ABSTRACT

Title of Document: MAPPING FOREST STRUCTURE AND HABITAT CHARACTERISTICS USING LIDAR AND MULTI-SENSOR FUSION

Anuradha Swatantran, PhD, 2011

Directed By: Prof Ralph Dubayah, Department of Geography

This dissertation explored the combined use of lidar and other remote sensing data for improved forest structure and habitat mapping. The objectives were to quantify aboveground biomass and canopy dynamics and map habitat characteristics with lidar and /or fusion approaches. Structural metrics from lidar and spectral characteristics from hyperspectral data were combined for improving biomass estimates in the Sierra Nevada, California. Addition of hyperspectral metrics only marginally improved biomass estimates from lidar, however, predictions from lidar after species stratification of field data improved by 12%. Spatial predictions from lidar after species stratification of hyperspectral data also had lower errors suggesting this could be viable method for mapping biomass at landscape level. A combined analysis of the two datasets further showed that fusion could have considerably more value in understanding ecosystem and habitat characteristics.

The second objective was to quantify canopy height and biomass changes in in the Sierra Nevada using lidar data acquired in 1999 and 2008. Direct change detection showed overall statistically significant positive height change at footprint level (ΔRH100 = 0.69 m, +/- 7.94 m). Across the landscape, ~20 % of height and biomass changes were significant with more than 60% being positive, suggesting regeneration from past disturbances and a small net carbon sink. This study added further evidence to the capabilities of waveform lidar in mapping canopy dynamics while highlighting the need for error analysis and rigorous field validation

Lastly, fusion applications for habitat mapping were tested with radar, lidar and multispectral data in the Hubbard Brook Experimental Forest, New Hampshire. A suite of metrics from each dataset was used to predict multi-year presence for eight migratory songbirds with data mining methods. Results showed that fusion improved predictions for all datasets, with more than 25% improvement from radar alone. Spatial predictions from fusion were also consistent with known habitat preferences for the birds demonstrating the potential of multi- sensor fusion in mapping habitat characteristics. The main contribution of this research was an improved understanding of lidar and multi-sensor fusion approaches for applications in carbon science and habitat studies.

.

MAPPING FOREST STRUCTURE AND HABITAT CHARACTERISTICS USING LIDAR AND MULTI-SENSOR FUSION

By

Anuradha Swatantran.

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2011

Advisory Committee:
Professor Ralph Dubayah, Chair
Research Associate Professor Michelle Hofton
Professor Christopher Justice
Professor James Dietz
Assistant Professor Tatiana Loboda

Dedication

For my parents,
Rajalakshmi and Raghunath Swatantran

Acknowledgement

I would like to express my deepest gratitude to my advisor Ralph Dubayah. He not only gave me innumerable opportunities to learn and grow as a researcher, but also tough challenges that pushed me beyond the boundaries of everything I knew. For teaching me to think like a scientist and 'look under every rock', I will always remain indebted to you. I am especially thankful to Michelle Hofton for helping me decipher cryptic lidar waveforms and for showing me the value of humor in stressful grad student life. I am very thankful to my other committee members; Chris Justice, Jim Dietz and Tatiana Loboda for giving me valuable and constructive feedback and helping me set up meetings with incredible ease, despite their busy schedules.

I could not have started this research without Wenge Ni-Meister, who introduced me to lidar remote sensing in the first place. This work would have also not been possible without lidar data from Bryan Blair and the LVIS team at NASA. I am extremely grateful to Dar Roberts for all his guidance with hyperspectral image processing, for patiently answering my endless queries and making valuable contributions in my paper. I acknowledge all my friends from the Viper Lab who made my stay at UC Santa Barbara a memorable experience. Many thanks to Marc Simard, Naiara Pinto and Zhiyu Zhang for all their guidance with radar data.

A big 'thank you' to Mike O'Connell, Anupam Anand, Amanda Whitehurst, Lee Ann King and all my lab mates for making the VCL lab lively and field work thoroughly enjoyable. You have enriched my life in more ways than you know. I am especially thankful to Sage Sheldon for walking me through the NASA proposal submission when I just moved here and was clueless. I will always cherish the good times we had at conferences after stressful semesters. I am thankful to all my friends particularly Debjani Ghatak, Prasad Pathak, Neeti Sinha and Angira Baruah for their valuable suggestions and discussions.

I am forever grateful to my friend, philosopher and guide Madhura Niphadkar, without whom this research would have remained a distant dream. I am indebted to all my teachers from Kamalnayan Bajaj School, Naik Sir and Prof C. K. Desai for teaching me the basics of everything. I owe my scientific curiosity and passion for work to my brother, Bharat Swatantran.

I cannot describe how fortunate I am to have the love, support and encouragement of two amazing families; the Swatantran and the Ranganathan family. Finally, I could have never come this far without my better half, Akshay Ranganathan, who not only encouraged me to pursue my dream but also walked 'the whole nine yards' with me to get there.

This dissertation was funded by a NASA Earth System Science Fellowship.

Contents

Dedication ii
Acknowledgement iii
Contents iv
List of Tables vi
List of Figures vii
Chapter 1 Forest Structure and Habitat Characteristics using Multi-sensor Data Fusion 1
1.1 Motivation *1*
1.2 Background *4*
1.2.1 Remote Sensing for carbon science 4
1.2.2 Remote sensing for biodiversity studies 5
1.2.3 Lidar, Radar and Hyperspectral Remote Sensing 6
1.2.4 Multi-sensor fusion 8
1.3 Dissertation Outline *9*
1.4 Summary *12*
Chapter 2 Mapping Biomass and Stress in the Sierra Nevada using Lidar and Hyperspectral Data Fusion 13
2.1 Introduction *13*
2.2 Background *15*
2.3 Study Area and Data *18*
2.3.1 Study Area 18
2.3.2 Data 19
2.4 Methods *21*
2.4.1 Field attributes 22
2.4.2 LVIS metrics 22
2.4.3 AVIRIS spectral metrics 24
2.4.4 MESMA fractions from AVIRIS 26
2.4.5 Land-cover classification from AVIRIS 29
2.5 Analysis *30*
2.5.1 Stand Level 30
2.5.2 Landscape Level 31
2.6 Results *32*
2.6.1 Stand Level 32
2.6.2 Landscape level 36
2.7 Discussion *42*
2.8 Conclusion *47*
Acknowledgements *49*
Chapter 3 Mapping Canopy Height and Biomass Dynamics in the Sierra Nevada using Waveform Lidar 50
3.1 Introduction *50*
3.2 Study Area and Data *52*

3.2.1 Study Area 52
3.2.2 Field Data 53
3.2.3 LVIS Data 54
3.3 Methods *56*
3.3.1 LVIS and field measured changes 57
3.3.2 Coincident footprint dynamics 58
3.3.3 Landscape level changes from LVIS 61
3.4 Results *63*
3.4.1 LVIS and field measured changes 63
3.4.2 Coincident footprint dynamics 65
3.4.3 Landscape level canopy height and biomass changes 68
3.5 Discussion *73*
3.5.1 LVIS and field-measured changes 75
3.5.2 Coincident footprint dynamics 76
3.5.3 Landscape changes in canopy height and biomass 77
3.6 Conclusion *80*
Chapter 4 Mapping Bird Habitat Quality in New Hampshire using Radar, Lidar and Multispectral Fusion 81
4.1 Introduction *81*
4.2 Study Area and Data *84*
4.2.1 Study Area 84
4.2.2 Bird Data 85
4.2.3 Radar Data 86
4.2.4 Small-Footprint lidar Data 88
4.2.5 Medium-footprint lidar data 89
4.2.6 Landsat Data 91
4.3 Methods *91*
4.3.1 Predicting prevalence 93
4.3.2 Variable importance 94
4.3.3 Prevalence at plot and landscape scales 95
4.4 Results *95*
4.4.1 Predictive capabilities of different datasets for prevalence 95
4.4.2 Variable importance in predictions 97
4.4.3 Predictions at plot and landscape scale 98
4.5 Discussion *104*
4.5.1 Predictive capabilities of different datasets for prevalence 104
4.5.2 Variable importance in predictions 106
4.5.3 Predictions at plot and landscape scale 108
4.6 Conclusion *110*
Chapter 5 Discussion and Conclusion 112
Bibliography 121

List of Tables

Table 1-1 Forest structure and habitat characteristics directly measured or modeled from Lidar / Fusion (adapted from Bergen et al., 2009) 7

Table 2-1 Distribution of field plots by Wildlife Habitat Relation (WHR) Type 20

Table 2-2 Hyperspectral metrics calculated using AVIRIS data for 1ha field plots included vegetation indices, red edge derivatives, Ligno-cellulose band ratios and MESMA fractions of green vegetation, soil/NPV and shade. 25

Table 2-3 Reference endmembers used in 10 3 endmember MESMA models for unmixing AVIRIS images. Soil and non photosynthetic vegetation (NPV) spectra were grouped into one class. Fit metrics - EAR, MASA and COB values were used to select the best representative spectra to un-mix the entire image. 28

Table 2-4 The predictive power of AVIRIS metrics alone , LVIS metrics alone , LVIS + AVIRIS metrics and LVIS metrics after species stratification of field data was tested over 125 1ha plots. Suitable predictor variables were selected using AIC criteria. The best model was obtained by predicting biomass with LVIS variables after stratifying field plots into WHR/species type 34

Table 2-5 Biomass was predicted using a single lidar equation before species stratification and a different equation for each species after stratification of field plots by WHR type. 37

Table 2-6 Error matrix for level 1 (genera/species) classification of AVIRIS images using 47 2-endmember MESMA models. Overall Accuracy = 87.7 % Kappa =0.86, Kappa variance = 0.001 38

Table 2-7 Error matrix showing classification accuracy at hectare level. WHR types from field polygons were used as ground reference. Overall accuracy = 0 69.5% 38

Table 3-1 Changes in LVIS RH metrics from nearly co-incident footprints. 66

Table 3-2 Transition matrix for LVIS canopy heights (RH100) co-incident footprints in 1999 and 2008. Reading down columns gives prospective transition probabilities. Reading across rows gives the retrospective transition probabilities. The diagonal shows the probability of staying in the same class. 67

Table 4-1 Common and scientific names for songbird species 86

Table 4-2 Metrics calculated from the different datasets. 92

List of Figures

Fig. 1-1 Improved understanding of interactions between vegetation, carbon flux and biodiversity requires quantitative assessments of forest vertical structure, floristic composition and biochemical characteristics, emphasizing the need for multi-sensor fusion approaches. 3

Fig. 1-3 Outline of dissertation chapters. 9

Fig. 2-1 Study area in the Sierra National Forest showing 1ha field plots. Plots are classified based on Wildlife Habitat Relation (WHR) type 18

Fig. 2-2 Example of an LVIS waveform centered on a field plot with area 0.0 7 ha. The amplitude of the waveform is proportional to energy reflected from canopy and ground. Metrics calculated from the waveform include ground elevation, quartile heights of energy return (RH25, RH50, RH75 & RH100) and canopy cover. 23

Fig. 2-3 Examples of reflectance spectra extracted from AVIRIS images over the study area. A set of 19 band ratios describing vegetation characteristics such as chlorophyll content, water content, stress were calculated from the visible, near infrared and short wave infrared wavelengths..... 26

Fig. 2-4 Subset of images showing endmember fractions generated using MESMA. Bright areas have high fractional abundance and dark areas have low abundance. GV, soil/NPV and shade fractions were summarized to calculate mean and std. deviation of values for 1ha plots. 28

Fig. 2-5 MESMA was used to map landcover and dominant vegetation types from AVIRIS images. AVIRIS maps at a resolution of 3.3m (a) were aggregated and dominant vegetation type at 1ha was identified (b). Labels show WHR types from field plots. *Forests classified with white fir as dominant vegetation type at 1ha also had a mixture of conifers and were grouped as the SMC type for biomass estimation. 30

Fig. 2-6 Biomass predicted for 125 field plots at 1ha scale using various metrics: (a) AVIRIS; (b) LVIS; (c) LVIS and AVIRIS metrics; (d) LVIS after species stratification of field data. 33

Fig. 2-7 Co-efficients of determination for predicted biomass with 95% confidence intervals. Narrower confidence intervals for LVIS + species stratification suggest a small improvement, but overlap of intervals shows that it is not statistically significant..... 35

Fig. 2-8 Landscape maps of biomass were generated from LVIS before (8a) and after species stratification of AVIRIS imagery (8b), using equations in Table 2-4 & 2-5. Forests dominated by pine and hardwood species (e.g. black rectangle) show more spatial variations in predicted biomass. 39

Fig. 2-9 Histograms showing differences between biomass predicted before (9a, 9c) and after (9b, 9d) species stratification of AVIRIS imagery. Stratification for hardwoods and pines increased predicted values in low (<50 Mg/ha) ranges and decreased values in high ranges (>200 Mg/ha) of biomass. 40

Fig. 2-10 Landscape maps of biomass, canopy cover, NDWI and D1GVI used for detecting water and chlorophyll stress in high biomass forests. 41

Fig. 2-11 Map showing high and low stress in stands with high biomass (>200 Mg/ha). Height of bars represents biomass values. 41

Fig. 2-12 Spatial distribution of non-photosynthetic vegetation (NPV) fractions within canopies for one AVIRIS image (1 ha level). Stands with high biomass and stress (Fig.2-11) also showed high NPV values. 42

Fig. 3-1 Study site in the Sierra National Forest, California showing field plots measured in 2000/2001 and 2008. 53

Fig. 3-2 Schematic layout of field plots in 2000/2001 and 2008. The 2008 square subplots were slightly larger than the inner circular plots to account for geolocation shifts between field and lidar data. Plots were oriented upslope. 54

Fig. 3-3 Flowchart showing summary of methods and LVIS RH/RH_E metrics used. 57

Fig. 3-4 Geolocation shifts between field plot locations and lidar data affected comparisons of LVIS and field metrics. The two 1999 LVIS footprint centers are 7m and 14.5m away from plot center respectively. The 2008 LVIS data (only footprint centroid shown) oversampled field plots but were not co-incident with field plot centers. 58

Fig. 3-5 Average ground elevation error between 1999 and 2008 (Δ ZG) as a function of distance between lidar footprints. Errors were lowest at 15% trimmed means and closest to zero for shots between 2 and 2.5m distance from each other. This suggests an average relative geo-location error of 2 - 2.5m between 1999 and 2008 footprints. 59

Fig. 3-6 Maximum field height was strongly correlated with LVIS RH100 for both years with similar errors. 64

Fig. 3-7 Field biomass was strongly correlated with LVIS RH50 for individual years with similar errors. Note that these equations were not used to predict biomass for 1999 and 2008 across landscape. 64

Fig. 3-8 Distribution of maximum field height changes relative to 1999 heights (a) 4 plots showed height decreases greater than 4m. The remaining show small changes, with increases slightly exceeding decreases. Biomass change relative to 1999 biomass (b)Two plots showed more than 50Mg/ha biomass gain and two less than 50Mg/ha loss. Both height and biomass changes were too small to be detected by LVIS metrics. 65

Fig. 3-9 Probability of transition to a higher height class is greater for shorter canopies, < 40m. In taller height classes (> 60m), losses increase exponentially and probability of staying in the same height class is very low..... 67

Fig. 3-10 Canopy Height (RH100) projections from the transition matrix at successive nine year time intervals (t) from 1999 to steady state in 2431 (t = 38)..... 68

Fig. 3-11 Canopy height changes were mapped across the landscape by directly calculating average ΔRH_E100 from 1999 and 2008 LVIS data. Inset shows losses and gains at 0.07 ha scale over the Teakettle Experimental Forest..... 70

Fig. 3-12 Statistically significant RH_E100 changes were mapped across the landscape using non parametric tests. Inset shows losses and gains at both scales over the Teakettle Experimental Forest. Histogram shows net positive change..... 70

Fig. 3-13 Regression equations derived for predicting biomass from LVIS and 2000/2001 field data (a) and all 2008 subplots (b). The 1999 equation was used to predict biomass for both years. 71

Fig. 3-14 Biomass changes predicted from LVIS (2008 – 1999) RH metrics at plot and hectare scales. Inset shows losses and gains at 0.07 ha scale over the Teakettle Experimental Forest. 71

