlands, semi-deciduous forests and evergreen forests, and scrub forest, shrublands and pastures, among others (*e.g.* Helmer et al., 2002, 2008; Kennaway and Helmer, 2007; Gould et al., 2008; Martinuzzi et al., 2008). Further classification challenges may arise by the presence of forest types in different successional stages, as well as by invasive species creating novel forest types (Lugo, 2009).

Recent studies in temperate zones found that LiDAR-derived information of canopy height can be powerful predictor of vegetation types, either alone or in combination with spectral imagery (Antonarakis et al., 2009; Geerling et al., 2009). In particular, LiDAR data can improve the delineation of forest classes that are spectrally similar but structurally different (Neuenschwander et al., 2009), a problem that confounds accurate classification in tropical dry forests. In addition, because topography influences the distribution of tropical forest types at a local scale (Lugo, 2005), the use of topographic information derived from the LiDAR derived digital elevation model (DEM) may prove useful for further refining forest classifications.

Understanding succession and land use legacy is important for managing tropical forests, as they affect the structure, function, and services of these ecosystems (Chazdon et al., 2009; Quesada et al., 2009). Today, tropical dry forests occur in a dynamic mosaic of different human land uses and forest patches in different stages of development (Quesada et al., 2009). In this context, the delineation of successional stages (*i.e.* young, mature, and primary) is a topic of major research interest, but of major technological challenges (Sánchez-Azofeifa et al., 2003, 2009; Quesada et al., 2009). Previous studies were able to separate successional stages in forests up to 30 yr of a age (*i.e.* in young to mid-secondary forests) using the Landsat-derived Normalized Difference Vegetation Index (NDVI) (Arroyo-Mora et al., 2005; Hartter et al., 2008). In these studies, the value of the NDVI resides in its sensitivity to tree canopy cover (Feeley et al., 2005), and the fact that tree canopy cover increases during the first 20-30 years (Arroyo-Mora et al.,

2005; Hartter et al., 2008). After that, however, the overall horizontal structure of the forest (including canopy cover) tends to stabilize (Lugo, 2005), limiting the use of NDVI to further track succession. As a result, assessing succession in tropical dry forests of > 30 yr old has been difficult (Sánchez-Azofeifa et al., 2003, 2009). Because the vertical structure of tropical dry forests can further change after the horizontal structure stabilizes (Lugo, 2005), LiDAR data may be helpful for tracking successional changes in more mature stands. This includes, for example, the identification of mature and old growth forests, which are of major value for biodiversity and conservation (Chazdon et al., 2009).

Scientists in other regions are using LiDAR to characterize succession. Pioneering work by Drake et al. (2002) showed that LiDAR was sensitive to successional changes in a tropical moist forest. In a temperate broadleaf forest, Weber and Boss (2009) successfully separated young, intermediate, and mature forests. In a temperate conifer forest, Falkowski et al. (2009) mapped six successional stages, ranging from stand initiation to old growth, with an accuracy of > 95%.

We investigated two major objectives in this study. First, we evaluated the use of airborne LiDAR data for classifying tropical dry forest types in an area exhibiting varied composition, including different native forest formations and a class dominated by an exotic species. Second, we explored the value of LiDAR data for identifying successional status in forests of > 30 years old. An overarching goal in both objectives was to compare the performance of LiDAR with similar classifications made from available Landsat ETM+ imagery and non-LiDAR topographic data.

Methods

Study area

We conducted this study in the Guánica Dry Forest Biosphere Reserve in southwestern Puerto Rico (17°58' N, 66°52' W; Fig. 1). Guánica is considered to be one of the best

examples of subtropical dry forests because it has been protected since the 1930s. The area is a local biodiversity hotspot (Gould et al., 2008), and is a core site for the upcoming NEON program (Keller et al., 2008). The reserve comprises two areas covering nearly 4,000 ha from the shoreline of the Caribbean Sea to an elevation about 230 m, and our study area corresponds to the largest of these portions (approx 3,000 ha; see Fig. 1). Annual rainfall is 860 mm with a major period of drought from December to March and a minor one between June and August (Medina and Cuevas, 1990). The substrate is derived from limestone with common presence of exposed rock outcrops.

Vegetation and land use history

Guánica forests are structurally smaller than tropical/subtropical dry forests from other parts of the world, although they are similar to other Caribbean islands (Murphy and Lugo, 1986). Three intergrading and edaphically determined forest types occupy most of Guánica, including: semi-deciduous forest (the most common vegetation type) in the internal hills and coastal areas, semi-evergreen forest in valleys and ravines with well developed soils, and scrub forest in coastal areas and limestone outcrops. A narrow fringe of dwarf forest occurs along coastal rocky areas exposed to the ocean winds. Finally, a distinctive stand dominated by the exotic *Prosopis pallida* (i.e. mesquite) occurs along an alluvial fan in the southern part of the reserve. Forested wetlands (i.e. mangroves) are not considered in this study. A detailed description of the forest types is provided in Table 1.

Guánica is a mosaic of mature secondary forests and primary forests. Most secondary forests are a result of land use abandonment that occurred between the 1930s and the 1950s. Before that, the area was used mostly for charcoal and fence-post production (involving logging and harvesting of stems and branches from trees), with smaller areas used for agriculture and forest plantations. Studies in the semi-deciduous forests revealed that Guánica has recovered significantly over the last decades (Murphy et al., 1995; Lugo, 2005). After 45 years of land use abandonment, for example, Molina Colon and Lugo (2006) found that the structure of stands previously used for charcoal production or agriculture recovered in >75% (i.e. tree height, basal area, stem density,

and crown area index), with slightly higher recovery values observed in the first group (i.e. charcoal production). According to Agosto Diaz (2008), the height of mature trees in Guánica may be reached after 60 to 70 years of land use abandonment. In addition, past agricultural lands are today dominated by *Leucaena leucocephala* (exotic), while the others are dominated by native tree species. Despite decades of recovery, past land use still explains most of the variation in found in the structure and composition of semi-deciduous forests (Agosto Diaz, 2008).

Data

The data consisted of location samples for the different forest types, information about land use history, airborne LiDAR data, Landsat ETM+ satellite imagery, and topographic information from the SRTM DEM.

For the first objective, we obtained location samples for the different forest types using previous studies and field visits. Most samples for the semi-deciduous forest class were based on field plots established by Agosto Diaz (2008). These plots were designed using a stratified random sample technique with the objective of identifying the human and environmental factors explaining variations in forest structure, and therefore they represented a useful training data set for our study. We visited the area during 2009 to collect additional vegetation location samples. We used GPS surveys and visual interpretation of 1-m spatial resolution color aerial photos supported by expert knowledge, to total 83 sample locations. We ensured that all samples represented an area of at least 30m x 30m of the same forest type (to coincide with the geospatial grain size used by this study), and were separated by > 60 meters (measured from the pixel border; consistent with Agosto Diaz, 2008). Finally, Agosto Diaz (2008) and E. Medina (personal communication) contributed with sample locations for the scrub forest and dwarf forest, respectively.

To support our second objective, we utilized historic land use information for 25 plots of semi-deciduous forest, as reported by Agosto Diaz (2008). The author

reconstructed the land cover and land use history of these plots using aerial photos from the 1930s, 1950s, and 1970s. Based on the inferred year of land use abandonment and the type of past land use, we identified and named four different successional classes. These included: (1) mid-secondary forests with a logging past (*i.e.* ~40 yr old stands previously logged and harvested for charcoal and fence post production), (2) late-secondary forests with a logging past (*i.e.* ~60 yr old stands previously logged and harvested for charcoal and fence posts), (3) late-secondary forests with an agricultural past (*i.e.* ~60 yr old stands previously cleared for agriculture), and (4) primary forests (*i.e.* undisturbed, ≥ 90 yr old stands). Although small in sample size ($n=3$), the past agricultural class allowed us explore the potential separability of different historic land uses (*i.e.* agriculture or charcoal production) within the same successional stage (*i.e.* late-secondary forest).

The airborne LiDAR data were collected during January and February 2004 by 3001 Inc. and the US government, covering Puerto Rico and the US Virgin Islands. The sensor recorded first, last and intermediate return height values, with varied post spacing. In the case of Guánica, the Western half was acquired at 2.8 m post spacing (resulting in 0.25 points/m²) and the Eastern half at 6.0 m post-spacing (0.06/m²). These are very low data densities relative to more recent acquisitions. The reported mean vertical error was 9.27 cm.

The satellite imagery consisted of 30-meter pixel Landsat ETM+ scenes from October 2002 (*i.e.* wet season) and January 2003 (*i.e.* dry season), with precision and terrain correction (*i.e.* Level 1T). The reported horizontal error was 2.3 m for the 2002 scene and 2.6 m for the 2003 scene.

Auxiliary topographic information came from the Shuttle Radar Topography Mission (SRTM), which provides digital elevation data for great part of the globe. We used the 30-meter gridded DEM product. The mean horizontal and vertical errors for the SRTM DEM are less than 12 m and 10 meters respectively, although they may vary locally (Rodríguez et al., 2006).

Data preprocessing

The LiDAR preprocessing involved several steps, following general methods found in other studies (*e.g.* Hudak et al., 2008; Falkowski et al., 2009). We first thinned the Western portion of the data (0.25 points /m²) to the same point density of the Eastern portion (0.06 points /m²). We then used an algorithm (Evans and Hudak, 2007) to classify the LiDAR data into ground vs. canopy returns, and calculated a suite of canopy metrics at a 30-meter grid size (see Table 2). We considered vegetation returns to those located > 30 cm above the ground to avoid confusion with limestone outcrops. Because most vegetation returns (*i.e.* 95%) were first or single returns, the metrics of canopy density by this study (calculated as # of vegetation returns above a specific height * 100 / total number of returns, within each 30m x 30m grid cell) can be used also as a physical measure of canopy cover (*i.e.* horizontal structure).