Fig. 3-15 Statistically significant biomass changes were mapped across the landscape. Inset shows losses and gains at both scales over the Teakettle Experimental Forest. Histogram shows net positive change. 72

Fig. 3-16 Spatial overlay of average canopy top elevation (ΔRH_E100) change map with aerial photos showing areas with significant gain (a -> b) and significant loss (c-> d). Note that visual change in percent canopy cover was only an approximate estimate and more reliable for large changes than small. 74

Fig. 3-17 Areas sensitive for wildlife showed overall increase in height. Forests under uneven aged management also had net height increase but less than protected areas. Low stress areas showed increase in height while high stress areas showed an overall decrease. Changes were statistically significant. 74

Fig. 4-1 Study area in the Hubbard Brook Experimental Forest showing plot transects laid out by Schwarz (2001). Bird point count data were collected over 371 plots by Doran et al., (2005). 85

Fig. 4-2 UAVSAR backscatter images from horizontal transmitted horizontal return (HH), vertical transmitted vertical return (VV) and Horizontal transmitted vertical return (HV) polarizations. 87

Fig. 4-3 Discrete return lidar canopy height map showing individual tree locations (inset) detected by TreeVaW. Crown radius and height are calculated for each tree by the algorithm. 89

Fig. 4-4 Components of an LVIS waveform (left) and canopy top height map showing vertical and spatial distribution of heights. Heights of 25%, 50%, and 75% lidar energy returns (RH metrics) are calculated from the waveform. The bold solid line shows the normalized cumulative energy return from which canopy cover at every 5m interval was calculated. 90

Fig. 4-5 Normalized difference vegetation index from Landsat data (leaf off) and seasonal NDVI change between from leaf off and leaf on data. High NDVI difference values show deciduous cover, low values show conifers. 91

Fig. 4-6 Random Forests regression results for bird prevalence using radar, lidar and Landsat data individually and in combination with each other. 96

Fig. 4-7 Comparison of variance explained from models using radar metrics alone with those using all metrics. 97

Fig. 4-8 Variable selection within 10 most important predictors for all species taken together. Model used : UAVSAR +LVIS +DRL + Landsat. 98

Fig. 4-9 Histograms of actual (left) versus predicted prevalence from models with all metrics combined.(right) . For species codes refer to Table 4-1. 100

Fig. 4-10 Histograms of actual (left) versus predicted prevalence from models with all metrics combined (right). For species codes refer to Table 4-1. 101

Fig. 4-11 Predicted prevalence from radar metrics alone (left) and all metrics together (right) . For species codes refer to Table 4-1. 102

Fig. 4-12 Predicted prevalence from radar metrics alone (left) and all metrics together (right). For species codes refer to Table 4-1. 103

Fig. 4-13 Quantile predictions of habitat quality using Random Forest Regression. 104

Chapter 1 Forest Structure and Habitat Characteristics using Multi-sensor Data Fusion

1.1 Motivation

Forest structure, including the vertical arrangement of vegetation within canopies, the horizontal distribution across landscapes, together with species composition comprises 'habitat' and influences biogeochemical fluxes of carbon in the earth-atmosphere system (Fig 1-1). Managing forests for carbon and/or biodiversity under increasing anthropogenic pressures and a changing climate therefore requires improved quantitative measurements of forest vertical structure, floristics and biogeochemical status. Improving our capabilities to simultaneously map these attributes can go a long way in advancing critical research in carbon science, ecosystem and biodiversity studies. For instance, accurate maps of vertical structure can reduce uncertainties in carbon storage and flux estimates, which are currently between 30 and 70% of the total terrestrial budget (Canadell et al., 2007). Better maps of structural and biogeochemical composition of vegetation can improve our understanding of many ecosystems processes. Three dimensional forest structure and composition combined with habitat preferences can inform biodiversity conservation. The overarching questions are how can such integrated information be obtained and how can currently available remote sensing resources bridge the gap in knowledge?

Remote sensing has greatly advanced forest mapping and monitoring capabilities in the last few decades by imaging the earth at various spatial and

temporal scales. Although multispectral airborne and space borne sensors have significantly improved spatial maps of vegetation and landcover, they have shown limited capabilities in measuring vertical structure critical for ecosystem studies. In addition, most multispectral metrics, such as the normalized difference vegetation index (NDVI) have not proven effective in monitoring the biogeochemical status of canopies, particularly in high biomass density forests (> 100 Mg/ha).

In recent years, many new sensors have been developed, each with complementary capabilities. Three of them are of particular interest in this research. The first is Light Detection and Ranging (lidar), which provides three dimensional forest structure measurements. The second is hyperspectral remote sensing that can be used to map biochemical characteristics and species level variations in vegetation. The third is radar remote sensing which is less sensitive to structure than lidar but has greater spatio-temporal coverage. While each of these technologies has been studied extensively, individually, the potential of combining them for mapping forest structure is yet to be thoroughly understood.

The overall goal of my dissertation is to explore applications of multi-sensor fusion for mapping forest structure and habitat characteristics. This is a vast area of research with many possible approaches. I use airborne data as a test bed for future space borne lidar /radar sensors such as the Deformation, Ecosystem Structure and Dynamics of Ice (DESDynI) and hyperspectral sensors such as Hyperspectral Infrared Imager (HyspIRI). My research focuses on three areas of interest: aboveground biomass estimation; canopy structural dynamics, and; quantification of habitat characteristics, all of which are key research requirements for carbon science

(Treuhaft et al., 2004) and biodiversity research (Bergen et al., 2006). Specifically, my objectives are to:

1. Test the combined use of lidar and hyperspectral data in mapping aboveground biomass using statistical fusion approaches.
2. Evaluate the potential of temporal waveform lidar in quantifying canopy height and biomass dynamics.
3. To integrate lidar, radar and multispectral data in a multi-sensor fusion approach for predicting bird habitat quality.

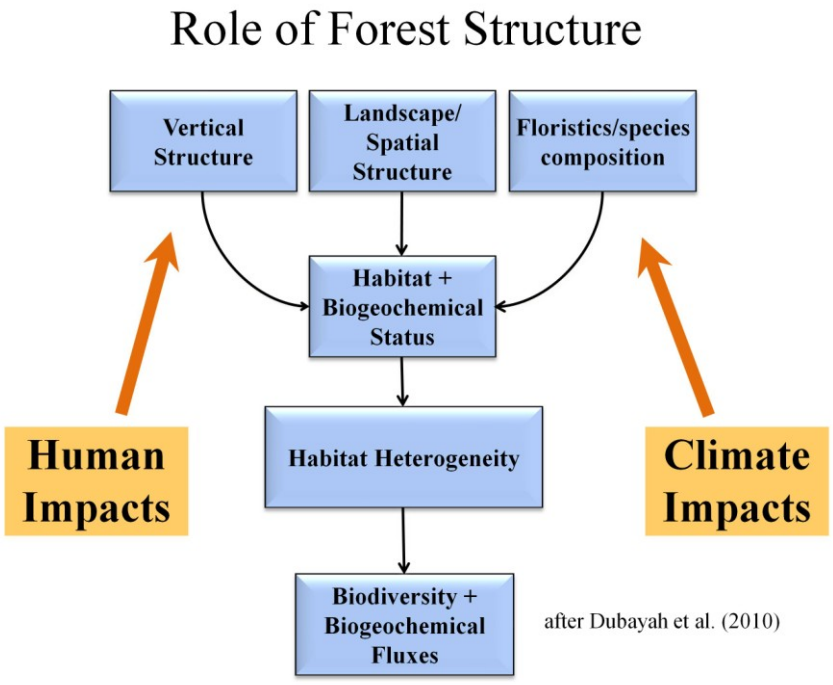

Fig. 1-1 Improved understanding of interactions between vegetation, carbon flux and biodiversity requires quantitative assessments of forest vertical structure, floristic composition and biochemical characteristics, emphasizing the need for multi-sensor fusion approaches.

1.2 Background

1.2.1 Remote Sensing for carbon science

Quantifying the amount and distribution of carbon in various reservoirs (oceans, atmosphere, terrestrial ecosystem, and fossil fuels) is critical for understanding ecosystem responses to climate change. Although sources and sinks in the other components of the carbon cycle are fairly well quantified, major uncertainties still exist in the terrestrial carbon budget (30 to 70%). These are associated with lack of quantitative estimates of carbon storage in forests and changes from disturbance (Canadell et al., 2007; Houghton and Goetz, 2008). Consequently, there are several gaps in our understanding of ecosystems and their interactions with bio-geochemical cycles.

One way of reducing these uncertainties is by quantifying aboveground biomass, about 50% of which is carbon. Remote sensing data have long been used to derive biomass by developing empirical relationships between spectral characteristics of images and field allometry. Although these methods are useful in low biomass ecosystems with densities less than 100 Mg/ha, they have limited sensitivity at moderate to high biomass levels (Imhoff et al., 1995). Biomass is closely related to vertical vegetation structurally; to a first approximation, forests that are taller have more biomass than similar forests that are shorter in stature. In addition, the most accurate allometric equations are developed from field measurements of tree height and diameter. Therefore, a better approach to improve biomass estimates is by improving three dimensional forest structure measurements from remote sensing.

While this has already been achieved to some extent with lidar remote sensing, more progress needs to be made in terms of accuracies and spatial coverage of biomass.

1.2.2 Remote sensing for biodiversity studies

Biodiversity management and conservation is increasingly a cause of concern, with rapid rates of species decline and extinction. According to the Millennium Ecosystem Assessment (2005), the rates of species extinctions have increased 1000 fold in the last century alone and losses continue at an alarming rate of around 50,000 species each year. Understanding habitat requirements of individual species is critical for the design and implementation of effective conservation strategies.

Many studies have shown the influence of vertical arrangement of foliage within canopies on habitat and niche selection by wildlife (MacArthur & MacArthur 1961; Robinson & Holmes, 1984; and Degraaf et al., 1998). As an example, the California spotted owl prefers old growth forests with tall trees, dense canopy cover, and presence of dead trees /snags for nesting (Verner et al., 1992). Another species of recent notoriety is the Ivory-billed woodpecker. Although, there have been no confirmed sightings for the bird, extant habitat preferences are for tall trees, large basal area, dense canopy cover with open mid-story, abundant standing dead trees, and dying vegetation (Tanner, 1942). Yet another example is the black-throated blue warbler that has a preference for taller deciduous trees with a well developed understory. In each of these cases, habitat requirements are well understood but it is not clear how they can be characterized over large spatial scales for decision making. There are many other wildlife species for which the habitat requirements are not well

understood. In both these situations, comprehensive maps of forest attributes are invaluable.

Although passive remote sensing has already proved useful for habitat mapping by providing information such as habitat fragmentation, landscape structure, patch characteristics, canopy cover and phenology, they have limited capabilities for providing three dimensional forest structure. Therefore, the need to explore multi-sensor fusion approaches to address these requirements.

1.2.3 Lidar, Radar and Hyperspectral Remote Sensing

Lidar, radar and hyperspectral data provide vegetation attributes that are a potentially powerful combination for carbon, ecological and habitat studies. Lidar is an active remote sensing system that calculates distance by measuring the time taken by a laser pulses (infra red wavelengths for vegetation) to reach a target and return to the source. Airborne lidar instruments either record heights at intermittent levels or fully digitize the return signal to provide accurate measurements of sub-canopy topography, canopy height, and vertical distribution of canopy elements (Blair et al, 1999). Three dimensional structure metrics from lidar are more accurate in quantifying aboveground biomass than other remote sensing methods (Drake et al., 2002a; Lefsky et al., 2002). Metrics describing foliage distribution within the canopy have also proven to be useful in identifying suitable habitat conditions for many wildlife species, particularly birds (Goetz et al., 2007; Nelson et al., 2005; Martinuzzi et al., 2009). The main drawback of lidar however is its limited spatial and temporal coverage; hence the need for fusion with other sensors.

Radar (Radio Detection and Ranging) is also an active remote sensing system operating in the microwave region of the electromagnetic spectrum. Radar backscatter can be polarized in different ways to increase sensitivity to vertical or horizontal structural properties. Although radar is less sensitive to structure than lidar it has greater spatial coverage and is not affected by cloud cover, unlike lidar.

Hyperspectral data provide information about vegetation biochemical attributes such as chlorophyll content, canopy moisture (Ustin et al, 2004) complementary to lidar and radar. Because of the higher spectral resolution (typically 350 nm – 2500 nm), hyperspectral data can be more useful than multispectral data in discriminating species level differences in vegetation (Dennison and Roberts, 2003). Lidar, radar, and hyperspectral sensors observe vegetation through different mechanisms and at different wavelengths, and as such they can be combined to provide comprehensive forest structure and habitat characteristics (Table 1-1) relevant for carbon, ecosystem, and habitat studies.

Table 1-1 Forest structure and habitat characteristics directly measured or modeled from Lidar / Fusion (adapted from Bergen et al., 2009)

Variable	Sensor
Canopy Height	Lidar/Radar
Canopy Cover	Optical/ Waveform lidar
3D foliage profile and cover at various levels within canopy	Waveform lidar
Life form/vegetation type/species composition	Multispectral/Hyperspectral
Coarse Woody Debris, snags, standing dead	Hyperspectral + Lidar
Biochemical properties (moisture, chlorophyll, lignin-cellulose)	Hyperspectral
Basal Area	Lidar , Fusion
Biomass	Lidar, Radar, Fusion
Temporal variation in spatial characteristics Temporal changes in height and biomass	Optical/Radar Lidar
Topography	Lidar/Radar
LAI	Lidar/ Radar/Hyperspectral
Stem Density	Lidar
Patch/Edge characteristics	Lidar/Radar/Hyperspectral
Phenology	Multispectral/Hyperspectral

1.2.4 Multi-sensor fusion

The process of dealing with data from multiple sources to achieve refined/improved information for decision making has been termed as 'Data Fusion' (Hall, 1992). In theory, combining data from different sensors and databases improves accuracies and provides inferences better than from one dataset alone (Hall and Llinas, 1997). There is considerable interest in fusion applications with remote sensing data but this area of research is still in its infancy. Fusion approaches with remote sensing are broadly classified as: pixel level; feature level, and; decision based fusion (Zhang, 2010). My approach towards fusion in this dissertation includes decision based fusion, where images are processed individually to extract information and then combined using statistical and machine learning methods. I use several image processing methods to process images for atmospheric effects, geolocation shifts and extract useful information with classification techniques. Waveform processing is used to obtain physically meaningful metrics such as ground elevation, canopy height and canopy cover from lidar data as well as statistical metrics such as waveform energy quantiles. Features/attributes obtained from datasets individually are then integrated using statistical and machine learning methods (described in detail in Chapter 2 and Chapter 4) to predict variables of interest such as aboveground biomass or habitat quality. Additionally, I use a range of data querying, data visualization and analytical methods to compare lidar and hyperspectral metrics to detect variations in structure and biochemical conditions of vegetation. While these approaches by themselves are not new methodological innovations, the ensemble of

techniques used here in combination with new datasets provides fresh insights into multi-sensor fusion.

1.3 Dissertation Outline

This dissertation is subdivided into five chapters. In Chapter 2, I integrate lidar and hyperspectral data for mapping aboveground biomass with an aim to understand whether fusion can be applied to reduce uncertainties in carbon storage. In addition, I combine the two sensors to detect 'stress' in high biomass forests as an indicator of canopy loss or mortality. In Chapter 3, I explore the use of temporal lidar data to map changes in height and biomass spatially with a larger goal of identifying potential carbon sources and sinks. While this chapter does not include fusion directly, I make semi-quantitative comparisons between lidar changes and optical imagery/ hyperspectral stress maps from my previous results. In Chapter 3 my goal is to explore the potential of multi-sensor fusion with radar, lidar and multispectral imagery for mapping bird habitat characteristics (Fig 1-1).

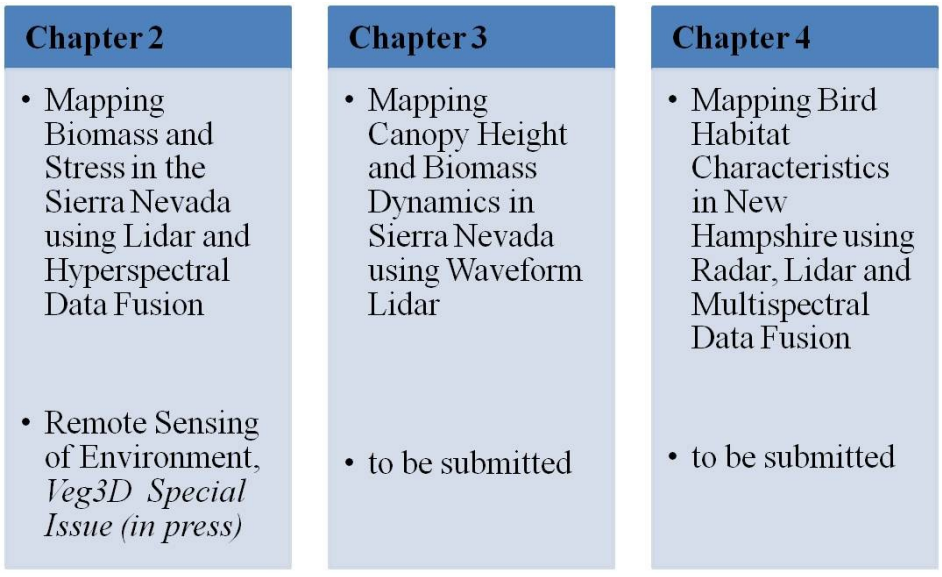

Fig. 1-2 Outline of dissertation chapters.

Chapter 1: Mapping biomass and stress with lidar and hyperspectral data

In this chapter, I draw from extensive studies that have derived canopy biophysical and chemical properties from hyperspectral remote sensing and validated the use of lidar in mapping forest structural characteristics. The primary objective here is to test fusion approaches between these two sensors for improving quantitative and spatial estimates of aboveground biomass. In the first approach, I combine a suite of spectral metrics from hyperspectral data with canopy height and cover metrics from lidar for biomass estimation. This would detect the contributions of canopy biochemical characteristics to structure from lidar in estimating biomass. The second approach is to test whether stratifying by species with hyperspectral data improves biomass estimates from lidar. The hypothesis is that species stratification would improve results because field biomass is derived from species specific allometry. Thirdly, since there is clear but unexplored synergy between these datasets for mapping canopy condition in addition to structure, I combine canopy height, biomass, canopy cover with indices of moisture, chlorophyll content to detect relatively high and low stress areas in forests with similar structural characteristics as an indicator of possible canopy losses and mortality.

Chapter 2: Mapping Canopy Height and Biomass Dynamics with Temporal Lidar Data

In addition to quantifying carbon storage, there is also a need to map changes as a result of natural and anthropogenic disturbances to reduce the larger uncertainties in carbon flux estimates. This is particularly important in temperate forests, where

aboveground biomass is fairly well quantified but changes in carbon sources and sinks from disturbance are not understood. Most research with lidar has been static; studies are only beginning to explore their use in canopy dynamics and carbon flux estimation. I build upon previous work with small (Kellner et al., 2009) and medium footprint lidar (Dubayah et al., 2010) for mapping potential carbon sources and sinks in the temperate montane forests of the Sierra Nevada. Going back to the results on canopy stress from Chapter 2, I test whether hyperspectral data can be used to identify areas with higher stress and that are thus likely to undergo canopy losses.

Chapter 3: Mapping Bird Habitat Characteristics Using Multi-sensor Fusion

This chapter brings together the lessons learned in previous chapters for an application in habitat mapping. The initial research plan was to integrate forest structure, composition and canopy dynamics in the Sierra Nevada from Chapters 2 and 3 to map habitat for the California Spotted Owl. However, this could not be done due to lack of bird data. I therefore shifted my focus to the Hubbard Brook Experimental Forest in New Hampshire, where bird species have been extensively studied and monitored since the 1960's. Additionally, recent mapping of the area with small and medium footprint lidar, radar and multispectral data provided an excellent opportunity to test fusion ideas in habitat mapping. In this study, I test the individual and combined capabilities of multiple polarized backscatter from radar, phenology from multispectral data, broader scale vertical structure from waveform lidar and individual tree mapping from small footprint (0.5m resolution) lidar for mapping habitat characteristics. My goal is to explore what variables can be derived from these

sensors, how they can be combined with machine learning methods for predicting bird habitat quality, and what are the combined accuracies of predictions.

1.4 Summary

The broad arc of my research, from hyperspectral/lidar fusion, to multi-date lidar, to radar/lidar and multispectral fusion illustrates both the enormous potential of fusion for habitat and carbon studies, as well as the tremendous amount of work required to translate ideas of what we think should work to actual models of applied value. As space borne lidar, hyperspectral and radar data become available over inaccessible forests across the globe, multi-sensor fusion will become more important and exceptionally useful in optimizing information for decision making. By exploring both the promise and limits of the next generation of sensors, I hope to lay the groundwork for future work that brings us from notional concepts on forest characterization to demonstrated efficacies that can be used to help us manage the environment in a changing world.

Chapter 2 Mapping Biomass and Stress in the Sierra Nevada using Lidar and Hyperspectral Data Fusion

2.1 Introduction

Improved estimates of forest aboveground biomass, hereafter "biomass" from remote sensing are critical for reducing uncertainties in the global carbon cycle (Rosenqvist et al., 2003; Hese et al., 2005) and are an important goal for future satellite missions. Although coarse-scale biomass estimates are well documented in temperate forests, they are mostly in the form of field measurements and averages over administrative units (Houghton, 2005). There is a need for higher resolution and spatially continuous estimates to quantify carbon flux and disturbance at scales at which land use activities occur (Houghton, 2005; Keith et al., 2009). Spatial distribution of carbon stocks in combination with species composition and vegetation stress can improve the understanding of ecosystem processes (Ustin et al., 2004; Chambers et al., 2007), carbon dynamics, and habitat structure (Bergen et al., 2007). The availability of such maps over difficult mountain terrain such as the Sierra Nevada can be particularly valuable for natural resource and wildlife habitat management.

Many studies have demonstrated the efficacy of waveform lidar in accurately measuring three-dimensional vegetation characteristics including biomass for different forest cover and types (Lefsky, 2002; Drake et al., 2002a). Lidar metrics are less prone to saturation effects even at high biomass levels (Lefsky, 2002; Drake et

al., 2002a; Hyde et al., 2007a) unlike most remote sensing indices, which saturate at moderate values (Gao, 1996; Huete, 1997). Acquiring wall-to-wall coverage of airborne lidar however, is expensive. A promising alternative is to extrapolate forest structure from lidar samples using continuous remotely sensed data. There is considerable interest in fusing sparse but accurate lidar measurements with optical (Hudak et al., 2002; Kimes et al., 2006; Hyde et al., 2007a; Asner et al., 2008) and radar sensors (Treuhaft et al., 2004) to improve prediction accuracy and spatio-temporal coverage of forest structure.

Imaging spectrometers or hyperspectral sensors provide many attributes complementary to canopy structure from lidar and can be used to discriminate vegetation types based on spectral characteristics.(e.g. Martin et al., 1998; Dennison and Roberts 2003; Clark et al., 2005). Studies have suggested that spectral attributes (Ustin et al., 2004; Bergen et al., 2006) and species composition (Rosenqvist, et al., 2003; Anderson et al., 2005) from hyperspectral data could improve biomass estimates in conjunction with lidar. However, it is still unclear as to how biophysical and biochemical attributes from hyperspectral data relate with structural attributes from lidar. There also remains considerable uncertainty on the efficacy of combining lidar with hyperspectral sensors for species-specific biomass mapping. Underlying causes of biomass change such as physiological stress, tree mortality and senescence cannot be detected from lidar alone, as it does not differentiate between healthy and stressed vegetation (Rosenqvist et al., 2003). While the ability of hyperspectral data to map stress is recognized (Roberts et al., 1997; Asner, 1998; Merton, 1998), the

combined use of the two sensors for mapping vertical structure and stress remains largely unexplored.

Our goal in this study was to explore fusion of waveform lidar from the Laser Vegetation Imaging Sensor (LVIS) with hyperspectral imagery from the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) for mapping biomass and stress in the diverse montane forests of the Sierra Nevada. In particular, we evaluated whether addition of spectral metrics from AVIRIS improved biomass estimates from LVIS. We also assessed whether species stratification using AVIRIS data prior to lidar estimation of biomass increased accuracy. Lastly, we explored the combined potential of the two sensors for mapping stress in the high biomass forests of the Sierra Nevada.

2.2 Background

Lidar and hyperspectral remote sensing are two potentially complementary technologies capable of providing comprehensive structural and biophysical characteristics of vegetation (Koetz et al., 2007). Lidar instruments record the time taken by laser pulses to reach the earth's surface from an aircraft/satellite and back to calculate distance to target. Discrete return lidar devices provide one or more laser returns that can be used for high resolution mapping of terrain and canopy elevation (Lefsky et al., 2002). Waveform lidar instruments digitize the entire outgoing and return signal to provide waveforms, from which various parameters such as subcanopy topography, canopy height, foliage profiles and vertical heterogeneity may be derived (Blair et al., 1999; Dubayah, 2000). Waveform metrics from small and

large footprint lidar have been used to predict biomass in tropical (Drake et al., 2002b; Clark et al., 2004) and temperate forests (Lefsky et al. 2002; Hyde et al. 2005; Anderson et al., 2005).

Hyperspectral sensors measure vegetation absorption and scattering characteristics in the visible, near infrared and short wave infrared wavelengths of the electromagnetic spectrum. Spectral indices or band ratios from hyperspectral data provide many attributes useful for ecological studies (Ustin et al., 2004) such as chlorophyll content (Elvidge and Chen, 1995), canopy water status (Gao, 1996; Serrano et al., 2000), vegetation stress (Merton, 1998) and lignin and cellulose content (Kokaly and Clark, 1999; Curran et al., 2001). Narrow band and derivative-based indices from hyperspectral data are relatively less affected by background soil reflectance (Elvidge and Chen, 1995), illumination, saturation (Gao, 1996; Pu et al., 2003, Roberts et al., 2004), and other factors that influence broadband vegetation indices such as the normalized difference vegetation index (NDVI). Measures of liquid water (e.g. equivalent water thickness, EWT) from hyperspectral data are highly sensitive to canopy properties such leaf area index (LAI) (Roberts et al., 2004). Measures of plant dry matter have been related to environmental stress (Asner, 1998) and could improve lidar estimates of biomass in areas with low canopy heights and sparse vegetation cover (Ustin et al., 2004; Treuhaft et al., 2004; Bergen et al., 2006).

Spectral Mixture Analysis (SMA) is a widely used remote sensing technique for obtaining ecologically relevant and meaningful components from an image pixel (Adams et al., 1986; Chambers et al., 2007). In SMA, two or more reference spectra/endmembers such as green vegetation, soil and shade are modeled as linear

combinations to estimate sub-pixel fractions of each component. A limitation of SMA is that it uses only one set of reference endmembers to model all pixels in an image. Multiple endmember spectral mixture analysis (MESMA) (Roberts et al., 1998) allows the number and type of reference endmembers to vary on a per-pixel basis, accounting for spectral variability in the landscape and improving the accuracy of resulting fractions. Because MESMA fractions are calculated using the entire spectrum, they are more robust than traditional vegetation indices and have successfully been used for estimating live fuel moisture (Roberts et al, 2006), LAI (Sonnetag et al., 2007) and green biomass in pastures (Numata et al., 2008). MESMA has also been used to map vegetation (Dennison and Roberts, 2003) and urban landcover (Franke et al., 2009).

Most studies on lidar and hyperspectral fusion have focused on land cover classification. Asner et al. (2008) used lidar to mask gaps and low canopy heights, improving detection of invasive species from AVIRIS for Hawaiian rainforests. Koetz et al. (2007), classified fuel composition from fused lidar and hyperspectral bands using Support Vector Machines (SVM). Classification accuracies from fusion were higher than from either sensor alone. Mundt et al. (2006) fused co-registered lidar and hyperspectral data to map sagebrush communities and suggested further use of classified vegetation maps in biomass calculations. Few studies have explored the combined potential of the two sensors for biomass estimation. Anderson et al. (2008) used Minimum Noise Fraction (MNF) transformed AVIRIS bands in combination with LVIS and reported an 8 – 10 % increase in biomass prediction accuracy for

northeastern temperate forests. There is a need to test similar approaches over a wider range of forest cover and types, while retaining the physical significance of variables.

2.3 Study Area and Data

2.3.1 Study Area

The study site (37°2'34.47"N, 119°9'33.81"W) covers an area of around 22,000 ha and lies along the western slopes of the Sierra National Forest (Fig. 2-1), in California, USA. The region has a Mediterranean climate with elevations ranging from 1000 m to 2500 m. Forests are dense and complex in structure with average biomass values of around 200 Mg/ha, and as high as 1000 Mg/ha in Giant Sequoia (*Sequoiadendron giganteum*) stands. Dominant species include red fir *(Abies magnifica)*, white fir *(Abies concolor*), ponderosa pine *(Pinus ponderosa),* and California black oak *(Quercus kellogi)* (Hunsaker et al., 2002).

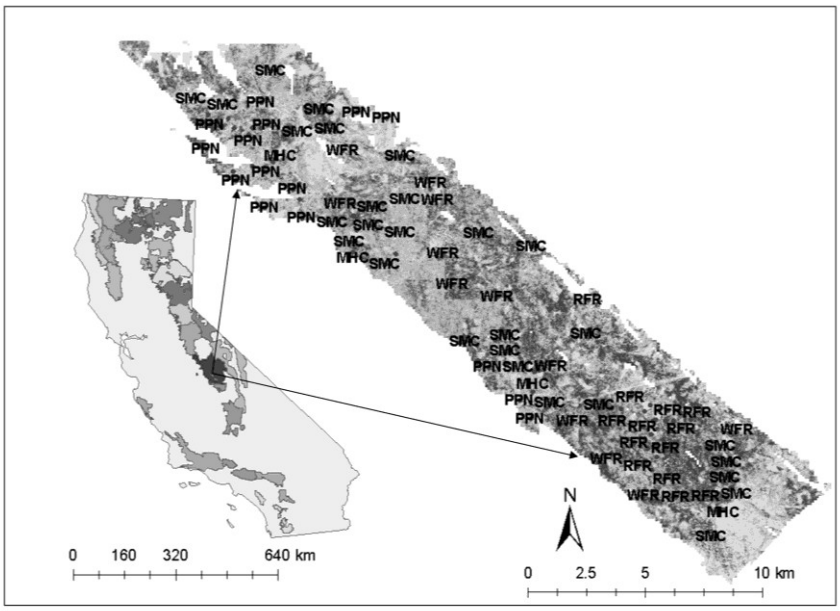

Fig. 2-1 Study area in the Sierra National Forest showing 1ha field plots. Plots are classified based on Wildlife Habitat Relation (WHR) type

2.3.2 Data

Field data

Field surveys were conducted in 2000 and 2001 (Hyde et al., 2005). A modified stratified random sampling scheme was used to measure structural parameters over 500 plots in the northern and southern Sierra Nevada. Field plots were laid out to provide a statistically representative measure of structural variability for the eight major Wildlife Habitat Relation (WHR) types: montane hardwood, montane hardwood conifer, red fir, white fir, sierran mixed conifer, pines, wet meadow, and barren (Mayer and Laudenslayer, 1988, Hyde et al., 2005). Structural variables for live trees such as height, diameter at breast height (DBH), crown form, canopy cover, species, heights of dead snags and snag decay classes were recorded for concentric plots with radii of 15 m (0.07 ha - footprint level) and 56.4 m (1 ha - stand level) respectively. For 1 ha plots, only large trees with dbh greater than 76 cm were measured. A detailed description of the methods used for field data acquisition is available in Pierce et al. (2002). The study area in the Sierra National forest had 285 measured plots out of which 125 1 ha plots had collocated lidar, hyperspectral and field measurements and were used for analysis (Table 2-1). The 0.07 ha plots were not used in this study because of increased geolocation errors between reprocessed lidar and field data.