We calibrated the Landsat data to reflectance values and then applied a radiometric normalization method (*i.e.* Canty and Nielsen, 2008) using ENVI 4.5 (ITT Visual Information Solutions ®). We calculated the NDVI and the Tasseled Cap transformation for the two scenes. The Tasseled Cap (Kauth and Thomas, 1976) reduces the image data to three ecologically meaningful bands or indices of brightness, greenness and wetness, which have shown useful for separating successional stages and forest types in other areas (*e.g.* Helmer et al., 2000; Song et al., 2007). Finally, we calculated differences in the band spectral values and indices between the two image dates (see Table 2).

We horizontally co-registered the 30m SRTM DEM to the 30m LiDAR DEM and calculated a suite of topographic variables from both DEMs using ArcGIS 9.2 (ESRI®) Spatial Analyst (see Table 2). Finally, we created a GIS layer quantifying the distance from the coast.

Data analysis

We used classification tree techniques for both objectives. Classification trees partition the data hierarchically using binary splits based on discrete values of the predictor variables, such that each split successively reduces the misclassification error in the partitioned group. For the first objective we used the Random Forest (RF) algorithm (Breiman, 2001), implemented in R (R development core 2005, Liaw and Wiener, 2002), which is a bootstrap-based extension of classification tree methods that has shown excellent results in classifications of ecological and remote sensing data (*e.g.* Lawrence et al., 2006; Cutler et al., 2007; Martinuzzi et al., 2009). In RF, classification rules are developed by combining 100s to > 1,000s of classification trees, which are constructed from random subsets of the training data and explanatory variables. The algorithm provides a robust internal measure of misclassification error (using observations that are randomly withheld in each tree development), which has shown to eliminate the need for a secondary dataset for accuracy assessment (Breiman, 2001; Lawrence et al., 2006; Falkowski et al. 2009). The algorithm provides measures of variable importance (derived from the permutation of the independent variables), which can be used to compare with ecological expectations based on published literature (Cutler et al., 2007). We identified the most parsimonious model using a RF-based method by Murphy et al. (2010), which iteratively reduces the number of variables using the variable's importance measure from RF, identifying the model with the lowest total error and maximum within-class error. Collinear variables were initially removed.

We classified forest types in two ways, first using single sources of remotely-sensed explanatory variables (*e.g.* LiDAR canopy alone, Landsat wet season alone, etc.) and then by combining multiple data sources. The variables Elevation (ELEV) and Distance to the coast (DIST) were added at the end, in a separate new mode. This is because the importance of these variables appears to be too specific of Guánica (*i.e.* they explain much of the distribution of the mesquite and dwarf forest) and we wanted to test the ability of LiDAR to classify forest types in spite of their relative location on the landscape (*i.e.* coastal or inland forest). Therefore, creating classification models with

and without the variables ELEV and DIST allowed us to better transfer our findings to other areas.

Classification trees were applied differently in classifying past land use. Because the sample size was smaller ($n=25$, restricted to Semi-deciduous forests), we were able to partition the data visually. We developed decision trees using the smallest number of binary partitions possible to explain the data (three splits; *i.e.* # of land use history classes – 1). We compared the results of using LiDAR canopy metrics vs. Landsat multi-season imagery. Topography was not included in objective 2 because it is not an important variable explaining variations in forest structure within Semi-deciduous forests (Agosto Diaz, 2008). For all the classifications made, we reported global and class-level errors rates (from 0 to 1) and the kappa statistic. The kappa statistic (Landis and Koch 1977) measures overall classification accuracy compensating for agreement due to chance. Kappa value of 1.0 denotes perfect agreement and 0.0 no agreement other than that which would be expected by chance, with ranges of 0.41- 0.60, 0.61-0.80, and 0.81- 0.99 describing moderate, substantial, and almost perfect accuracies.

Results

We first classified forest types using one source of remotely-sensed explanatory variables. The classification using LiDAR canopy metrics alone yielded the highest accuracy (kappa = 0.69), followed by the LiDAR DEM metrics (kappa = 0.58; Table 3). The classifications using Landsat or STRM DEM data, on the other hand, had kappa values ranging between 0.30 and 0.37. In addition, the classification from the LiDAR DEM had a higher accuracy than the classification from SRTM DEM (kappa values of 0.58 and 0.30 respectively). In these models, the semi-evergreen forest showed the largest error difference, with a misclassification rate of 0.05 from the LiDAR DEM vs. 0.57 from the SRTM DEM. Finally, classification accuracies from the two Landsat scenes were very similar (kappa values of 0.37 and 0.32).

The most important predictors in the LiDAR-canopy model included the median absolute deviation of vegetation heights (HMAD), the 90th percentile of vegetation heights (H90th), and the percent of returns > 1.0 m (CDENSITY2, Table 3). The NDVI and the wetness and greenness indices were important in the Landsat-based models. Finally, the variable curvature (CURV) appeared among the most important variables in the classifications using topographic data (*i.e.* LiDAR and SRTM DEMs).

The combination of the two different sources of LiDAR explanatory variables (*i.e.* canopy and topographic metrics) reduced the error from 0.23 (observed using LiDAR canopy metrics alone) to 0.14, resulting in a kappa statistic of 0.81 (Table 3). The class-level errors ranged between 0.13 and 0.16. On the other hand, combining the two Landsat seasons did not substantially improve the classification accuracy over the dry season imagery alone (kappa values of 0.39 and 0.37 respectively). The class-level errors for the Landsat multi season classification ranged between 0.21 and 0.76, with two of the three most extensive forest types in Guánica (*i.e.* semi-deciduous and semi-evergreen) greater than 0.40.

The addition of SRTM variables to the multi-date Landsat model increased the classification accuracy from a kappa of 0.39 observed in the Landsat-only model to 0.51 (Table 3). The class-level error ranged between 0.21 and 0.67, with the largest error found in the semi-evergreen forest. Finally, the addition of Landsat variables did not improve the accuracy observed using LiDAR explanatory variables (canopy plus DEM; *i.e.* kappa 0.82 vs. 0.81). The variables included in the models developed from multiple sources of remotely sensed based explanatory variables were typically a combination of the most important variables observed in the models using one source at a time.

Adding the variables elevation (ELEV) and distance to the coast (DIST) improved the classification accuracy. The kappa value increased from 0.81 to 0.90 in the LiDAR-based model, and from 0.51 to 0.63 in the Landsat-SRTM one (Table 3). The final models included ELEV and DIST in addition to other variables that were important in the

previous classifications. The resultant LiDAR-based map of forest types for Guánica is shown in Figure 2.

In addressing our second objective we classified past land use within the semi-deciduous forest using LiDAR (canopy) vs. Landsat variables (Fig. 3). The model constructed from LiDAR had the highest accuracy (kappa values of 0.88 and 0.77 respectively), and both models separated the mid and late secondary forest classes equally. The LiDAR based model, however, separated the primary forests perfectly (*i.e.* error rate = 0.00) while the Landsat based models produced some confusion (error rate = 0.18). The three explanatory variables included in LiDAR-based model included the percentage of vegetation returns between 0.3m and 3m (STRAT1), the 75th percentile of vegetation heights (H75th), and the median absolute deviation of vegetation heights (HMAD; Fig 3). On the other hand, the Landsat based model included the Band 4, Band 5 and the Brightness index, all from the dry season.

Discussion and conclusions

Because improved remote sensing technologies are needed to advance the research and management of tropical dry forests, we explored the use of LiDAR data to classify forest types and quantify forest succession, which are challenging aspects of remote sensing in these ecosystems. Furthermore, our study focused on mature secondary and primary forests and involved native- and exotic-dominated forest types, which are research-need areas in the face of tropical forest conservation and global change (Chazdon et al., 2009; Lugo, 2009).

We found that LiDAR data were powerful predictors of tropical dry forest types, achieving higher accuracies than classifications made from previously available remote sensing data (*i.e.* Landsat ETM+ and SRTM). The best results were obtained when combining LiDAR data of canopy structure and topography. As a result, forest types that have been difficult to separate in previous efforts using Landsat, such as semi-deciduous

forests and shrublands/scrub forest or semi-deciduous forests and semi-evergreen forests (e.g. Helmer et al., 2002, 2008; Kennaway and Helmer, 2007; Gould et al., 2008; Martinuzzi et al., 2008), were possible to separate using LiDAR data.

In addition, we detected no enhanced value for classifying forest types in Guánica after integrating Landsat with LiDAR data. In fact, with a misclassification rate of 0.14 (and 0.07 when adding the variables elevation and distance to the coast; Table 3), the results from LiDAR yielded little room for improvement. Expanding to other land cover types is however important to better understand the role of LiDAR for mapping and characterizing tropical dry forest landscapes.

The forests of Guánica vary in terms of canopy height, canopy cover and vertical complexity, and are profoundly affected by soil conditions (Lugo 2005). The most important LiDAR-derived predictors seem to reflect those variations, providing a novel and meaningful picture of the vegetation (Fig 4). The metric H90th, for example, tends to correspond well with the mean tree height of a stand (Hopkinson et al., 2006; Li, 2008). We likewise observed an increase in canopy height as going from the dwarf forest to the scrub forest, semi-deciduous (and mesquite), and finally to the semi-evergreen forest, with median values of about 1m, 3m, 5m, and 10m that match the general descriptions found in the literature (see Lugo, 2005; Agosto Diaz, 2008). The metric CDENSITY2, a surrogate of canopy cover in this study, showed that well-developed forests such as the semi-deciduous, semi-evergreen and mesquite have high canopy cover (e.g. > 70% for the semi-deciduous and consistent with Murphy and Lugo [1986]), while the naturally open scrub forest reported much lower values, as expected (Fig. 4). The fact that the dwarf forest appeared with little or no canopy cover is an artifact of the threshold established at > 1 m, above which little or no vegetation exists for this forest type. Finally, the metric HMAD (median absolute deviation of vegetation heights) seems to reflect known patterns of vertical heterogeneity. According to Lugo (2005), the semi-evergreen is the most structurally diverse forest type in Guánica because it supports the largest variations in tree height and diameter. In our study, these forests showed the highest HMAD values, agreeing with Lugo (2005).