Table 2-1 Distribution of field plots by Wildlife Habitat Relation (WHR) Type

Species/WHR type	Number of Plots 1 ha (LVIS, AVIRIS, field)
WTM (wet meadow)	8
BAR (barren)	3
RFR (red fir)	32
WFR (white fir)	27
SMC (Sierran mixed conifer)	37
PPN/JPI (pines)	18
GSQ (Giant Sequoia)	1
MHC/MHW	11
Total	137

Lidar Data

The Laser Vegetation Imaging Sensor (LVIS) is a medium footprint, waveform digitizing, scanning laser altimeter, designed, and developed at NASA's Goddard Space Flight Center. LVIS operates at altitudes up to 10 km with a 7 º field of view and uses laser pulses with a wavelength of 1064 nm for profiling vertical vegetation structure (Blair et al., 1999). NASA flew LVIS over the Sierra National Forest in summer 1999 at an altitude of 7km with trees in leaf-on condition. The lidar shots had a nominal footprint radius of 12.5m. The data had a swath width of 1km and covered an area of 175 sq. km. The subset used for this study had around 892,444 lidar footprints. Footprints were contiguous along track and overlapping across track. (See http://lvis.gsfc.nasa.gov/)

Hyperspectral Data

The Airborne Visible Infrared Imaging Spectrometer (AVIRIS) designed and developed at the Jet Propulsion Laboratory, measures upwelling radiance from the earth's surface in wavelengths between 350 nm – 2500 nm in 224 contiguous bands with a bandwidth of 10 nm (Green et al., 1998). Fine spatial resolution AVIRIS images were acquired over the Sierra Nevada in July 2003. Radiometrically corrected images were processed to retrieve apparent surface reflectance using the MODTRAN based forward inversion approach as described in Green et al. (1993) and Roberts et al. (1997). The images were geometrically corrected using Digital photo Ortho Quads (DOQQ). The AVIRIS data consisted of three overlapping scenes covering a total area of 22,000 ha. Each image had a nominal spatial resolution of 3.3 m with 224 spectral bands. Bands with a poor signal-to noise ratio from atmospheric interference of water vapor and carbon dioxide were eliminated, resulting in 118 bands for analysis.

2.4 Methods

The data sets used in this study had different geographical projections and were brought into a common frame of reference using the Universal Transverse Mercator Projection (UTM 19N) and NAD 1983 datum. Spatial overlay of AVIRIS and LVIS data showed good geolocation for analysis at the 1 ha level and no further rectification was performed. All hyperspectral processing was done using VIPER tools ENVI- Add on Module© (Roberts et al., 2007) and a hyperspectral metrics add on module.

2.4.1 Field attributes

Species-specific allometric equations from the USDA Forest Service (Waddell and Hiserote, 2003) were used to calculate biomass for all trees with dbh greater than 76cm within 1 ha plots. Tree height, species, and dbh measurements from field data were used as inputs for calculating biomass of bole, bark, branches, and foliage separately for each tree. Biomass values for individual trees were then added to obtain aboveground biomass per hectare for large trees within each field plot. Field plots were classified based on WHR type for analysis by species/vegetation type. Although WHR types consist of species associations, most of them have a dominant plant genus/species. The MHC/MHW plots consisted of mixed hardwoods and conifers with broadleaf oaks (Quercus sp.) as the dominant vegetation type. Plots classified as PPN were mostly composed of ponderosa pine (Pinus ponderosa). SMC plots had mixed conifers including pines (Pinus sp.), firs (Abies sp.), and incense cedar (Libocedrus decurrens) with shade tolerant white fir (Abies concolor) becoming increasingly dominant after regeneration from fire (Zald et al., 2008). RFR plots were almost entirely composed of red fir (Abies magnifica) trees.

2.4.2 LVIS metrics

An LVIS waveform essentially consists of a signal with amplitudes proportional to energy reflected from intercepted surfaces within canopy and ground. LVIS footprints are geo-located to the global reference ellipsoid WGS 84, using a combination of GPS and Inertial Navigation System (INS) information (Blair et al., 1999; Hofton and Blair, 2002). Ground elevation is determined by identifying the center of the lowest mode in the waveform greater than mean signal noise (Fig. 2-2).

Canopy elevation is the height at which the signal increases beyond a certain threshold (usually 3σ of the background noise) at the top of the waveform (Hofton and Blair, 2002). The difference between canopy elevation and ground elevation gives the canopy height metric or height of 100% laser energy return (RH100). The 1999 LVIS data were reprocessed using algorithms for ground detection and an improved horizontal geolocation algorithm prior to the start of this analysis (Blair et al., 2006). For each LVIS waveform, quartile heights of laser energy return i.e. height of 25% (RH25), 50% (RH50) and 75% (RH75) energy return were calculated in addition to RH100 (Fig. 2-2). Canopy cover was calculated from the ground energy return of each waveform normalized by the canopy and background reflectivity ratio. We used a ratio of 1.6, derived from a previous study (Hyde et al., 2005). LVIS metrics were calculated for lidar shots within 1 ha plots and summarized to obtain minimum, maximum, mean and standard deviation for all metrics.

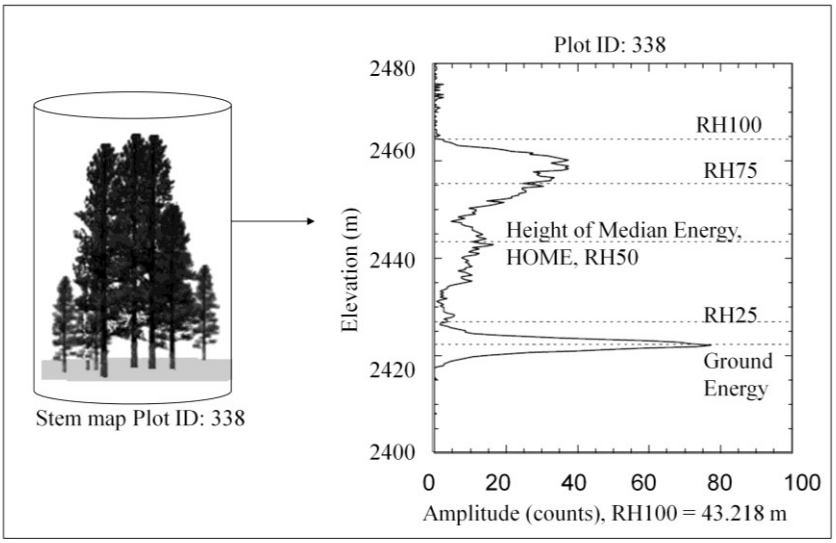

Fig. 2-2 Example of an LVIS waveform centered on a field plot with area 0.0 7 ha. The amplitude of the waveform is proportional to energy reflected from canopy and ground. Metrics calculated from the waveform include ground elevation, quartile heights of energy return (RH25, RH50, RH75 & RH100) and canopy cover.

2.4.3 AVIRIS spectral metrics

Reflectance spectra (e.g. Fig. 2-3) were extracted from AVIRIS images over field plots to construct a spectral library for the study area. A set of 19 hyperspectral indices (Table 2-2) were calculated to quantify vegetation attributes from each spectrum and aggregated to obtain mean and standard deviation of values for 1 ha plots. These mainly included vegetation indices, derivatives of the chlorophyll red edge, water band ratios, and ligno-cellulose band ratios. Indices based on the green, red, and blue wavelengths were found to be more robust than NDVI in estimating vegetation fractions (Gitelson et al., 2002). We tested the use of green band vegetation indices (VARIGREEN, VIGREEN) in addition to NDVI, Enhanced Vegetation Index (EVI), and Ratio Vegetation Index (RVI) for biomass prediction. Water absorption features in the infrared regions of the spectrum (e.g. 980nm 1450nm, 1940nm) are sensitive to canopy biophysical properties (Serrano et al., 2000, Roberts et al., 2004). The Normalized Difference Water Index (NDWI), Equivalent Water Thickness (EWT), and Ratio Water Index (RWI) were used as measures of canopy water content. The red edge or the rapid change in chlorophyll reflectance in the visible and near infrared portion of the spectrum provides a measure of chlorophyll content (Elvidge and Chen, 1995) and vegetation stress (Merton, 1998). We used the first and second derivatives of the red edge as measures of chlorophyll content. Wavelength, asymmetry and area of lignin and cellulose absorption features (2045nm -2218nm) (Kokaly and Clark, 1999; Curran et al., 2001) were used to identify non-photosynthetic vegetation; their usefulness in improving biomass estimates in combination with LVIS data was also tested.

Table 2-2 Hyperspectral metrics calculated using AVIRIS data for 1ha field plots included vegetation indices, red edge derivatives, Ligno-cellulose band ratios and MESMA fractions of green vegetation, soil/NPV and shade.

Variable	Hyperspectral metric (mean , std. deviation)	Reference
IDL_DGVI	First Derivative of Red Edge normalized to 626-795 baseline	(Elvidge & Chen, 1995)
1DZ_DGVI	First Derivative of Red Edge normalized to 626 baseline	(Elvidge & Chen, 1995)
2DZ_DGVI	Second derivative of Red Edge normalized to 626 baseline	(Elvidge & Chen, 1995)
EVI	Enhanced Vegetation Index	(Heute, 1997)
EWT	Equivalent Water Thickness	(Roberts, et al)
LC2 BAND RATIOS	Area, Asymmetry, Depth, Width, Wavelength of SWIR ligno-cellulose absorption feature (range 2045-2218 nm)	(Kokaly & Clark, 1999)
NDVI	Normalized Difference Vegetation Index NDVI	(Tucker, 1979)
NDWI	Normalized Difference Water Band Index	(Gao,1996)
RWI	Ratio Water Band Index	(Penuelas et al.,1993)
RVI	Ratio Vegetation Index	(Jordan, 1969)
RVSI	Red Edge Vegetation Stress Index	(Merton,1998)
REDEDGEWAVE	Red Edge Wavelength	(Pu et al., 2003)
REDEDGEMAG	Magnitude of Red Edge	(Pu et al., 2003)
VARIGREEN	Visible Atmospherically Resistant Vegetation Index Green	(Gitelson,2002)
VIGREEN	Vegetation Index Green	(Gitelson,2002)
SHADE FRACTION	fraction of vegetation shade within each pixel (calculated using MESMA)	MESMA, Roberts et al,1998
GV FRACTION	fraction of green vegetation within each pixel (calculated using MESMA)	MESMA, Roberts et al,1998
SOIL+ NPV FRACTION	fraction of soil + non photosynthetic vegetation within each pixel (calculated using MESMA)	MESMA, Roberts et al,1998

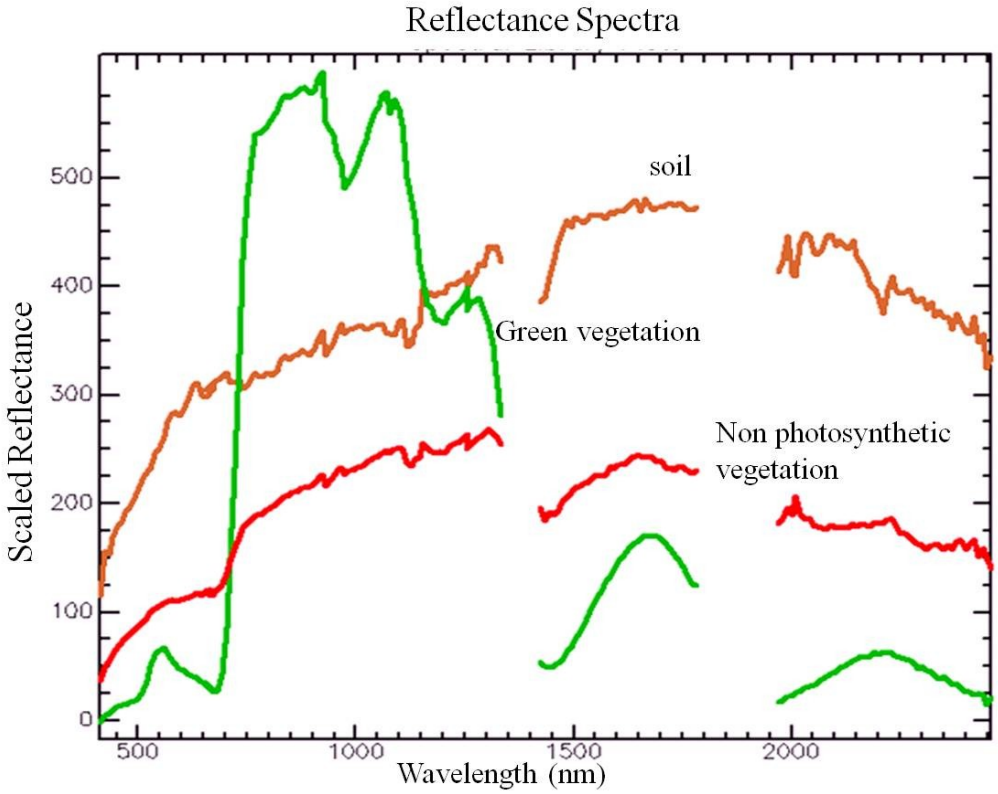

Fig. 2-3 Examples of reflectance spectra extracted from AVIRIS images over the study area. A set of 19 band ratios describing vegetation characteristics such as chlorophyll content, water content, stress were calculated from the visible, near infrared and short wave infrared wavelengths

2.4.4 MESMA fractions from AVIRIS

The reflectance of an image pixel over a forested area is typically composed of varying combinations of bare soil, shade/shadows, green vegetation (GV) from foliage and non-photosynthetic vegetation (NPV) from dead bark, leaf litter or senescent vegetation (Roberts et al., 2004; Chambers et al., 2007). Multiple endmember spectral mixture analysis (MESMA) involves creation of regionally specific libraries by using reference spectra from an image, field, or modeled spectra. Each spectrum in the reference spectral library is modeled as a combination of another spectrum and shade (Dennison et al., 2004).

Three fit metrics are used to identify representative spectra or endmembers for each class: Count Based endmember (COB) (Roberts et al., 2003; Franke et al., 2009), Endmember Average Root mean square error (EAR) and Minimum Average Spectral Angle (MASA). COB values are used to select endmembers that model spectra within the same class (In COB) better than those in other classes (Out COB). EAR values are used to select spectra with lowest root mean square error in modeling other spectra of the same class (Dennison and Roberts, 2003). Spectra with low average spectral angle values (MASA) are selected as reference endmembers. A detailed description of MESMA and fit metrics can be found in Dennison et al., (2004).

A library was created for the Sierra Nevada from AVIRIS images by extracting reference spectra for grass, shrubs, trees, soil, and NPV using field data and image interpretation. Each spectrum in the library was modeled as a combination of another spectrum and shade. We selected endmembers with high In COB values followed by those with low MASA and EAR values (Table 2-3). Several models with varying combinations of endmembers were tested using SMA/MESMA. For this study, we used 10 three-endmember (soil, green vegetation, and shade) models for unmixing the AVIRIS images. Soil and NPV endmembers were combined into one class. The resulting image consisted of fractional abundances of green vegetation, soil/NPV, and shade for each pixel at 3.3m nominal spatial resolution (Fig. 2-4). MESMA fractions were then summarized to calculate mean and standard deviation of values for 1 ha field plots.

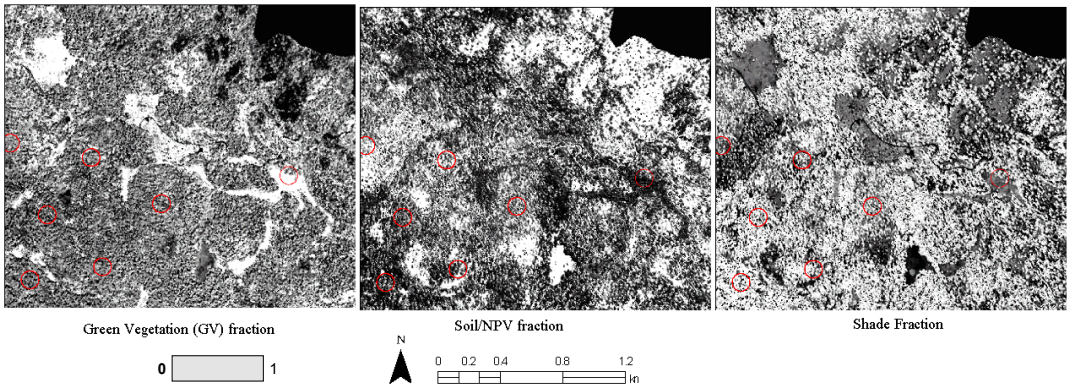

Fig. 2-4 Subset of images showing endmember fractions generated using MESMA. Bright areas have high fractional abundance and dark areas have low abundance. GV, soil/NPV and shade fractions were summarized to calculate mean and std. deviation of values for 1ha plots.

Table 2-3 Reference endmembers used in 10 3 endmember MESMA models for unmixing AVIRIS images. Soil and non photosynthetic vegetation (NPV) spectra were grouped into one class. Fit metrics - EAR, MASA and COB values were used to select the best representative spectra to un-mix the entire image.

Endmember	Class	Brightness	EAR	MASA	In_CoB	Out_CoB	CoBI
Rock/Soil	Soil/NPV	421.53	0.13	0.66	10	0	0
Soil	Soil/NPV	345.05	0.04	0.095	0	0	0
NPV	Soil/NPV	254.47	0.04	0.13	0	0	0
NPV	Soil/NPV	260.65	0.04	0.133	6	2	0.08
NPV	Soil/NPV	203.46	0.08	0.24	17	1	0.46
Green Vegetation	GV	188.92	0.09	0.39	0	0	0
Green Vegetation	GV	210.73	0.09	0.41	0	0	0

2.4.5 Land-cover classification from AVIRIS

We also used MESMA to classify landcover and dominant vegetation type from AVIRIS images. A spectral library was constructed from the AVIRIS images using field knowledge, coarse vegetation type maps (USDA Forest Service CALVEG data, 2007), lidar height maps, and image interpretation. We isolated patches of vegetation with dense canopy cover and extracted relatively pure spectra for oaks (Quercus sp.), white fir (Abies concolor), red fir (Abies magnifica), mixed firs (Abies sp.), and pines (Pinus sp.). Field knowledge was used to avoid plots with abundant ground cover of chaparral (Ceonothus sp. and Arctostaphylos sp.) and reduce mixing with canopy dominant spectra. Spectral metrics from AVIRIS such as NDVI, NDWI, EWT, and lignocellulose band ratios were also useful for separating non-photosynthetic vegetation, bare soil, and spectra for dominant vegetation types.

Each spectrum in the library of 183 spectra was unmixed with another spectrum and shade resulting in 182 unique two endmember models for each spectrum. Fit metrics EAR, MASA, and COB (see sec. 2.4.4) were used to select suitable reference endmembers for landcover classification. We selected 47 spectra from several classes including pines, hardwoods, grass, soil, NPV, and chaparral. All AVIRIS images were unmixed using 47 two-endmember MESMA models to map landcover/ dominant vegetation type (Fig. 2-5a). Outputs included dominant landcover type in each pixel and the corresponding fractional abundance. Pixels mapped as soil, rock, NPV, chaparral, and grass in the AVIRIS vegetation map were excluded. A vector grid of 1ha polygons was placed over the species map and class statistics were calculated for pixels with pines, firs (red and white), and hardwoods

within each polygon. The dominant class in each polygon was recorded to create an aggregated 1 ha species map (Fig. 2-5b).

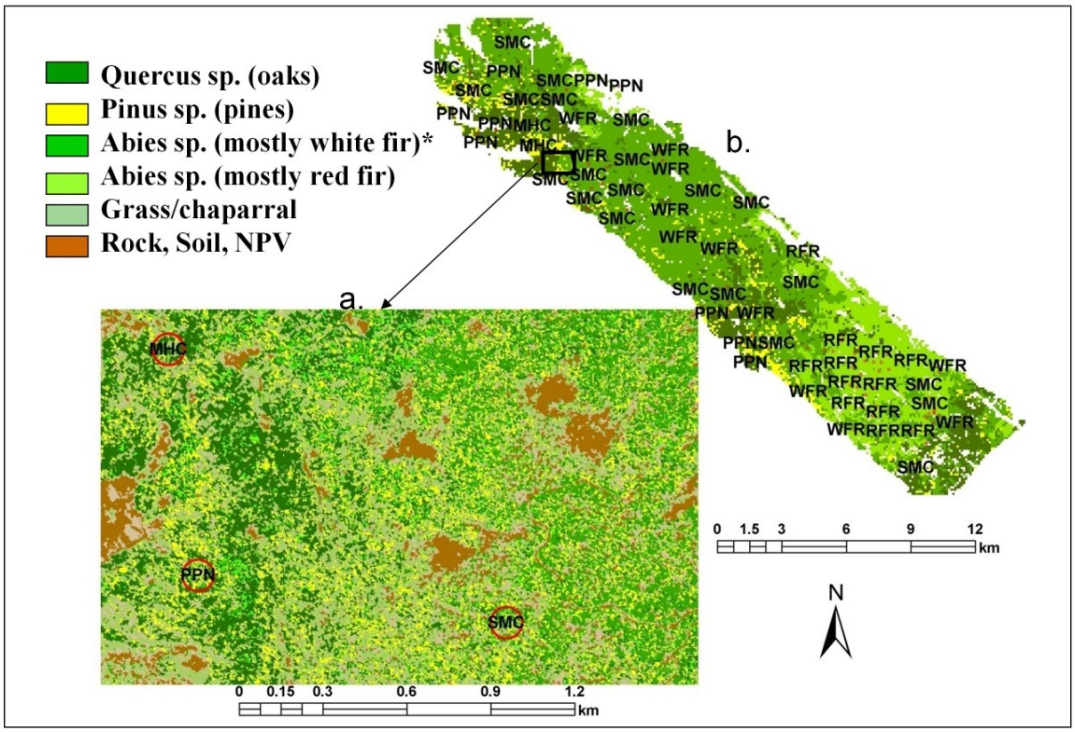

Fig. 2-5 MESMA was used to map landcover and dominant vegetation types from AVIRIS images. AVIRIS maps at a resolution of 3.3m (a) were aggregated and dominant vegetation type at 1ha was identified (b). Labels show WHR types from field plots. *Forests classified with white fir as dominant vegetation type at 1ha also had a mixture of conifers and were grouped as the SMC type for biomass estimation.

2.5 Analysis

2.5.1 Stand Level

The final data for stand level (1 ha) analysis included field-measured biomass, LVIS metrics, spectral indices, and MESMA fractions from AVIRIS. Wet Meadow (WTM) and barren (BAR) plots were excluded and 125 other plots with collocated LVIS, AVIRIS, and field data were used. We tested several linear and multiple

stepwise regression models to predict biomass before and after species stratification. Models for predicting total and species-specific biomass were tested using AVIRIS metrics alone, LVIS variables alone and a combination of AVIRIS and LVIS metrics. Three parameters were used to select the best models; high co-efficient of determination values (r^2), low Root Mean Squared Error (RMSE), and p value < 0.05 for r-squared as well as predictor variables. The Akaike information criteria (AIC) were used to select suitable predictor variables for all models. We also tested variable selection using Bayesian model averaging (BMA). Confidence intervals for co-efficients of determination were calculated to compare the statistical significance of different models.

2.5.2 Landscape Level

Equations from stand level analysis were used to generate landscape maps to study variations between biomass from LVIS alone and species-specific biomass from fusion. The vegetation map at 1 ha had four classes: hardwoods, pines, white fir/mixed conifers, red fir. White fir (Abies concolor) and red fir (Abies magnifica), were grouped into one class at lower elevations (<2000m) because of mixing. At the hectare level, mixing of species/genera within each class was unavoidable, but our vegetation classes closely matched the WHR type classification from the USDA Forest Service map at a coarser resolution. For example, polygons grouped as hardwoods were similar in distribution to the MHC/MHW type; pines were similar to the PPN type. Polygons with white fir as the dominant species also had a mix of pines, red fir and some hardwoods. These polygons were grouped under the

firs/mixed conifer type. Biomass was calculated for each dominant vegetation type using equations derived from lidar and field data (Table 2-5).

Spatial patterns of AVIRIS vegetation indices, water band indices and red edge derivatives were analyzed in combination with lidar heights, canopy cover, and biomass predicted from fusion. We used the NDWI as an indicator of water content (Gao et al., 1996; Serrano et al., 2000; Maki et al., 2004) and the normalized first derivative of red edge, D1GVI as an indicator of chlorophyll content (Merton, 1998; Smith, 2004). The species biomass map from fusion was combined with NDWI and D1GVI maps to detect 1 ha stands with biomass greater than 200Mg/ha, canopy cover greater than 40%, NDWI less than 0.05 and D1GVI less than 0.1. Stands with biomass greater than 200 Mg/ha, canopy cover greater than 40 %, NDWI greater than 0.05 and D1GVI greater than 0.1 were identified as areas with relatively low stress. To rule out effects of soil reflectance on water band indices (Gao, 1996) we further analyzed NPV fractions within canopies alone by masking other landcover types.

2.6 Results

2.6.1 Stand Level

AVIRIS variables explained around 60% of the variability in biomass (r^2= 0.60 RMSE = 92.13 Mg/ha) with water band indices being the most important variables (Fig. 2-6a). LVIS height metrics were found to be consistently better predictors of total and species specific biomass. The best model for stand level prediction had an r^2=0.77, RMSE = 70.12 Mg/ha, with RH75 being the single best

predictor (Fig. 2-6b). AVIRIS metrics showed marginal improvement in biomass prediction (but not statistically significant) when combined with LVIS metrics for 1 ha plots (r^2= 0.80, RMSE = 64.18 Mg/ha) (Table 2-4, Fig. 2-6c). AVIRIS variables including water band ratios (RWI, NDWI, EWT) and shade fractions from MESMA showed strong correlation with LVIS heights (r^2= 0.69, RMSE = 5.2 m). Mean and standard deviation of shade fractions alone explained more than 50% variability in all LVIS metrics (for example, r^2= 0.54, RMSE = 6.25 m for RH100).

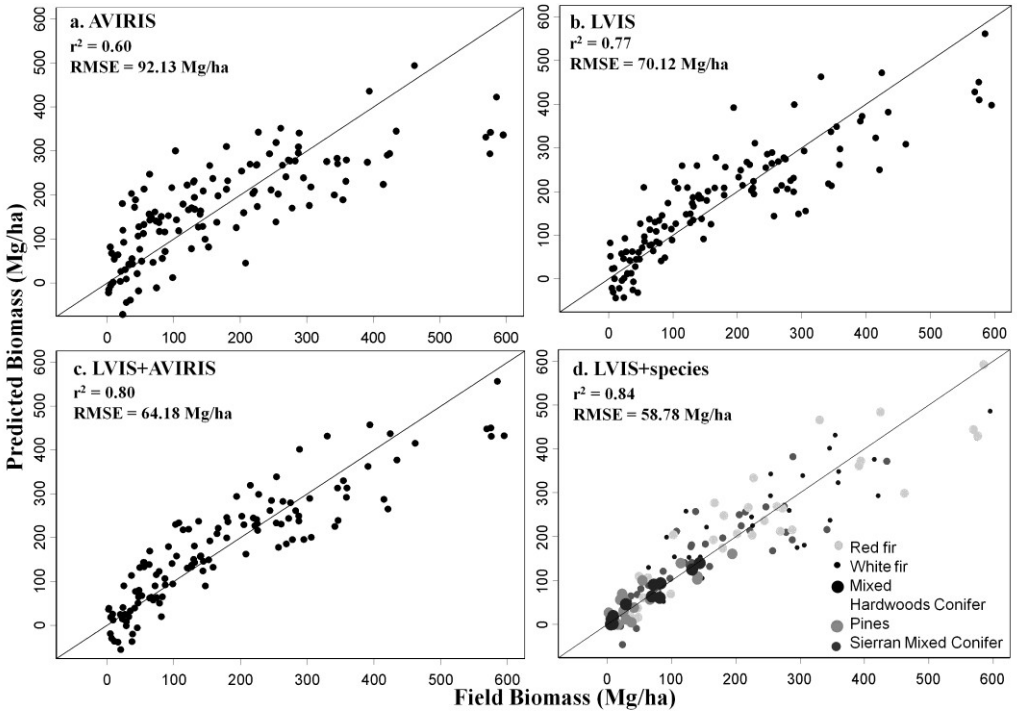

Fig. 2-6 Biomass predicted for 125 field plots at 1ha scale using various metrics: (a) AVIRIS; (b) LVIS; (c) LVIS and AVIRIS metrics; (d) LVIS after species stratification of field data.

Species-specific biomass relationships were analyzed for 125 plots (excluding barren and meadow plots). Classification of field plots by vegetation type/species before biomass estimation from LVIS improved prediction accuracy (r^2 = 0.84, RMSE = 58.78 Mg/ha) (Fig.2-6d). The maximum increase in predicted values were

observed for MHC/MHW plots with hardwoods as dominant vegetation (r^2 = 0.94, RMSE = 12.7 Mg/ha). For other dominant vegetation types, there was little change or even a slight increase in RMSE with species stratification (Table 2-5). RH75 was again the single best predictor of biomass for almost all vegetation types. AVIRIS metrics showed strong correlation with biomass for pines and hardwoods (r^2 greater than 0.7). Relationship between AVIRIS metrics and field biomass decreased considerably (r^2 less than 0.45) in high biomass plots of red fir (Abies magnifica) and mixed conifers.

Table 2-4 The predictive power of AVIRIS metrics alone , LVIS metrics alone , LVIS + AVIRIS metrics and LVIS metrics after species stratification of field data was tested over 125 1ha plots. Suitable predictor variables were selected using AIC criteria. The best model was obtained by predicting biomass with LVIS variables after stratifying field plots into WHR/species type

Regression Model	**Predicted Biomass 56.4 m (1 ha) n = 125**	
	R^2 (95% C.I.)	**RMSE (Mg/ha)**
Field Biomass ~AVIRIS (Biomass = 8865*sd NDWI -1624.13*sd NDVI - 5421.92*sd EWT+ 3658.35 mean NDWI - 1218.99*mean EWT -28.90* mean RVI+ 20166.10 sd RVSI -125305.76* mean Rededgemag)	**0.60 (0.49-0.69)**	**92.13**
Field Biomass ~ LVIS metrics (Biomass = 60.58*mean RH75 - 698.24* mean CCover - 27.84*mean RH100 + 149.18)	**0.77 (0.69-0.83)**	**70.12**
Field Biomass ~ LVIS + AVIRIS metrics (Biomass = 55.83*mean RH75-24.76*mean RH100-776.32*mean CCover-526.10* mean1DZ_DGVI + 820.65 mean NDWI -5299.80 sdlc2dpth + 339.99)	**0.80 (0.74-0.86)**	**64.18**
Species Specific Biomass from Field ~ LVIS (Equations in table 6) sd = standard deviation	**0.84 (0.79-0.88)**	**58.78**

Statistical Significance of Models

Confidence intervals for coefficients of determination were calculated for all the models used to predict biomass (Fig. 2-7). Prediction using lidar variables alone showed a statistically significant improvement over the model using AVIRIS variables alone. Addition of AVIRIS variables to LVIS did not show a significant improvement over LVIS metrics alone. Species stratification prior to lidar estimation of biomass reduced prediction errors from LVIS alone by 12%, but the reduction was again not statistically significant. However, confidence intervals for the model using species stratification were narrower than the other models.

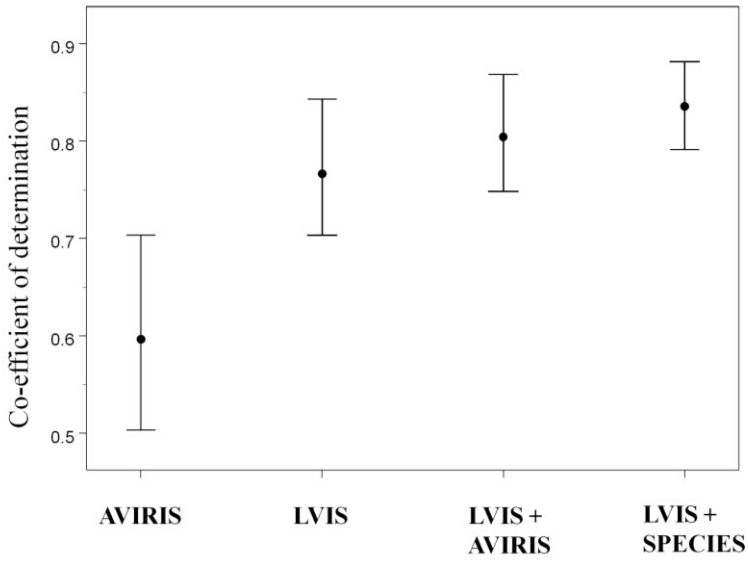

Fig. 2-7 Co-efficients of determination for predicted biomass with 95% confidence intervals. Narrower confidence intervals for LVIS + species stratification suggest a small improvement, but overlap of intervals shows that it is not statistically significant.