Distinctive characteristics of the invasive Mesquite forest also related to the HMAD LiDAR metric. Our study showed that the mesquite forest is similar to the semi-deciduous forest in terms of horizontal cover (*i.e.* CDENSITY2) and height (H90th; Wilcoxon W Test [W] = 87.5, p-value = 0.68 for CDENSITY2 and W = 107.5, p-value = 0.71 for H90th). However, this exotic-dominated forest appeared less vertically diverse (in HMAD values) than the semi-deciduous forest (Wilcoxon W Test [W] = 167, p-value = 0.005; see Fig. 4). This might be a result of the different number of tree strata naturally present in these two forest types, *i.e.* one for the mesquite (Stromberg, 1993) vs. one to three for the semi-deciduous (Lugo, 2005). This difference in strata, in turn, will cause the laser returns to be more scattered along the vertical dimension, resulting in different HMAD values. These findings agreed with Asner et al. (2008a,b), who found that exotic species can transform the 3D structure of the forest, and that LiDAR can be sensitive to those transformations. In this sense, there are great opportunities to continue investigating the role of this powerful technology in the context of biological invasions and novel forests. For example, we observed little variation in H90th between the different plots of Mesquite (Fig. 4), suggesting the presence of a tree cohort established likely after a single disturbance event (*e.g.* agricultural abandonment), rather than gradual recruitment over time.

Topographic (*i.e.* DEM-based) variables were important for refining classification of tropical dry forest types, agreeing with previous remote sensing studies (*e.g.* Helmer et al., 2008; Martinuzzi et al., 2008; Sesnie et al., 2008). However, although both LiDAR and SRTM DEMs added predictive power to the classifications, the LiDAR DEM did a better job in discriminating vegetation types. This difference was largely attributed to the semi-evergreen forest, which was better separated from the rest with the LiDAR DEM. We believe that this is a result of a lower ability of SRTM data to depict topographic changes under forest canopies compared to LiDAR. While LiDAR returns can reach the true ground even under very dense forest canopies (Hofton et al., 2002), elevations retrieved from SRTM are located somewhere between the canopy top surface and the ground (Hofton et al., 2006). As a result, valleys and ravines, which are the distinctive

locations of semi-evergreen forests, were well defined in the LiDAR DEM, but were more difficult to distinguish from the surrounding terrain in the SRTM DEM (see Fig. 5). In these areas, the lower surface elevations might be canceled by the simultaneous presence of taller forests, resulting in a final SRTM elevation that is not very different from the nearest cells. The different values of curvature (i.e. CURV, which describe convexity and concavity of slope profiles with positive and negative values respectively) observed in the LiDAR and SRTM DEMs for the semi-evergreen forests might further confirm the previous point (Fig. 4).

LiDAR metrics of canopy structure were useful for separating mid and late secondary forests, as well as stands of primary forests. Identifying these habitats is critical for conservation and management, yet has been difficult in previous studies (Sánchez-Azofeifa et al., 2009). We also observed that the LiDAR variables that best classified succession were those describing vertical canopy structure (rather than horizontal, *i.e.* STRAT1, H75th, and HMAD). This supports the idea that horizontal structure might have stabilized in these mature forests, while changes in the vertical dimension might still be occurring (Lugo, 2005). Although more research is needed to better understand the biophysical meaning of these metrics, our findings reinforced the value of LiDAR to assess land use legacy (Falkowski et al., 2009). Finally, the fact that an exotic species (*Leucaena leucocephala*) dominates forest recovery in one of the past land uses perfectly classified (*i.e.* the one with agricultural history), adds further value to use of LiDAR for assessing novel forests. Expanding and testing succession assessments in other tropical dry forest types and land use histories are warranted.

Our findings provide also valuable information for further applications using optical imagery (*e.g.* Landsat). For instance, our study supported the use of dry season imagery to assess succession in tropical dry forests, similar to Arroyo-Mora et al. (2005) and Kalacska et al. (2007). However, we found that the variables that appeared useful for classifying succession in mature forests by this study (*i.e.* Landsat bands 4, 5 and brightness, achieving a kappa of 0.77) were different to those identified by previous authors in younger forests, *i.e.* NDVI (see Arroyo-Mora et al., 2005 and Hartter et al.,

2008). This might be related to the fact that NDVI captures differences in horizontal structure between successional stages, which are more evident in young forests than in mature ones (Kalacska et al. 2005; Lugo 2005). The successional classes by this study showed very similar NDVI and canopy cover values (CDENSITY2), which might reflect this point (Fig. 6). Finally, the variables identified by this study were consistent with Helmer et al. (2002) assessing succession in other tropical forests.

Our study highlights the value of LiDAR remote sensing data -even low density LiDAR- for supporting forest land cover mapping and land use legacy assessments in tropical dry forests, and complements recent findings by Kalacska et al. (2007) using hyperspectral data. Continued application of these promising technologies (see Asner and Martin, 2009) is important to advance the research and management of these valuable ecosystems.

Acknowledgements

Funding for this work was provided by the USGS Gap Analysis Program and administrated by the US Forest Service IITF and the University of Idaho (U of I). E. Næsset, M. Falkowski, and J. Stoker provided valuable information for the LiDAR processing. R. Agosto, E. Medina, J. Chinea, and S. Van Bloem facilitated vegetation information. Technical support was provided by the University of Idaho's Geospatial Laboratory for Environmental Dynamics, the US Forest Service IITF GIS and Remote Sensing Lab, and the US Forest Service Rocky Mountains Research Station, Moscow Forest Science Laboratory. Work at IITF is conducted in collaboration with the University of Puerto Rico.

References

- Agosto Diaz, R. E. (2008). Human and environmental factors explaining the structural and compositional variability in a sub-tropical dry forest. MS Thesis, University of Puerto Rico, 55 pp.
- Antonarakis, A. S., Richards, K. S., & Brasington, J. (2009). Object-based land cover classification using airborne LiDAR. *Remote Sensing of Environment*, 112, 2988-2998.
- Arroyo-Mora, J. P., Sánchez-Azofeifa, G. A., Kalacska, M. E. R., Rivard, B., Calvo-Alvarado, J. C., & Janzen, D. H. (2005). Secondary forest detection in a Neotropical dry forest landscape using Landsat 7 ETM+ and Ikonos imagery. *Biotropica*, 34, 497-507.
- Asner, G. P., Jones, M. O., Knapp, D. E., Kennedy-Bowdoin, T., Jones, M. O., Martin, R. E., Boardman, J. W., & Field, C. B. (2007). Carnegie Airborne Observatory: in-flight fusion of hyperspectral imaging and waveform light detection and ranging for three-dimensional studies of ecosystems. *Journal of Applied Remote Sensing*, 1, [Online: doi 10.1117/1111.2794018].
- Asner, G. P., Hughes, R. F., Vitousek, P. M., Knapp, D. E., Kennedy-Bowdoin, T., Boardman, J., Martin, R. E., Eastwood, M., & Green, R. O. (2008a). Invasive plants transform the three-dimensional structure of rain forests. *Proceedings of the National Academy of Sciences*, 18, 4519-4523.
- Asner, G. P., Knapp, D. E., Kennedy-Bowdoin, T., Jones, M. O., Martin, R. E., Boardman, J. & Hughes, R. F. (2008b). Invasive species detection in Hawaiian rainforests using airborne imaging spectroscopy and LiDAR. *Remote Sensing of Environment*, 112, 1942, 1955.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5-32.
- Canty, M. J., & Nielsen, A. A. (2008). Automatic radiometric normalization of multitemporal satellite imagery with the iteratively re-weighted MAD transformation. *Remote Sensing of Environment*, 112, 1025-1036.

- Chazdon, R. L., Peres, C. A., Dent, D., Sheil, D., Lugo, A. E., Lamb, D., Stork, N. E., Niguel, E., & Miller, S. E. (2009). The potential for species conservation in tropical secondary forests. *Conservation Biology*, 23, 1406-1417.
- Cowen, W. B., & Goward, S. N. (2004). Landsat's role in ecological applications of remote sensing. *Bioscience*, 54, 535-545.
- Cutler, R. D., Edwards, T. C. Jr., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random Forests for classification in ecology. *Ecology*, 88, 2783-2792.
- Drake, J. B., Dubayah, R. O., Clark, D. B., Knox, R. G., Blair, J. B., Hofton, M. A., Chazdon, R. L., Weishampel, J. F., & Prince, S. D. (2002). Estimation of tropical forest structural characteristics using large-footprint lidar. *Remote Sensing of Environment*, 79, 305-319.
- Evans, J. S. & Hudak, A. T. (2007). A multiscale curvature algorithm for classifying discrete return LiDAR in forested environments. *IEEE Transactions on Geoscience and Remote Sensing*, 45, 1029-1038.
- Evans, J. S., & Cushman, S. A. (2009). Gradient modeling of conifer species using random forests. *Landscape Ecology*, 24, 673-683.
- Falkowski, M. J., Evans, J. S., Martinuzzi, S., Gessler, P. E., & Hudak, A. T. (2009). Characterizing forest succession with Lidar data: an evaluation for the Inland Northwest USA. *Remote Sensing of Environment*, 113, 946-956.
- Farr, T. & Kobrick, M. (2000). Shuttle radar topography mission produces a wealth of data. *Eos Trans. AGU* 81, 583-585.
- Feeley, K. J., Gillespie, T. W., & Terbough, J. W. (2005). The utility of spectral indices from Landsat ETM+ for measuring the structure and composition of tropical dry forests. *Biotropica*, 37, 508-519.
- Geerling, G. W., Vreeken-Buijs, M. J., Jesse, P., Ragas, A. M. J., & Smits, A. J. M. (2009). Mapping river floodplain ecotypes by segmentation of spectral (CASI) and structural (Lidar) remote sensing data. *River Research and Applications*, 25, 795-813.
- Gesch, D., Oimoen, M., Greenlee, S., Nelson, C., Steuck, M., & Tyler, D. (2002). The National Elevation Dataset. *Photogrammetric Engineering and Remote Sensing*, 68, XXX.