2.6.2 Landscape level

Classified land cover map

An error matrix was generated to assess the accuracy of the AVIRIS image classification. 183 reference spectra were modeled using 47 spectra in the selected MESMA model. Classification results showed an overall accuracy of 87.7% for level 1 (genera/species) with a kappa value of 0.86 (Table 2-6). Errors were higher because of mixing between white fir (Abies concolor) and red fir (Abies magnifica) spectra. Level 2 classification (plant functional type/ genera) had a higher accuracy of 93%. It was noted that around 14 spectra in the original library were left un-modeled. Visual comparisons showed the dominant vegetation types in the AVIRIS maps were similar to WHR types in the USDA Forest Service map. At the hectare level, we generated an error matrix using WHR types from field polygons as reference (Table 2-7). Overall accuracy for this classification was 69.5%. Accuracy could be lower (45%) in areas with greater mixing between pines, firs, and hardwoods. Accuracy was also lower because the reference maps were classified as discrete polygons and were at a coarser scale than the AVIRIS classified map.

Table 2-5 Biomass was predicted using a single lidar equation before species stratification and a different equation for each species after stratification of field plots by WHR type.

WHR TYPE	Biomass predicted using single lidar equation Total biomass (actual vs. predicted) r^2 = 0.7667 RMSE = 70.12 Mg/ha		Biomass predicted using a different equation for each species. Species specific biomass (actual vs. predicted) r^2 = 0.84 RMSE = 58.78 Mg/ha	
	r^2 (95% C.I.)	RMSE (Mg/ha)	r^2 (95% C.I.)	RMSE (Mg/ha)
Montane Hardwood Conifer (MHC) n= 10	0.82 (0.41 – 0.96)	22.6	0.94 (0.77 - 0.99) Biomass = 6.03*mean RH75 – 29.60	12.7
Sierran Mixed Conifer (SMC) n = 37	0.77 (0.60-0.88)	60.35	0.78 (0.61 – 0.89) Biomass = 49.69*mean RH75-20.98* mean RH100 -553.06*mean CCover + 77.21	59.73
Red Fir (RFR) n = 32	0.82 (0.67-0.91)	71.02	0.83 (0.68 – 0.91) Biomass = 81.57*mean RH75-40.7512* mean RH100-1009.77*meanCCover+255.02	69.82
White Fir (WFR) n = 27	0.70 (0.45-0.85)	75.59	0.70 (0.45 – 0.85) Biomass =70*mean RH75-31.55*mean RH100 -669.16*mean CCover+86.42	74.63
Pines (PPN) n = 18	0.7542 (0.46-0.90)	28.22	0.75 (0.43 – 0.90) Biomass = 5.95*mean RH75 – 45.84	28.7

Table 2-6 Error matrix for level 1 (genera/species) classification of AVIRIS images using 47 2-endmember MESMA models. Overall Accuracy = 87.7 % Kappa =0.86, Kappa variance = 0.001

Matrix	Firs/ mixed	granite	grass	chaparral	NPV	Pines *(Pinus sp.)*	Hardwoods *(Quercus sp.)*	Red Fir *(Abies magnifica)*	soil	White Fir *(Abies concolor)*
Firs/mixed	2	0	0	0	0	0	0	0	0	1
granite	0	20	0	0	0	0	0	0	3	0
grass	0	0	5	0	0	0	0	0	0	0
Chaparral	0	0	0	24	0	0	0	0	0	0
NPV	0	0	0	0	17	0	0	0	0	0
Pines *(Pinus sp.)*	1	0	0	0	0	6	0	1	0	0
Hardwoods *(Quercus sp.)*	0	0	0	1	0	1	3	0	0	0
Red fir *(Abies magnifica)*	4	0	0	0	0	0	0	8	0	0
soil	0	0	0	0	1	0	0	0	15	0
White fir *(Abies concolor)*	1	0	0	0	0	0	0	1	0	7

Table 2-7 Error matrix showing classification accuracy at hectare level. WHR types from field polygons were used as ground reference. Overall accuracy = 0 69.5%

WHR type/ field polygons	Dominant vegetation type AVIRIS 1ha map				Total Classified
	Hardwoods	Pines	White fir / mixed conifers	Firs/ mostly red	total
MHC/MHW	6	7	5	0	18
PPN	1	2	5	0	8
SMC	0	4	35	0	39
RFR	0	0	10	30	40
Total ground truth	7	13	55	30	105
Overall Accuracy	69.5%				

Biomass

Biomass maps predicted from LVIS before and after species stratification showed large differences in spatial variability, mainly in forests with hardwoods and pines as dominant vegetation type (Fig. 2-8). Histograms of biomass distribution before and after species stratification showed increases in predicted values for both

hardwoods and pines in low biomass ranges (<50 Mg/ha) and decreases in high ranges (> 200 Mg/ha) (Fig. 2-9). High biomass firs and mixed conifer stands showed little variation in predicted values before and after species stratification. Histograms of biomass for hardwoods and pines derived using classified AVIRIS maps and from USDA Forest Service vegetation maps showed similar trends in low and high biomass ranges.

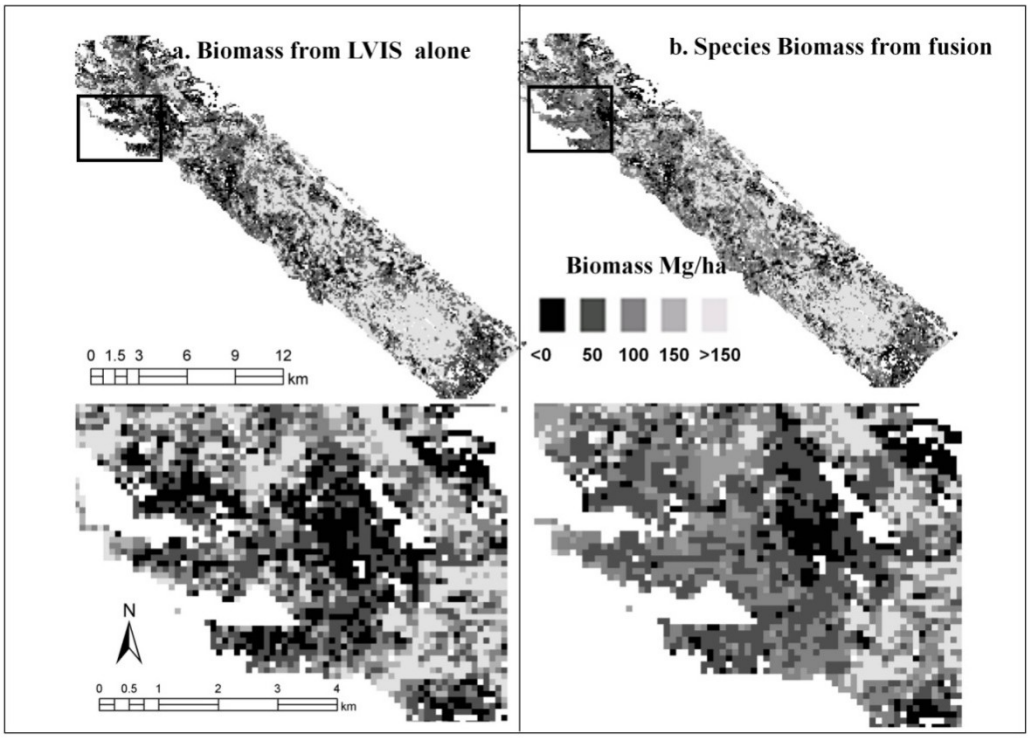

Fig. 2-8 Landscape maps of biomass were generated from LVIS before (8a) and after species stratification of AVIRIS imagery (8b), using equations in Table 2-4 & 2-5. Forests dominated by pine and hardwood species (e.g. black rectangle) show more spatial variations in predicted biomass.

Stressed Biomass

Combined analysis of AVIRIS and LVIS metrics revealed spatial patterns that could not be detected from either sensor alone (Fig. 2-10). For most of the study area, water band indices, red edge derivatives and vegetation indices had very low values

over barren land, exposed rock surfaces, and higher values over dense forests and wet meadows. However, we found low values for water band indices and vegetation indices in some high biomass (>200 Mg/ha) forests, particularly in red fir stands around the Teakettle Experimental Forest (Smith et al., 2005) and some mixed conifer stands. (Fig.2-11). NPV fractions were also high in the areas where we detected water and chlorophyll stress (Fig.2-12).

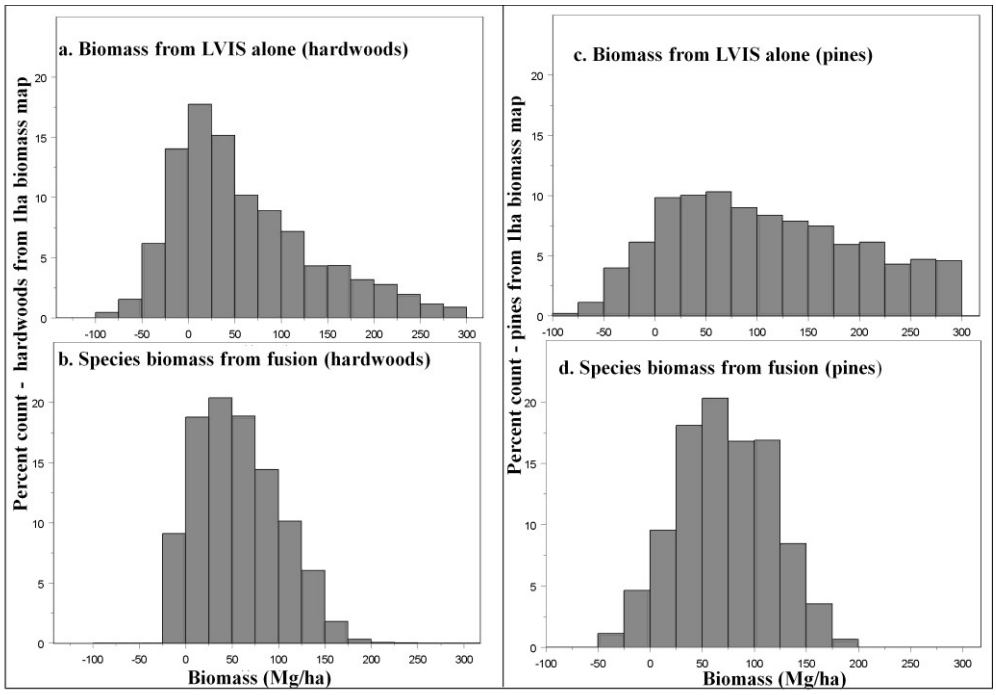

Fig. 2-9 Histograms showing differences between biomass predicted before (9a, 9c) and after (9b, 9d) species stratification of AVIRIS imagery. Stratification for hardwoods and pines increased predicted values in low (<50 Mg/ha) ranges and decreased values in high ranges (>200 Mg/ha) of biomass.

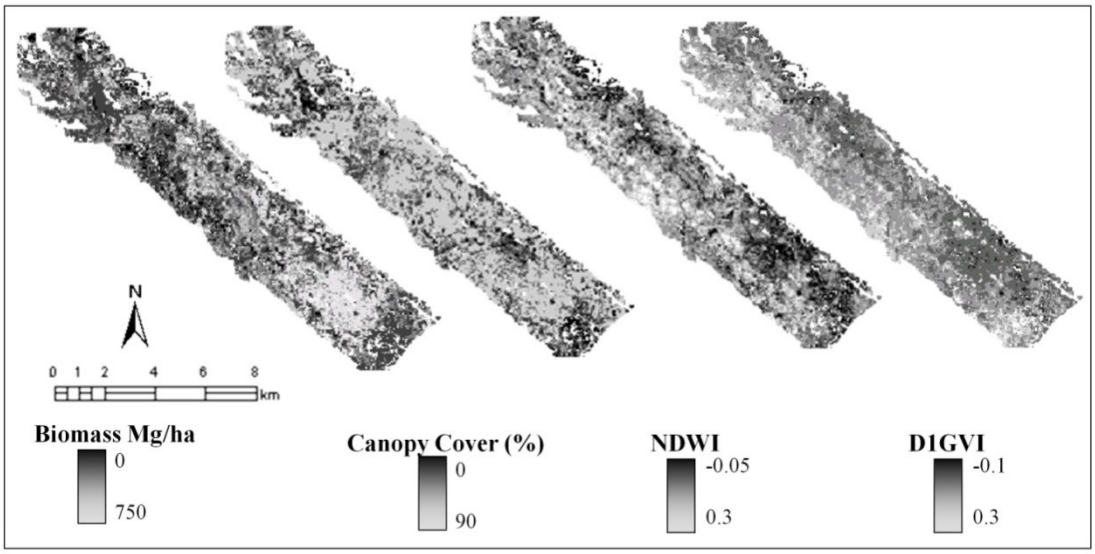

Fig. 2-10 Landscape maps of biomass, canopy cover, NDWI and D1GVI used for detecting water and chlorophyll stress in high biomass forests.

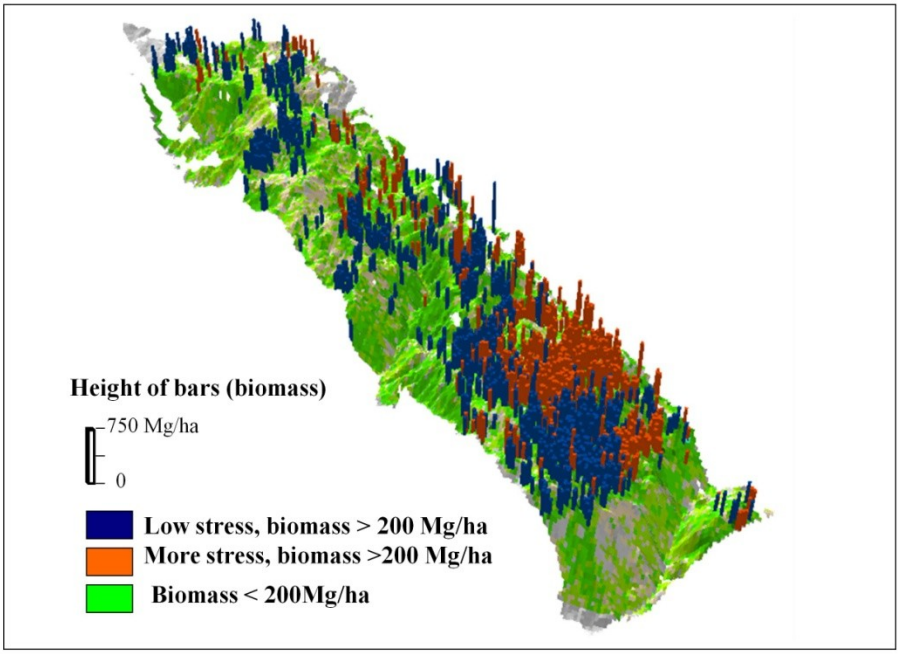

Fig. 2-11 Map showing high and low stress in stands with high biomass (>200 Mg/ha). Height of bars represents biomass values.

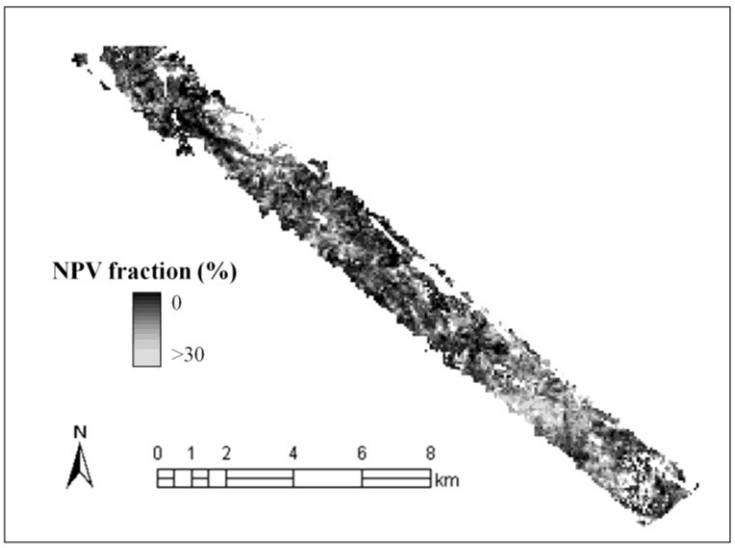

Fig. 2-12 Spatial distribution of non-photosynthetic vegetation (NPV) fractions within canopies for one AVIRIS image (1 ha level). Stands with high biomass and stress (Fig.2-11) also showed high NPV values.

2.7 Discussion

Our first objective was to test the efficacy of combining hyperspectral metrics with lidar variables for biomass prediction. AVIRIS band indices and MESMA fractions added little explanatory value to LVIS, even though they explained around 60% of the variability in biomass at the stand level. This was because of strong correlations between LVIS and AVIRIS metrics, particularly water band indices and shade fractions. Similar relationships between field measured canopy height and shade fractions were reported by Numata et al. (2008). Shade fractions are related to canopy structure, so this correlation is not unexpected. Roberts et al. (2004) showed that liquid water is highly sensitive to LAI, which may explain the observed correlation between water band indices and structural metrics from LVIS that respond most strongly to photosynthetic leaf material. Our results also showed that narrow

band AVIRIS metrics such as red edge derivatives were more sensitive to biomass than NDVI, similar to Elvidge and Chen (1995), Roberts et al. (1997), and Roberts et al. (2004). However, these metrics suffered saturation effects over the high biomass range of this study area.

Although AVIRIS metrics did not add much predictive power in our moderately high biomass test area, shade fractions and water band indices may be useful in areas with lower biomass and little or no lidar coverage. Approximately 40% of the world's forests fall in the low canopy height, low biomass category (Bergen et al., 2006), where lidar performance is largely untested. The potential of hyperspectral and other optical imagery in extrapolating forest structure from lidar samples in such areas requires further investigation.

Previous studies combining lidar with multispectral (Hyde et al., 2007a), radar (Hyde et al., 2007b) and hyperspectral sensors (Anderson et al., 2008) have shown that lidar was more useful than other sensors for biomass prediction. Our results further support this. Drake et al. (2002a) and Anderson et al. (2008) have shown the predictive power of the RH50 metric and suggested the use of canopy cover to improve biomass estimates from LVIS. We additionally included RH75 and canopy cover to the variables used by Hyde et al. (2005) for the Sierra Nevada. Although both RH50 and RH75 were strongly correlated with biomass, RH75 was consistently selected as the best predictor variable in all regression models. One probable reason could be the species composition and vertical foliage distribution in this study area. Further analysis of lidar waveforms, foliar profiles, and stem densities within lidar

footprints is required to understand the physical significance of RH75 in biomass estimation for the Sierra Nevada.

The issue of the efficacy of fusing lidar and hyperspectral data for species level biomass estimation remains open. Similar to Anderson et al. (2008), our results show that a combination of LVIS and AVIRIS metrics improves biomass estimates marginally than using either sensor alone. Anderson et al. (2008) found that AVIRIS metrics explained most of the variability in species fractions of biomass for northeastern temperate forests. Our results show that LVIS metrics were better predictors of species level biomass (Table 2-5) while AVIRIS metrics were mostly redundant when combined with LVIS. One reason could be the difference in tree species in the Sierra Nevada as compared to Bartlett. Most of the species in Bartlett are broadleaf deciduous, while in the Sierra Nevada they are conifer dominants. Another reason could be that the predictive power of AVIRIS is higher when lidar relationships with biomass are weaker as observed in the Bartlett Experimental Forest. A study by Roth (2009) showed similar results for the Smithsonian Environmental Research Center (SERC) study site. Lidar metrics in the Sierra Nevada study area were strongly correlated with biomass, so addition of AVIRIS probably did not show much improvement.

The overlap of confidence intervals of the co-efficients of determination before and after species stratification suggests that overall predictive power for biomass was not significantly higher at the species level for our study area. Part of the reason for this could be the relatively small sample size used in this study. The dominance of high biomass mixed conifers and low abundance of deciduous species

in the study area could have also affected the results. Yet another factor could be the relatively coarse spatial scale of 1 ha used in our study, one that is large enough to encompass various species and canopy configurations. These limitations aside, stratification seemed to perform better at lower biomass levels. Increased prediction accuracy, lower RMSE values, and narrow confidence intervals suggest a small improvement with species stratification (Fig. 2-7).

We tested both linear and non-linear variables for all regression models. Best-fit models were obtained with linear combinations of variables. Although there is an apparent non-linear trend in Fig. 2-6a and Fig. 2-6c., it is because of the poor predictive power of the models in low biomass plots (<50Mg/ha). The RMSE values from the regression models should be interpreted in terms of model-to-model comparisons rather than an absolute measure of accuracy in a mapping perspective.

Spatial predictions of biomass from LVIS were quite different before and after species stratification by AVIRIS. Relative to species-level equations, a single lidar equation underestimated values in the lower ranges and overestimated it in the higher ranges of biomass, particularly for hardwoods and pines. Using a different lidar equation for hardwoods and pines reduced apparent errors in lower ranges of biomass for both these vegetation types (Fig. 2-9b & 2-9d). The trend towards reduced error and improved prediction accuracy was clear (Fig. 2-7) even at stand level analysis for hardwoods but not for pines.

Fusion of lidar and hyperspectral sensors at species level and in areas with low biomass is an important remote sensing research requirement (Rosenqvist et al., 2003; Treuhaft et al., 2004; Bergen et al., 2006). Our study shows that species

stratification could potentially improve predictions from sparse lidar samples, in low biomass regions better than fusion with spectral metrics. More work is needed to confirm these results over larger samples and homogenous stands. Improving classification accuracies for individual species by using field spectra may further refine spatial prediction of biomass from AVIRIS. Also the optimum level of classification (plant functional type, genera or species) and scale (1 ha or less) must be studied further.

Intuitively, we would expect species stratification to provide an improvement because the data used for biomass ground truth is routinely derived from forestry tables on a species-level, just as we did in our research here. However, there is the larger, and unanswered question, of whether lidar metrics are sensitive to species-level differences in canopy vertical structure, canopy gap spatial pattern, stem density and stem spatial pattern, among others, that should be predictive of biomass, and at what spatial scales. While species-specific predictions as applied in this study could improve estimates over other forested areas, the true impact of a priori stratification may never be known unless this problem is explored thoroughly.

We did not expect a significant change in species composition within the time lag between lidar and hyperspectral data acquisition. However, some uncertainty in spectral metrics related to changes in structure and stress may have affected the outcome. Another limitation was that only large trees (>76cm dbh) were measured in 1 ha plots. Footprint level plots (0.07 ha) included measurements of all trees above 10cm dbh but were not included in this study because of increased geolocation errors between reprocessed LVIS data and 2000/2001 field plot centers. Better geolocation

of field, lidar, and hyperspectral data may help, but this may also only increase correlation between metrics rather than improve biomass estimates.

Spatial maps of various AVIRIS metrics in combination with LVIS maps showed increased water stress in many high biomass red fir (*Abies magnifica*) and mixed conifer stands. High values of NPV fractions within canopies in addition to low vegetation and water band indices, suggests increased stress and mortality in these areas. Moisture stress was high in open stands with more canopy gaps as well as in dense stands, consistent with findings from Smith et al., (2005). Our results are similar to recent studies linking water stress and increased tree mortality in the Sierra Nevada (van Mantgem et al., 2009; Lutz et al., 2009).

Areas within the Teakettle Experimental Forest (North et al., 2002), where red fir was the dominant vegetation type also showed a large number of NPV spectra in the 2003 AVIRIS images. Subsequent field observations in 2008 showed abundant dead trees as well as evidence of logging in these areas. Further analysis is required to confirm whether stress maps from 2003 AVIRIS images showed early indications of the tree mortality observed in 2008. Presumably, lidar/hyperspectral data could be used to map areas of high stress and mortality in response to climate change as suggested by Van Mantgem et al. (2009).

2.8 Conclusion

Species stratification may improve predictions from lidar, a result only suggested by our work, as overall predictive ability did not improve significantly; however, confidence intervals were narrowed and biomass showed very different

spatial variability when mapped across the landscape. Extrapolating structure from lidar samples with stratified optical data can be a promising strategy for mapping low biomass forests from future space borne lidar sensors such as DESDynI. Such species-specific biomass maps have the potential to be exceptionally useful for carbon and ecosystem modeling.

AVIRIS indices and MESMA fractions provide quantitative measures of canopy condition and can be of considerable value in ecological applications, when combined with lidar. We demonstrated one such application here, by mapping stress in high biomass forests of Sierra Nevada. Stress maps can serve as early indicators of mortality, drought, and fire susceptibility in old growth forests and help improve forest management practices. Classified vegetation maps can be further used to study regeneration from fire or combined with small footprint lidar data to map individual tree biomass/mortality.

Lidar can provide measures of vertical structure such as canopy height, understory cover, and foliage diversity while species composition, stress, and decadence can be obtained from hyperspectral data. Fusion of the two sensors is therefore, powerful for biodiversity and habitat studies. Future research will focus on combining the two sensors for mapping potential habitats for rare and endangered bird species.

Acknowledgements

This research was funded in part by "Multisite Integration of LIDAR and Hyperspectral Data for Improved Estimation of Carbon Stocks and Exchanges", (P.I. Dar Roberts), NASA Carbon Cycle Science grant (NNG05GE56G) and a NASA Earth and Space Science graduate fellowship (NNX06AF91H). Error matrix codes for assessing accuracy of MESMA models were made available by Phillip Dennsion. We are thankful to Kerry Halligan, Dylan Parenti and the Viper Lab team at University of California, Santa Barbara for software training and help with AVIRIS data processing. We thank Peter Hyde for providing useful inputs from previous studies in the Sierra Nevada.

Chapter 3 Mapping Canopy Height and Biomass Dynamics in the Sierra Nevada using Waveform Lidar

3.1 Introduction

Changes in forest structure from disturbances such as fires, wind throw, insect outbreaks and logging reduce the amount of carbon stored in the form of aboveground biomass. On the other hand, recovery from disturbances and growth in canopies increases biomass and carbon storage. The lack of quantitative estimates of such changes leads to large uncertainties in carbon flux in forests (Houghton et al., 2010). Improved measurements of canopy dynamics are critical for reducing these uncertainties (Frolking et al., 2009) and understanding how ecosystems respond to disturbances (NRC, 2007).

Field data on canopy dynamics are generally sparse spatially with most in the form of averages at coarse resolutions (Houghton, 2005). Although passive remote sensing data have been extensively used in change detection (Coppin et al., 2004), the focus has been on mapping areal extents of change and estimating carbon flux with models. It has been suggested that these models can be improved or validated independently with direct measurements of canopy height and biomass changes from active remote sensing (Houghton et al., 2010).

Light Detection and Ranging (lidar) is an active remote sensing system that can measure and map three-dimensional vegetation attributes, including canopy height and cover, at various spatial scales. Lidar metrics have been used to derive aboveground biomass in tropical and temperate forests (Lefsky, 2002a; Drake et al.,

2002a; Hyde et al., 2005; Anderson et al., 2006; Dubayah et al., 2010) with greater accuracies than other remote sensing data. Most lidar maps of forest structure, however, are static and do not provide information on changes (Frolking et al., 2009). There is considerable interest in directly mapping canopy dynamics from temporal lidar data, but the errors and accuracies are largely unexplored.

In a recent study, Kellner et al. (2009) used small footprint lidar data with high spatial resolution to quantify gaps and analyze height transitions over old-growth tropical forests in Costa Rica. By comparing lidar height distributions from 1997 and 2006 with projected equilibrium conditions, they showed that these forests were in steady state, i.e., canopy height increases from growth balanced losses from tree mortality. Dubayah et al. (2010) further quantified canopy height and biomass changes across successional types at La Selva with medium footprint (25 m) lidar data and also mapped the distribution of potential carbon sources and sinks. This study evaluated the strengths and challenges of measuring changes from medium footprint lidar and discussed its implications for future space-borne sensors.

Simultaneous observations of airborne lidar and field data over a span of nearly a decade in the Sierra Nevada in the western United States provide an opportunity to test similar approaches in temperate montane forests. These forests are increasingly becoming a cause of concern because of catastrophic fires, insect attacks and higher-than-usual tree mortality from moisture stress (van Mantgem et al., 2007). Empirical maps of changes can help improve management efforts for fire, timber and habitat conservation.

The goal of this study was to evaluate the use of temporal lidar data from the Laser Vegetation Imaging Sensor (LVIS) in quantifying and mapping canopy dynamics over the montane forests of the Sierra Nevada. We tested whether LVIS measurements in 1999 and 2008 could predict canopy height and biomass changes observed in the field. We also evaluated the potential of directly measuring canopy dynamics from lidar data at footprint, plot (0.07 ha) and hectare scales. Results from this research can add to the growing body of knowledge on lidar remote sensing applications for forest monitoring and carbon modeling.

3.2 Study Area and Data

3.2.1 Study Area

The study site lies in the Sierra National Forest in California, USA (Fig. 3-1). The region is characterized by a Mediterranean climate and has elevations ranging from 1000 m to 2500 m. Distribution of vegetation is largely determined by climate and topography (Raumman and Soulard, 2007). Dominant tree species include red fir *(Abies magnifica)*, white fir *(Abies concolor*), ponderosa pine *(Pinus ponderosa),* and California black oak *(Quercus kellogi).* The landscape is a mosaic of patchy and heterogeneous forests (North, 2002), shaped by natural and anthropogenic disturbances such as fire (Collins et al., 2006), insect outbreaks, (Das et al., 2008), thinning treatments for fire suppression (Rambo & North, 2009) and commercial timber harvests (Raumman and Soulard, 2007). A major portion of the study site is under uneven-aged forest management. Patches of mixed tree species are in different

stages of regeneration from past disturbances (Kern et al., 2008). Areas reserved for old growth and wildlife habitats are relatively undisturbed (Hunsaker et al., 2002*)*.

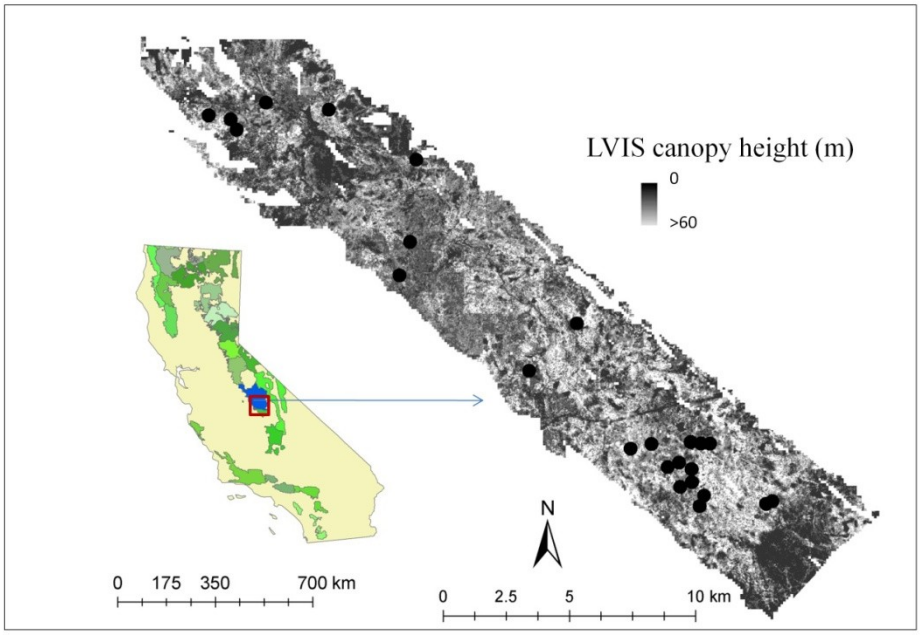

Fig. 3-1 Study site in the Sierra National Forest, California showing field plots measured in 2000/2001 and 2008.