- Gessler, P. E., Moore, I. D., McKenzie, N. J., & Ryan, P. J. (1995). Soil-landscape modeling and spatial prediction of soil attributes. *International Journal of Geographic Information Systems*, 9, 421-432.
- Gould, W. A., Alarcón, C., Fevold, B., Jiménez, M. E., Martinuzzi, S., Potts, G., Quiñones, M., Solórzano, M., & Ventosa, E. (2008). The Puerto Rico Gap Analysis Project – Volume 1: Land cover, vertebrate species distributions and land stewardship. US Forest Service IITF-GTR-39, 165 pp.
- Guariguata, M. R., & Ostertag, R. (2001). Neotropical secondary forest succession: changes in structural and functional characteristics. *Forest Ecology and Management*, 148, 185-206.
- Harterter, J., Lucas, C., Gaughan, A. E., & Aranda, L. L. (2008). Detecting tropical dry forest succession in a shifting cultivation mosaic of the Yucatán Peninsula, Mexico. *Applied Geography*, 28, 134-149.
- Helmer, E. G., Brown, S., & Cohen, W. B. (2000). Mapping montane tropical successional stage and land use with multi-date Landsat imagery. *International Journal of Remote Sensing*, 21, 2163-2183.
- Helmer, E. H., Ramos, O. R., López, T. M., Quiñones, M., & Diaz, W. (2002). Mapping forest type and land cover of Puerto Rico, a component of the Caribbean biodiversity hotspot. *Caribbean Journal of Science*, 38, 165–183.
- Helmer, E. H., Kennaway, T. A., Pedreros, D. H., Clark, M. L., Marcano-Vega, H., Tieszen, L. L., Ruzycski, T. R., Schill, S. R., & Carrington, C. M. S. (2008). Land cover and forest formation distributions for St. Kitts, Nevis, St. Eustatius, Grenada and Barbados from decision tree classification of cloud-cleared satellite imagery. *Caribbean Journal of Science*, 44, 175-198.
- Hofton, M. A., Rocchio, L. E., Blair, J. B., & Dubayah, R. (2002). Validation of large-footprint lidar sub-canopy topography. *Journal of Geodynamics*, 34, 477–489.
- Hofton, M., Dubayah, R., Blair, J. B., & Rabine, D. (2006). Validation of SRTM elevations over vegetated and non-vegetated terrain using medium footprint lidar. *Photogrammetric Engineering and Remote Sensing*, 72, 279–285.
- Hopkinson, C., Chasmer, L., Lim, K., Treitz, P., & Creed, I. (2006). Towards a universal lidar canopy height indicator. *Canadian Journal of Remote Sensing*, 32, 139–152.

- Hudak, A. T., Crookston, N. L., Evans, J. S., Hall, D. E., & Falkowski, M. J. (2008). Nearest neighbor imputation modeling of species-level, plot-scale structural attributes from LiDAR data. *Remote Sensing of Environment*, 112, 2232-2245.
- Hudak, A. T., Evans, J. S., & Smith, A. M. S. (2009). LiDAR utility for natural resource managers. *Remote Sensing*, 1, 934-951.
- Iverson, L. R., Dale, M. E., Scott, C. T., & Prasad, A. (1997). A GIS-derived integrated moisture index to predict forest composition and productivity of Ohio forests (U.S.A.). *Landscape Ecology*, 12, 331-348.
- Janzen, D. H. (1988). Tropical dry forests: the most endangered major ecosystem. In: Wilson, E. O. (Ed.). *Biodiversity*. Washington, DC: National Academic Press. pp. 130-137.
- Kalacska, M. E. R., Sánchez-Azofeifa, G. A., Calvo-Alvarado, J. C., Rivard, B., and Quesada, M. (2005). Effects of season and successional stage on leaf area index and spectral vegetation indices in three Mesoamerican tropical dry forests. *Biotropica*, 37, 486-496.
- Kalacska, M., Sánchez-Azofeifa, G. A., Rivard, B, Caelli, T., White, H. P., & Calvo-Alvarado, J. C. (2007). Ecological fingerprint of ecosystem succession: Estimating secondary tropical dry forest structure and diversity using imaging spectroscopy. *Remote Sensing of Environment*, 108, 82-96.
- Kalacska, M., & Sánchez-Azofeifa, A., Eds. (2008). *Hyperspectral remote sensing of tropical and subtropical forests*. Taylor and Francis Group, LCC, Boca Raton, FL.
- Kauth, R. J., & Thomas, G. S. (1976). The tasseled cap – a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. *Proceedings on the Symposium on Machine Processing of Remotely Sensed Data*, 4b: 41-51, June 6 – July 2, West Lafayette, Indiana.
- Keller, M., Schimel, D. S., Hargrove, W. W., & Hoffman, F. M. (2008). A continental strategy for the National Ecological Observatory Network. *Frontiers in Ecology and the Environment*, 6, 282-284.
- Kennaway, T., & Helmer, E. H. (2007). The forest types and ages cleared from land development in Puerto Rico. *GIScience & Remote Sensing*, 44, 356-382.

- Kennaway, T. A., Helmer, E. H., Lefsky, M. A., Brandeis, T. A., & Sherrill, K. R. (2008). Mapping land cover and estimating forest structure using satellite imagery and coarse resolution lidar in the Virgin Islands. *Journal of Applied Remote Sensing*, 2, 023551.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* 33, 159-174.
- Lawrence, R. L., Wood, S. D., & Sheley, R. L. (2006). Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (RandomForest). *Remote Sensing of Environment*, 100, 356-362.
- Lefsky, M. A., Cohen, W. B., Parker, G. G., & Harding, D. J. (2002). Lidar remote sensing for ecosystem studies. *Bioscience*, 52, 19-30.
- Li, Y. (2009). A comparison of forest height prediction from FIA field measurement and LiDAR data via spatial models. In: McWilliams, Will; Moisen, Gretchen; Czaplowski, Ray, comps. 2009. 2008 Forest Inventory and Analysis (FIA) Symposium; October 21-23, 2008: Park City, UT. Proc. RMRS-P-56CD. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 1 CD.
- Liaw, A., & Wiener, M. (2002). Classification and regression by Random-Forest. *R News*, 2, 18-22.
- Lugo, A. E. (2005). Los Bosques. In: R. Joglar (Ed), *Biodiversidad de Puerto Rico*. Instituto de Cultura Puertorriqueña, San Juan, PR, pages 395-548.
- Lugo, A. E. (2009). The emerging era of novel tropical forests. *Biotropica*, 41, 589-591.
- Martinuzzi, S., Gould, W. A., Ramos Gonzalez, O. M., Martinez Robles, A., Calle Maldonado, P., Pérez-Buitrago, N., Fumero Caban J. J. (2008). Mapping tropical dry forest habitats integrating Landsat NDVI, Ikonos imagery, and topographic information in the Caribbean Island of Mona. *Revista de Biología Tropical*, 56, 625-639.
- Martinuzzi, S., Vierling, L., Gould, W., Falkowski, M., Evans, L., Hudak, A., & Vierling, K. (2009). Mapping snags and understory shrubs for a LiDAR-based assessment of wildlife habitat suitability. *Remote Sensing of Environment*, 113, 2533-2546.

- Molina Colón, S., & Lugo, A. E. (2006). Recovery of a subtropical dry forest after abandonment of different land uses. *Biotropica*, 38, 354-364.
- Murphy, M. A., Evans, J. S., & Storfer, A. S. (2009). Quantifying *Bufo boreas* connectivity in Yellowstone National Park with landscape genetics. *Ecology* (in press).
- Neuenschwander, A. L., Magruder, L. A., & Tyler, M. (2009). Landcover classification of small-footprint, full-waveform lidar data *Journal of Applied Remote Sensing*, 3, [Online: DOI:10.1117/1.3229944]
- Quesada, M., Sánchez-Azofeifa, G. A., Alvarez-Añorve, M., Stoner, K. E., Avila-Cabadilla, L., Calvo-Alvarado, J., et al. (2009). Succession and management of tropical dry forests in the Americas: Review and new perspectives. *Forest Ecology and Management*, 258, 1014-1024.
- R Development Core Team. (2005). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Rodríguez, E., Morris, C. S., & Belz, E. (2006). A global assessment of the SRTM performance. *Photogrammetric Engineering & Remote Sensing*, 72, 249-260.
- Sánchez-Azofeifa, G. A., Castro, K. L., Rivard, B., & Kalacska, M. R. (2003). Remote sensing research priorities in tropical dry forest environments. *Biotropica*, 35, 134-142.
- Sánchez-Azofeifa, G. A., Quesada, M., Rodriguez, J. P., Nassar, J. M., Stoner, K. E., Castillo, A., Garvin, T., Zent, E. L., Calvo, J., Kalacska, M. R., Fajardo, L., Garmon, J., & Cuevas-Reyes, P. (2005). Research priorities for Neotropical dry forests. *Biotropica*, 37, 477-485.
- Sánchez-Azofeifa, G. A., Castro-Esau, K. L., Kurz, W. A., & Joyce, A. (2009). Monitoring carbon stocks in the tropics and the remote sensing operational limitations: from local to regional projects. *Ecological Applications*, 19, 480-494.
- Sesnie, S. E., Hagell, S. E., Otterstrom, S. M., Chambers, C. L., & Dickson, B. G. (2008). SRTM-DEM and Landsat ETM+ data for mapping tropical dry forest cover and biodiversity assessment in Nicaragua. *Revista Geográfica Acadêmica*, v.2, n.2, 53-65.

- Silver, W. L., Ostertag, W., & Lugo, A. E. (2000). Restoration of abandoned tropical agricultural and pasture lands. *Restoration Ecology*, 8, 394-407.
- Song, C., Schroeder, T. A., & Cohen, W. B. (2007). Predicting temperate conifer forest successional stage distributions with multitemporal Landsat Thematic Mapper imagery. *Remote Sensing of Environment*, 106, 2008-237.
- Stromberg, J. C. (1993). Riparian mesquite forests: a review of their ecology, threats, and recovery potential. *Journal of the Arizona-Nevada Academy of Science*, 27, 111-124.
- Turner, W., Spector, S., Gardiner, N, Fladeland, M., Sterling, E. & Steininger, M. (2003). Remote sensing for biodiversity science and conservation. *Trends in Ecology and Evolution*, 18, 306-314.
- Vierling, K. T., Vierling, L. A., Gould, W., Martinuzzi, S., & Clawges, R. (2008). Lidar: Shedding new light on habitat characterization and modeling. *Frontiers in Ecology and the Environment*, 6, 90-98.
- Weishampel, J. F., Drake, J. B., Cooper, A., Blair, J. B., & Hofton, M. (2007). Forest canopy recovery from the 1983 hurricane and subsequent salvage damage measured with airborne LiDAR. *Remote Sensing of Environment*, 109, 142-153.
- Weber, T. C., & Boss, D. E. (2009). Use of LiDAR and supplemental data to estimate forest maturity in Charles County, MD, USA. *Forest Ecology and Management*, 258, 2068-2075.

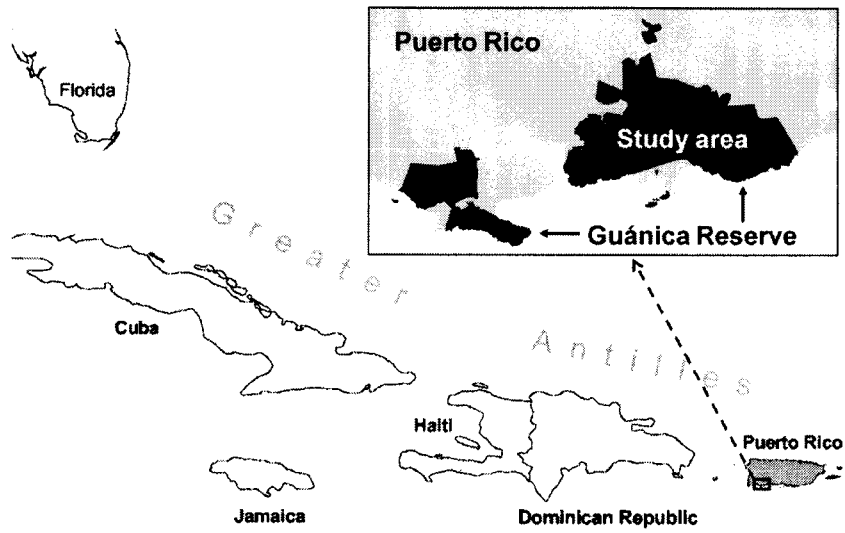


Figure 1. Locations of the study area.

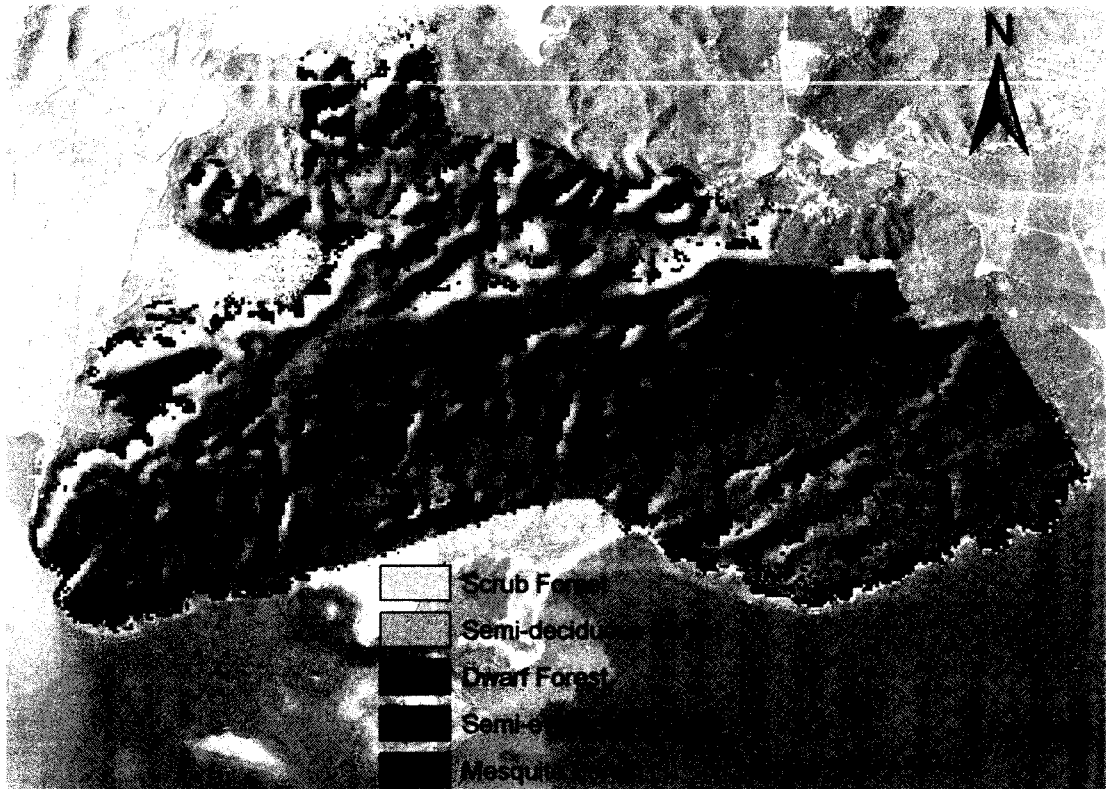


Figure 2. Map of forest types for Guánica derived from LiDAR data.

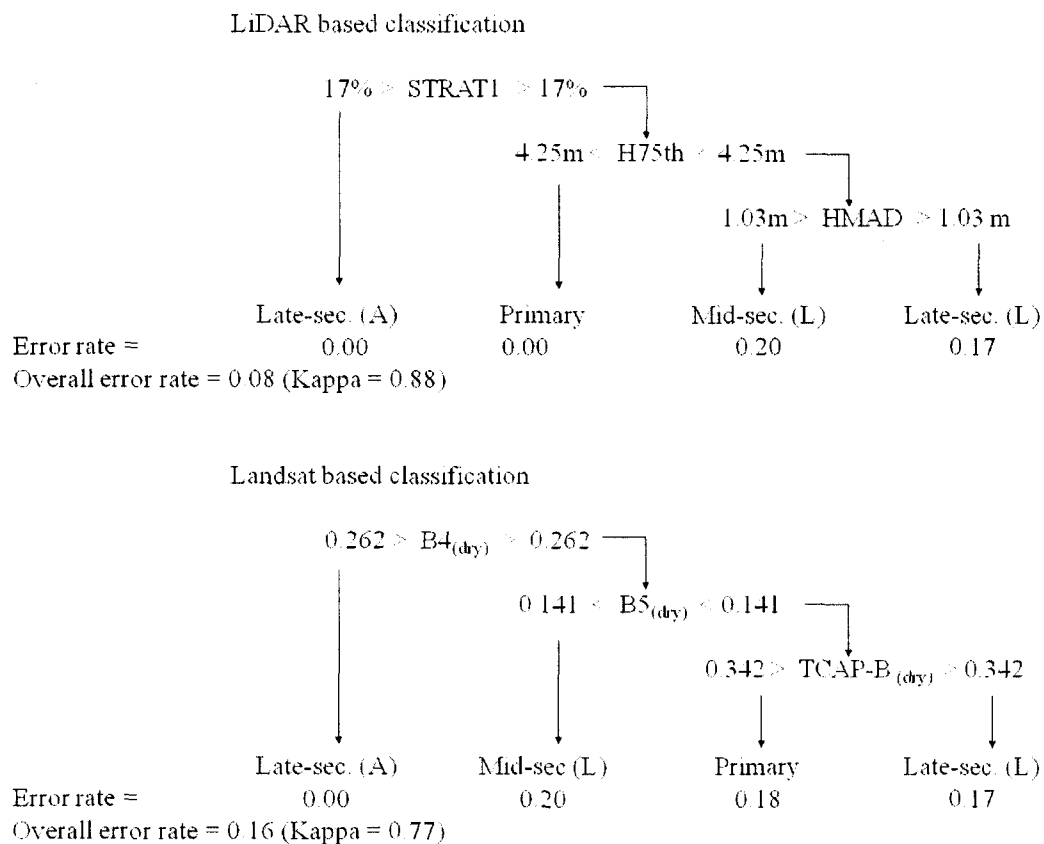


Figure 3. Classification trees decisions and resultant accuracy statistics for different successional stages using LiDAR (top) or Landsat ETM+ (bottom) variables. Mid-sec. = Mid-secondary; Late-sec. = Late-secondary; (A) and (L) describe agricultural or logging-based past land uses respectively.

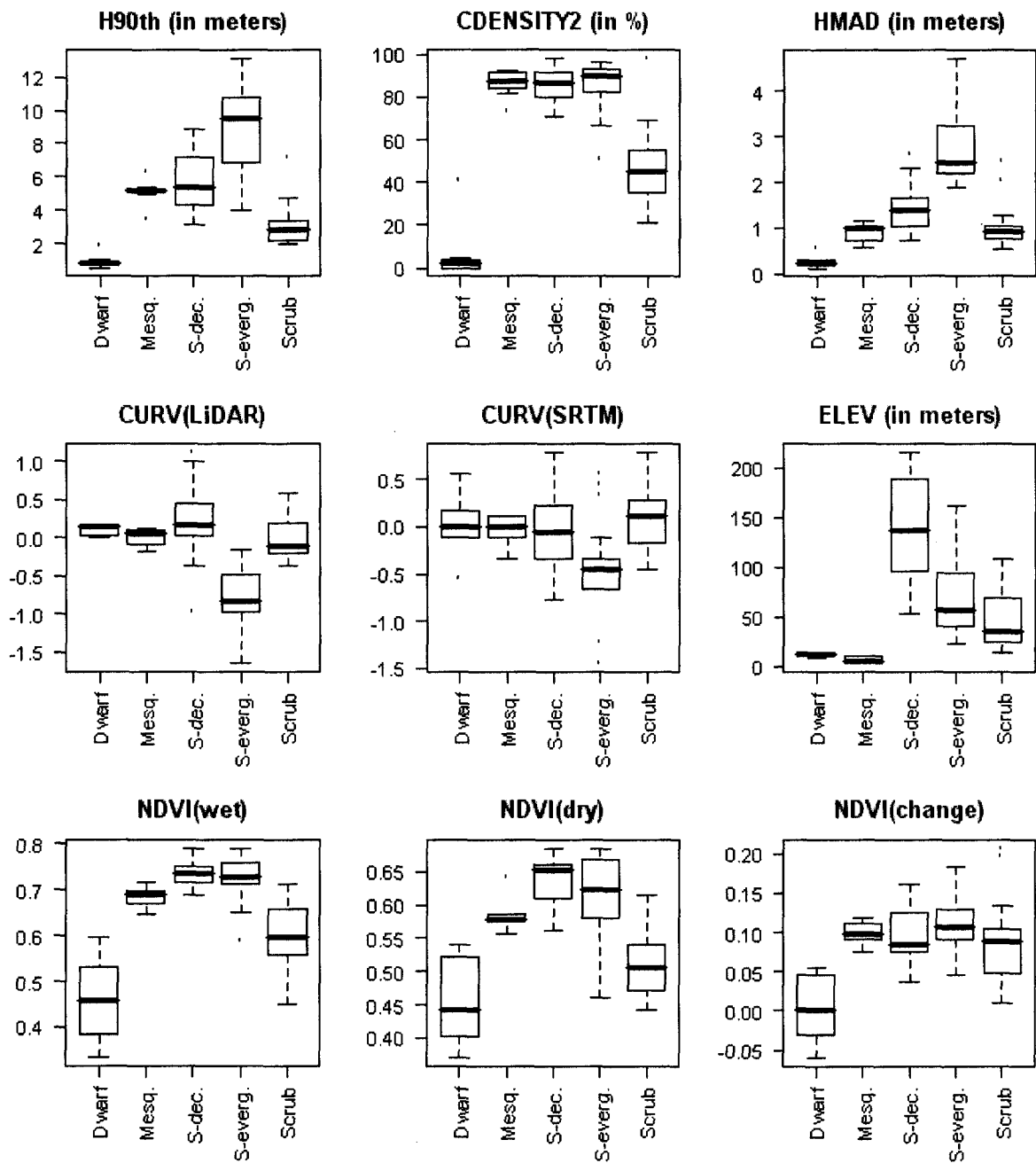


Figure 4. Forest types vs. remotely sensed predictor variables. Mesq. = Mesquite; S-dec. = Semi-deciduous; S-everg. = Semi-evergreen.

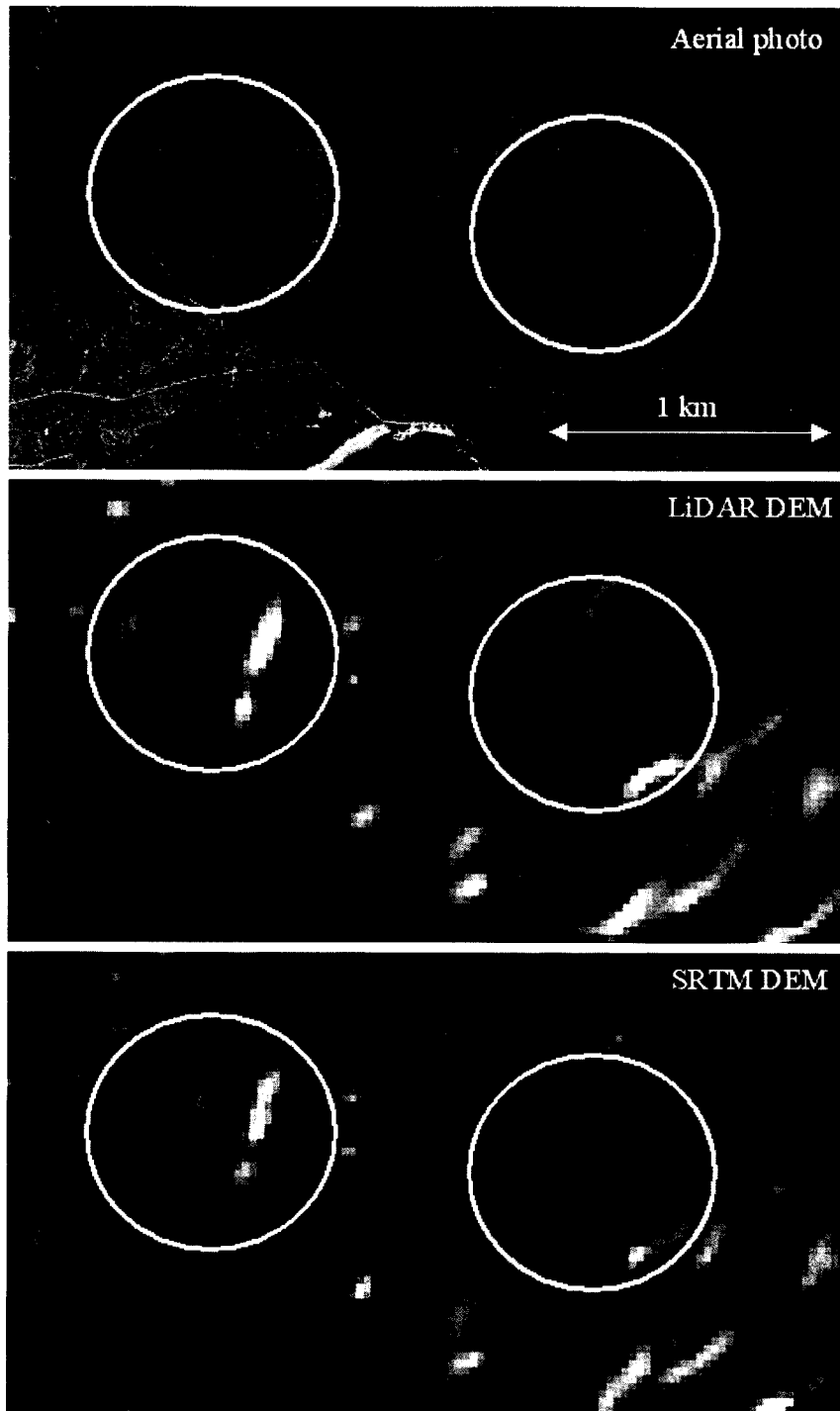


Figure 5. Topographic detailed captured by the 30-meter pixel LiDAR and SRTM digital elevation models.

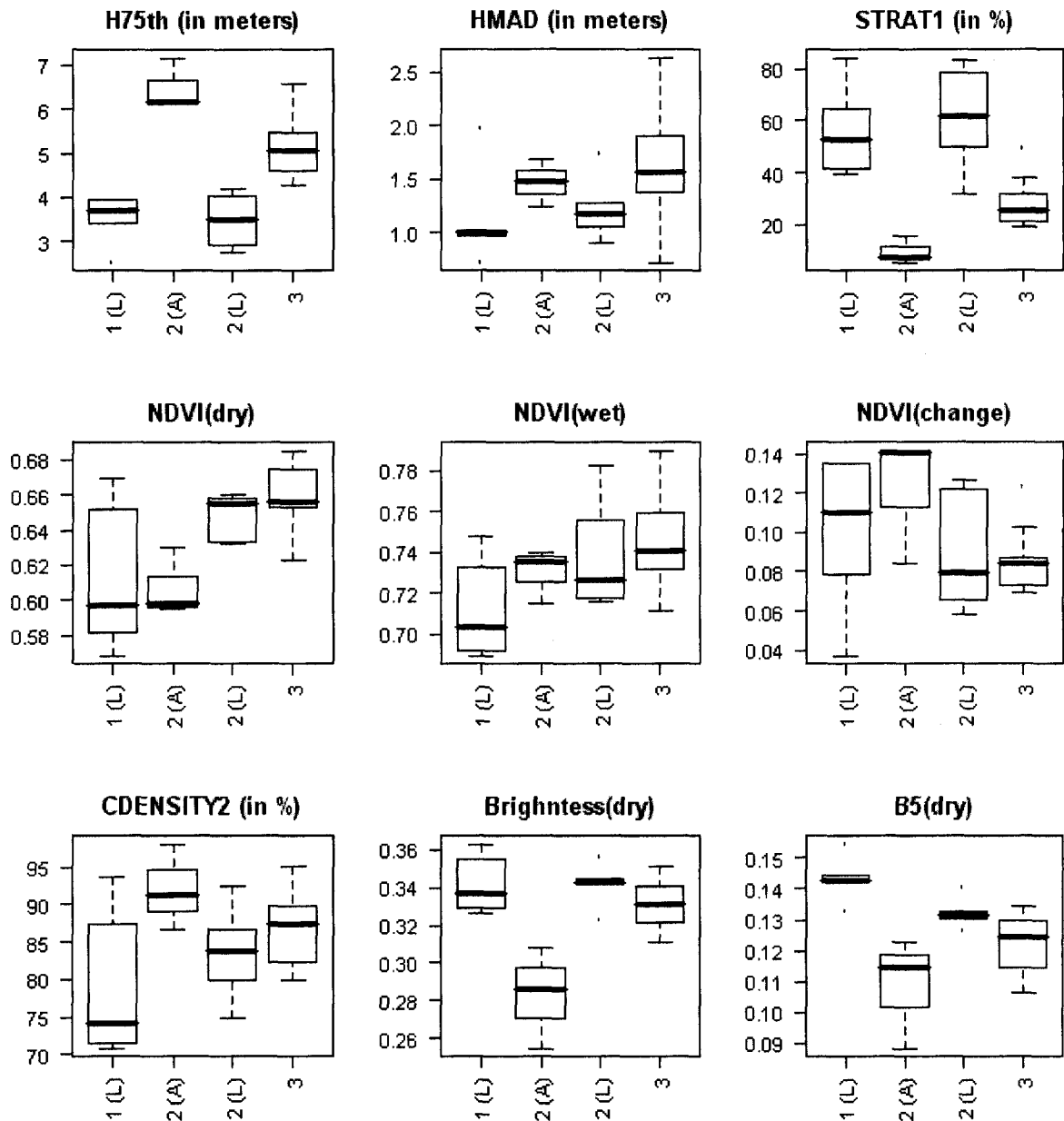


Figure 6. Successional classes vs. remotely sensed predictor variables. 1 = Mid-secondary forest; 2 = Late-secondary forest; 3 = Primary forest; (A) and (L) describe agricultural or logging-based past land uses respectively.

Table 1. Forest types in Guánica based on Lugo et al. (1978), Lugo 2005, Murphy and Lugo (1986), Farnsworth (1993), and Gould et al. (2008). The number of field samples used by this study (N) is also presented.

Class name	N	Description
Semi-deciduous forest	28	Covers about 60 % of Guánica. Average tree height is 4.3 m with tallest trees between 7 m and 12 m. Most trees are less than 10 cm in diameter, with up to 14,000 trees per hectare and basal area about 20 m ² /ha. Canopy cover fluctuates from 97 % in the wetter month to 77 % in the drier month, and leaf area index from 4.3 to 2.1.
Semi-evergreen forest	21	Represents about 20 % of the area, in moist ravines and valleys with thicker soils, abundant leaf litter, and springs or runoff catchments. Trees are taller with some species reaching 10 to 15 meters, and high species diversity. Includes evergreen forests.
Scrub forest	19	Covers about 15% of Guánica, also known as shrubland/cactus. Open shrubland with cactus and widely spaced stunted trees found on areas with poor soil development and much exposed bedrock.
Dwarf forest	8	Along coastal rocky areas exposed to ocean winds and salt spray, and composed by gnarled and twisted trees with horizontal stem and canopy growing close to the ground. Believed to support the oldest trees in Puerto Rico.
Mesquite forest	7	Relatively homogenous stand of <i>Prosopis pallida</i> with a dense herbaceous understory.

Table 2. Remote-sensing based explanatory variables.

	Variable name	Description
LiDAR variables of canopy structure	CDENSITY1	Canopy density 1; % of returns > 0.3 m
	CDENSITY2	Canopy density 2; % of returns > 1.0 m
	CDENSITY3	Canopy density 3; % of returns > 2.0 m
	STRAT1	Stratum 1; % of vegetation returns between 0.3 m and 3 m
	STRAT2	Stratum 2; % of vegetation returns between 3 m and 8 m
	STRAT3	Stratum 3; % of vegetation returns > 8 m
	H10th	10th percentile of vegetation heights
	H25th	25th percentile of vegetation heights
	HMEDIAN	Median height of vegetation returns
	H75th	75th percentile of vegetation heights
	H90th	90th percentile of vegetation heights
	HMEAN	Mean height of vegetation returns
	HMAX	Maximum height of vegetation returns
	HMAD	Median absolute deviation of vegetation heights
	HSD	Standard deviation of vegetation heights
	HSKEW	Skewness of vegetation heights
	HKURT	Kurtosis of vegetation heights
	HIQR	Interquartile range of vegetation heights
Landsat ETM+ spectral variables	B1	Band 1
	B2	Band 2
	B3	Band 3
	B4	Band 4
	B5	Band 5
	B7	Band 7
	NDVI	Normalized Difference Vegetation Index
	TCAP-B	Tasseled cap brightness

TCAP-G Tasseled cap greenness

TCAP-W Tasseled cap wetness

(Note: The suffix "wet" or "dry" before the variable's name is added to identify the image season when needed [e.g. NDVI_{dry}]. Similar, the symbol "Δ" before the variable's name is added to denote multitemporal change [e.g. ΔNDVI]).

Topographic variables	ELEV	Elevation
	SLP	Slope
	ASP	Aspect
	CURV	Curvature
	SRAD	Area Solar Radiation
	IMI	Integrated Moisture Index (Iverson et al. 1997)
	CTI	Compound Topographic Index (Gessler et al. 1995)
(Note: Derived from both 30-meter LiDAR and SRTM DEMs)		
Other	DIST	Distance to the coast

Table 3. Forest type classification models, including sources of predictor variables (top section), accuracy statistics (center) and variables included in the final models (bottom). Models 1-5 developed from single sources of predictor variables; models 6-9 from multiple sources, and models 10-11 after adding elevation (ELEV) and curvature (CURV).

Model	1	2	3	4	5	6	7	8	9	10	11
LiDAR canopy	X					X			X	X	
ETM+ wet season		X					X	X	X		X
ETM+ dry season			X				X	X	X		X
LiDAR DEM				X		X			X	X	
SRTM DEM					X			X			X
ELEV and DIST										X	X
Scrub forest	0.16	0.21	0.37	0.68	0.58	0.16	0.21	0.21	0.11	0.05	0.11
Semi-deciduous forest	0.21	0.46	0.43	0.25	0.46	0.14	0.46	0.21	0.14	0.07	0.25
Dwarf forest	0.13	0.38	0.38	0.50	0.75	0.13	0.38	0.38	0.13	0.13	0.25
Semi-evergreen forest	0.29	0.86	0.67	0.05	0.57	0.14	0.76	0.67	0.14	0.10	0.57
Mesquite forest	0.43	0.57	0.43	0.14	0.29	0.14	0.29	0.43	0.14	0.00	0.00
Overall error rate	0.23	0.51	0.47	0.31	0.53	0.14	0.46	0.36	0.13	0.07	0.28
Kappa	0.69	0.32	0.37	0.58	0.30	0.81	0.39	0.51	0.82	0.90	0.63

Model 1 - HMAD, H90th, CDENSITY2, HMAX, H75th, CDENSITY3, STRAT2

Model 2 - NDVI, TCAP-W, TCAP-G

Model 3 - NDVI, TCAP-W, TCAP-G, TCAP-B

Model 4 - CURV, IMI, CTI, SRAD, SLP

Model 5 - SRAD, CURV, SLP, ASP

Model 6 - CURV, CDENSITY2, HMAD, H90th, CTI, CDENSITY3, SLP, HMAX, H75th, STRAT2, HMEDIAN, CDENSITY1, STRAT1

Model 7 - NDVI(wet), TCAP-W(dry), TCAP-G(wet), TCAP-W(wet), ΔTCAP-B, NDVI(dry), TCAP-

G(dry), $\Delta B2$, $\Delta B5$

Model 8 - NDVI(wet), TCAP-G(wet), TCAP-W(dry), NDVI(dry), TCAP-W(wet), $\Delta B2$, $\Delta TCAP-B$, TCAP-G(dry), CURV, $\Delta B5$, TCAP-B(wet), SLP, ASP, $\Delta B1$, CTI

Model 9 - CURV, HMAD, H90th, CDENSITY2, CTI, HMAX, H75th, CDENSITY3, B2(wet), STRAT2, TCAP-G(wet), SLP, HMEDIAN, TCAP-W(wet)

Model 10 - CURV, ELEV, DIST, HMAD, CDENSITY2, H90th, CTI, CDENSITY3

Model 11 - ELEV, DIST, NDVI(wet), TCAP-W(dry), TCAP-G(wet), TCAP-W(wet), NDVI(dry), $\Delta TCAP-B$, $\Delta B2$, CURV, TCAP-G(dry), SLP, $\Delta B5$, TCAP-B(wet), CTI, ASP

CONCLUSIONS AND FUTURE WORK

This dissertation evaluated the consequences of integrating LIDAR remote sensing data into current assessments of wildlife habitat distribution and conservation. The main findings were:

- (1) *Species distribution models for a large number of wildlife species can immediately benefit from LiDAR-derived variables of canopy structure (Chapter 1).* In the Inland Northwest (USA), for example, this represents about 30% of the avian species and 20% of the mammal species, including several species of great conservation need. The most important habitat variables appeared to be: tree canopy cover, shrub canopy cover (including understory shrubs), some measure of stand biomass (mean tree diameter / basal area / age), shrub height, and size and density of snags. Furthermore, the number of wildlife species that could potentially benefit from such data is expected to increase since the structural habitat preferences of many species are either unknown or often not reported.
- (2) *LiDAR data can be used to map the distribution of key wildlife habitat features in forests, such as snags and understory shrubs (Chapter 2).* In a mixed-conifer forest of northern Idaho, the combination of LiDAR-derived metrics of canopy height and topography allowed us to predict the distribution of snags and understory shrubs with an accuracy of > 80%. This represents an important step towards improved assessments of forest structure and wildlife habitat relative to previous efforts using traditional remote sensing technologies.
- (3) *LiDAR improve the mapping of wildlife habitat suitability at local scales (Chapter 2).* For the first time, avian-habitat relationships that have been known for decades (e.g. woodpeckers with specific snag sizes, flycatchers with dense understory areas, etc) were possible to map with the help of LiDAR. This resulted in refined patterns of habitat distributions that better reflect species' preferences for particular conditions of forest structure. In this sense, LiDAR data helped reduce the gap between the environmental variables that are important for assessing wildlife habitat, and the variables that we can effectively map.

- (4) *Novel data of vegetation structure impacts traditional GAP analysis assessments (Chapter 3)*. We observed that the incorporation of vegetation structure into a recent Idaho GAP Analysis Project resulted in: improved species-habitat models (i.e. that better reflect known species-habitat relationships), smaller predicted habitats (i.e. 15% to 68% smaller than those previously developed without information about vegetation structure), higher model accuracy by reduced commission errors (one of the major criticism to GAP data), and lower values of species richness calculated at the Landsat pixel scale. On the other hand, species richness values calculated at the GAP hexagon scale, as well as the species representation within the network of protected lands, did not change after adding vegetation structure information. Surprisingly, all of these changes resulted from the incorporation of only one aspect of vegetation structure (i.e. tree canopy cover).
- (5) *LiDAR data facilitate forest land cover mapping (Chapter 4)*. Forest land cover maps are the main source of environmental geospatial data for assessing wildlife habitat and species distribution; however, accurate maps of detailed forest types have been difficult to obtain. In a complex tropical forest, we observed that the use of LiDAR data increased forest type classification from 60% (using Landsat) to about 90%. Both canopy and topographic data from LiDAR seemed important for refining forest land cover classifications.
- (6) *LiDAR data facilitate forest succession assessments (Chapter 4)*. Because forest succession is one of the major forces affecting forest structure, it is also a key variable for determining wildlife habitat. However, forest successional stages have been typically difficult to map with traditional remote sensing data (e.g. Landsat). Our efforts in temperate and tropical landscapes showed the value of LiDAR for accurately separating successional classes in forests (i.e. accuracies >90%; see Falkowski et al., 2009 and Chapter 4). This included the ability to identify remaining patches of old growth forests, which are important components for biodiversity and conservation, and represent one of the most endangered ecosystems in the United States (Noss et al., 1995).

Expanding and testing the findings from this thesis are important future steps. Although very promising for advancing wildlife habitat and conservation assessments with remote sensing, the research presented by this thesis warrants further applications in other landscapes and involving other wildlife species in order to fully evaluate the potential of LiDAR data. For example, can snags and understory shrubs be mapped from LiDAR in other forest types? Which are the most important variables of vegetation structure for assessing wildlife habitat in broad-leaf forests? To what degree can LiDAR improve wildlife habitat assessments for critical groups of species such as amphibians?

Great opportunities also exist to further advance wildlife habitat modeling using LiDAR-based structure mapping. These include, for example, improving our understanding of the scales at which species respond to the environment, and identifying potential relationships between ecosystem 3-D structure and habitat quality. While most of the studies –including this thesis- have used LiDAR to identify areas where wildlife species are most likely to occur (i.e. habitat suitability approach, or presence/absence analyses), only few efforts have used LiDAR to identify areas where species “can do” better (i.e. to evaluate habitat quality or fitness, and considering survival and reproduction, e.g. Hill et al., 2004). Because such information is critical for managers, and because our knowledge about the local habitat characteristics that influence fitness is poor, evaluating the use of LiDAR for assessing habitat quality emerges as an important need.

In addition, measuring vegetation structure at spatial scales that are relevant for wildlife is critical for developing accurate habitat-association models. However, such work has been limited due to logistical constraints to measuring vegetation structure in the field. As a result, our knowledge of the relationships between wildlife and vegetation structure is restricted to field plots of 1-ha or less in size. In this thesis, for example, I mapped habitat suitability using indices that are based on information at 1-ha scale (Chapter 2). LiDAR, on the other hand, allows forest structure characteristics to be quantified at varying grain and extent, and therefore provides a unique opportunity to identify the scale at which species respond to vegetation structure. By doing so, for example, a recent study by Seavy et al. (2009) found that avian relationships with canopy

height in a riparian forest were best explained at scales that ranged between 0.2 ha to 50 ha, depending on the species. These findings can be used to further enhance the biological meaning of geospatial data sets, and therefore to advance current assessments of wildlife habitat distribution and conservation such as GAP analysis.

The increasing availability of LiDAR data opens new opportunities for the GAP Analysis Program. In previous GAP efforts, the lack of geospatial data about forest structure for supporting wildlife habitat models, and the difficulties for accurately separating forest land cover types, have been major limitations (e.g. Laba et al. 2002; Scott et al., 2002; Lawry et al., 2005; McKerrow et al., 2006). This thesis showed that LiDAR data can be used to refine GAP habitat distribution models, as well as to improve forest land cover mapping; and furthermore, that refined habitat data can significantly impact many of the results from GAP.

Today, ten states in the country have wall-to-wall LiDAR data available (Jason Stoker, personal communication). With remote sensing data at the core of GAP, the presence of these powerful -and freely available- datasets is definitively good news. Improving GAP assessments with LiDAR data is recommended, as it would increase the relevance of GAP data for conservation planning today, and help prepare GAP for tomorrow, as LiDAR collections become even more widely available. A critical step for making this happen would be to develop approaches that facilitate the operational use of large (i.e. State-wide) LiDAR datasets.

Finally, further efforts should also explore the use of satellite-based LiDAR data, as well as the integration of LiDAR with radar, multispectral and/or hyperspectral imagery for assessing wildlife habitat. Indeed, a recent paper by Bergen et al. (2009) evaluated the potential of future LiDAR and radar spaceborne missions for assessing biodiversity and habitat, and Asner & Martin (2009) showed that the integration of LiDAR with hyperspectral data can tremendously advance assessments of vegetation diversity in forests.

References

- Asner GP, & Martin RE. (2009). Airborne spectranomics: mapping canopy chemical and taxonomic diversity in tropical forests. *Frontiers in Ecology and the Environment*, 7, 269-276.
- Bergen KM, Goetz SJ, Dubayah RO, Henebry GM, Hunsaker CT, Imhoff ML, Nelson RF, Parker GG, & Radeloff VV. (2009). Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications LiDAR and radar spaceborne missions. *Journal of Geophysical Research*, 114, G00E06:13.
- Hill RA, Hinsley SA, Gaveau DLA, & Bellamy PE. (2004). Predicting habitat quality for great tits (*Parus major*) with airborne laser scanning data. *International Journal of Remote Sensing*, 25, 4851–55.
- Falkowski MJ, Evans JS, Martinuzzi S, Gessler PE, & Hudak AT. (2009). Characterizing Forest Succession with Lidar Data: an Evaluation for the Inland Northwest USA. *Remote Sensing of Environment*, 113, 946-956.
- Laba ML, Gregory S, Braden J, Ogurcak D, Hill E, Fegraus E, Fiore J, & DeGloria SD. (2002). Conventional and fuzzy accuracy assessment of the New York Gap Analysis Project land cover map. *Remote Sensing of Environment*, 82, 443–456.
- Lowry J H, Ramsey RD, Boykin K, Bradford D, Comer P, Falzarano S, Kepner W, Kirby J, Langs L, Prior-Magee J, Manis G, O'Brien L, Sajwaj T, Thomas KA, Rieth W, Schrader S, Schrupp D, Schulz K, Thompson B, Velasquez C, Wallace C, Waller E, & Wolk B. (2005). Southwest Regional Gap Analysis Project: Final Report on Land Cover Mapping Methods, RS/GIS Laboratory, Utah State University, Logan, Utah.
- McKerrow AJ, Williams SG, & Collazo JA. (2006). The North Carolina Gap Analysis Project: Final Report. North Carolina Cooperative Fish and Wildlife Research Unit, North Carolina State University, Raleigh, North Carolina. xiv + 110 pp. + appendices.
- Noss RF, La Roe ET III, & Scottt JM. (1995). Endangered ecosystems of the United States: a preliminary assessment of loss and degradation. Biological Report 28, US Department of the Interior, National Biological Service, Washington, DC.

Scott JM, Peterson CR, Karl JW, Strand E, Svancara LK, & Wright NM. (2002). A Gap Analysis of Idaho: Final Report. Idaho Cooperative Fish and Wildlife Research Unit. Moscow, ID, USA.

Seavy, NE, Viers, JH, & Wood, JK. (2009). Riparian bird response to vegetation structure: a multiscale analysis using LiDAR measurements of canopy height. *Ecological Applications*, 19, 1848-1857