3.2.2 Field Data

A detailed description of field data acquisition in the Sierra Nevada is available in Pierce et al., (2002) and Hyde et al., (2005). Field surveys were conducted in 2000/2001 (Hyde et al., 2005) in over 285 concentric circular plots with 0.07 ha and 1 ha areas, respectively (Fig. 3-2). Structural variables including location, tree height, diameter at breast height (dbh), crown form and species were recorded for all trees greater than 10 cm dbh within the inner 0.07 ha plot. In the outer 1 ha plots, dbh and height for trees greater than 76 cm were recorded. In 2008, we re-measured

25 of these plots to obtain a statistically representative sample of canopy height, cover, biomass and topography. The 2008 1 ha plots were subdivided into 9 square subplots (3X3) with 33.3 m on each side (0.11 ha) and were oriented upwards along slopes (Fig 3-2). The square subplots were slightly larger than the 2000/2001 inner circular plots to account for collocation offsets between lidar and field data. In the central subplot, all trees mapped in 2000/2001 were identified and re-measured. In the 8 outer subplots, dbh was recorded for trees greater than 10 cm. The hectare data from 2000/2001 were not used in this study because only large trees were measured.

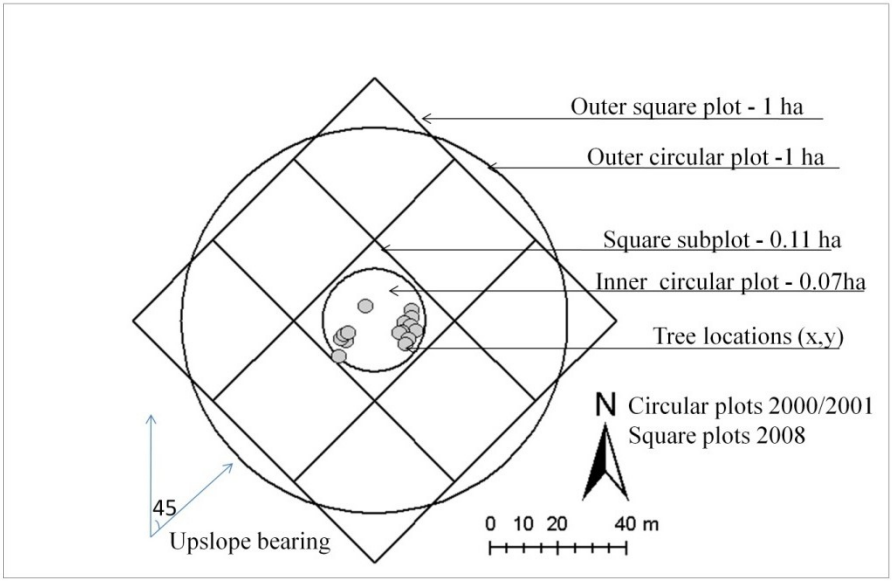

Fig. 3-2 Schematic layout of field plots in 2000/2001 and 2008. The 2008 square subplots were slightly larger than the inner circular plots to account for geolocation shifts between field and lidar data. Plots were oriented upslope.

3.2.3 LVIS Data

The Laser Vegetation Imaging Sensor (LVIS) is a large footprint waveform digitizing, scanning laser altimeter, designed and developed at NASA's Goddard Space Flight Center. LVIS operates at altitudes up to 10 km with a 7º field of view and uses laser pulses with a wavelength of 1064 nm for profiling vertical vegetation

structure (Blair et al., 1999). The entire outgoing and return signal is digitally recorded to provide a waveform describing the reflected energy from intercepted surfaces. NASA flew LVIS over the Sierra National Forest in the summer of 1999 with trees in leaf-on condition. The lidar shots had a nominal footprint radius of 12.5 m. Footprints were contiguous across track and overlapping along track, covering an area of about 175 sq. km. (See http://lvis.gsfc.nasa.gov/). Final coverage was denser in some areas because of repeated flight lines. The site was re-flown in 2008 to acquire coincident data. Parts of the study area were not mapped in one or both years because of cloud cover, flight path irregularities, and noise. These 'no data' areas were larger in 1999 as compared to 2008, which oversampled the study area. Because of overlapping flight lines, the number and spatial distribution of lidar footprints varied between years across the landscape.

The digitized LVIS return signals were geolocated with respect to the WGS84 ellipsoid. LVIS waveforms were processed using Gaussian decomposition methods outlined in Hofton and Blair (2002). Ground elevation was determined by finding the mode of the lowest peak in the waveform with amplitude greater than 3 σ of the noise. The canopy top was identified as the largest return above the noise threshold at the top of the waveform. The difference between the ground return and canopy top was measured as canopy height (RH100). Other metrics calculated from the waveform were heights of 25% (RH25), 50% (RH50), 75% (RH75) energy return (Drake, 2002a; Hyde et al., 2005; Anderson et al., 2006), canopy cover (Ni Meister et al., 2001). A detailed description of LVIS waveform processing and associated errors is available in Hofton et al. (2006) and Dubayah et al. (2010).

Because of potential ground finding errors, canopy elevation metrics, i.e., LVIS RH + elevation, hereafter 'RH_E' were used wherever possible (see Dubayah et al., 2010 for a detailed description). RH metrics were used for predicting biomass separately from 1999 and 2008 LVIS data because RH_E metrics were not applicable to individual year predictions. At footprint scale, we minimized elevation errors and therefore used RH metrics (described below) to measure changes. For height changes across the landscape, RH_E metrics were used because this method of correction is only applicable to co-incident footprints as opposed to characterizing changes over target 0.07 ha and 1 ha areas that use averages of footprints within each year.

3.3 Methods

The datasets used in this study were projected to UTM 11N and WGS 84 ellipsoid. LVIS waveforms were processed using IDL 7.2© and analyses were performed using ESRI© Arc GIS 9.3 and R© statistical software. We first tested whether lidar plot average change metrics were sensitive to canopy height and biomass changes observed in the field between 1999 and 2008 using linear and stepwise regression methods. Next, we directly calculated height changes in nearly coincident LVIS footprints and analyzed canopy height dynamics using transition probability matrices. Finally, we mapped canopy height and biomass changes across the landscape at 0.07 ha and 1 ha scales and validated these changes outside of plot data areas with optical imagery (See Fig. 3-3).

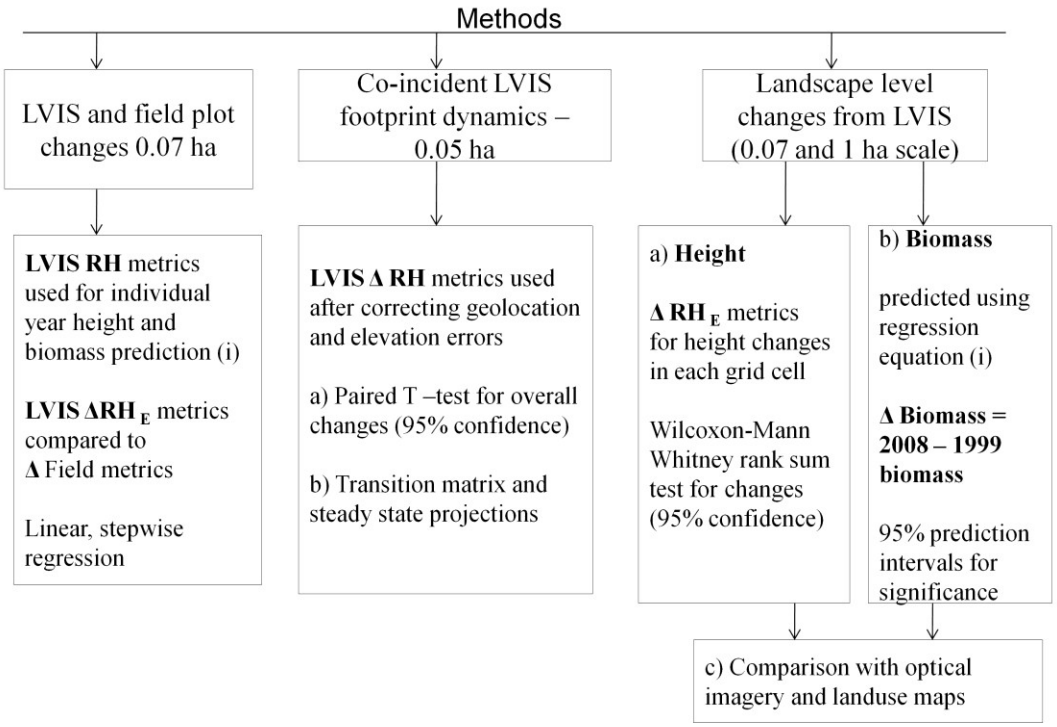

Fig. 3-3 Flowchart showing summary of methods and LVIS RH/RH_E metrics used.

3.3.1 LVIS and field measured changes

Out of the 25 re-measured 0.07 ha field plots, 18 had a tree-to-tree stem plot match in both years. One plot was removed because the 1999 LVIS footprints were close to the edge of the plot (Fig. 3-4) leaving 17 plots for analyses. Maximum, average and Lorey's (basal area weighted average) heights were calculated for both years. Aboveground biomass was calculated as the sum of bole, bark, live branch, and foliage for each tree using species-specific allometric equations from the California Inventory Manual (Waddell & Hiserote, 2003). For dead trees, biomass also calculated using bark and bole biomass, excluding foliage and live branches (Kim et al., 2009). Total biomass changes (2008-1999) for all live and dead trees within each plot were summarized. Neither 1999 nor 2008 were LVIS footprints exactly

coincident with field plot centers, so we averaged ΔRH_E metrics for all footprints that had their centroids within the 0.07 ha plots. Lidar metrics were validated with field attributes for each year separately. Changes in lidar and field data were compared using linear and stepwise regressions.

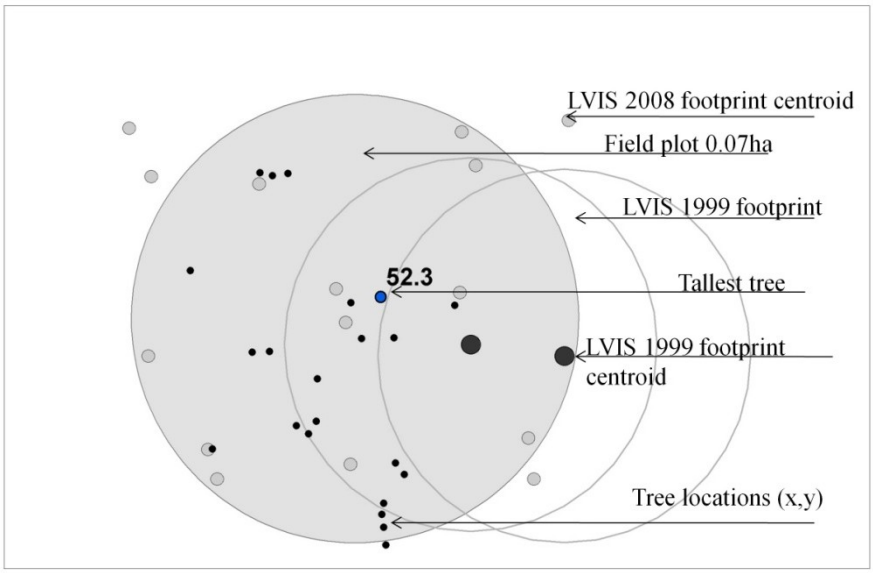

Fig. 3-4 Geolocation shifts between field plot locations and lidar data affected comparisons of LVIS and field metrics. The two 1999 LVIS footprint centers are 7m and 14.5m away from plot center respectively. The 2008 LVIS data (only footprint centroid shown) oversampled field plots but were not co-incident with field plot centers.

3.3.2 Coincident footprint dynamics

Height changes

Assuming that coincident footprints have no elevation difference in the time period studied (e.g., from tectonics, erosion etc.), we determined the relative geolocation shift between the two datasets by finding the distance at which average elevation errors were near zero. LVIS footprint centers from 1999 and 2008 were selected from the entire coverage (not the study area alone), at consecutive 0.5 m

distance intervals, i.e., 0-0.5 m, 0.5-1 m, 1-1.5 m and so on. Average elevation errors were plotted as a function of distance

(Fig.3-5).

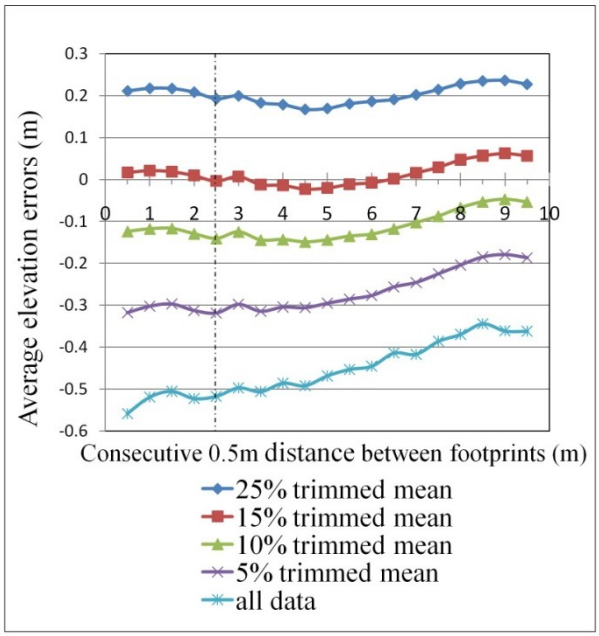

Fig. 3-5 Average ground elevation error between 1999 and 2008 (Δ ZG) as a function of distance between lidar footprints. Errors were lowest at 15% trimmed means and closest to zero for shots between 2 and 2.5m distance from each other. This suggests an average relative geo-location error of 2 - 2.5m between 1999 and 2008 footprints.

Trimmed means (Caprra & Rivest, 1995) of errors were also plotted along with original errors for all distances. Trimming allows the removal of outliers or extreme values affecting the mean, most of which in this case were related to instrument noise and ground finding errors in the data rather than actual change. The lowest average elevation errors were obtained with 15% trimmed means (15% extreme values on either side of the mean removed). Errors after trimming were closest to zero for a footprint distance of 2-2.5m and increased thereafter. This suggests an average relative geolocation shift of 2-2.5 m between the two datasets.

We therefore used only those lidar footprints that were between 2 and 2.5 m of each other with 15% trimmed means for analyzing footprint level height changes. Paired t-tests were used to determine whether the overall height changes from co-incident footprints were significant.

Height transition matrix

Transition probability matrices or Markov chains have been extensively used to analyze forest transitions from one ecological state to another (Usher, 1981; Hall et al., 1991). A transition matrix gives the probability of an element in 'j' class at time't' to be in 'i' class at time 't+1' (Biondini & Kandus, 2006). A basic assumption is that the future state of an element in the matrix depends only on the current state and not on past conditions (Perry & Millington, 2008). Transition matrices have been used to study canopy height dynamics in tropical forests using lidar observations (Kellner et al., 2009; Dubayah et al., 2010) and temperate forests using aerial photographs (Tanaka & Nakashizuka, 1997). Matrices were generated in this study using LVIS heights (RH100) from the co-incident footprints corrected for geolocation errors. The rows of the height transition matrix show the likelihood of canopies occupying past height classes or retrospective probabilities, while the columns show those occupying future classes or prospective probabilities. The diagonal shows probabilities of staying in the same class or no net change. The triangle above the diagonal shows probabilities canopy height loss while the one below the diagonal shows probabilities of growth.

Forest transitions are expected to eventually reach a steady or equilibrium state, similar to a Markov chain. By analyzing transition probability matrices, it is possible not only to compare changes within a given time interval, but also project distributions into the future at successive time intervals, provided the disturbance regime remains fairly unchanged (Usher, 1981; Caswell, 2000). Canopy height projections over successive nine-year time intervals were calculated as $x(t) = x*P^t$, where x is the height distribution at time t and P is the transition matrix (Perry & Millington, 2008). Steady-state canopy height projections were obtained by solving for the dominant right-handed eigenvector of the matrix (Caswell, 2000).

3.3.3 Landscape level changes from LVIS

Direct canopy height changes (0.07 ha and 1 ha)

In addition to calculating changes at footprint scale, we also analyzed spatial patterns of change across the landscape. The 0.07 ha scale was used to create a map comparable to field plot scale and the 1 ha scale was used because it is a scale used for analyzing biomass dynamics with models (See Dubayah et al., 2010). Canopy elevations (Rh_E100) from LVIS footprints within 0.07 ha and 1 ha grids were averaged and subtracted to map changes across the landscape. To determine how many grid cells showed statistically significant change at 95% confidence, we compared height changes within each cell using the non-parametric Wilcoxon-Mann-Whitney rank sum test. This test was preferred to the two sample t-tests because the number of lidar footprints within each grid cell was small, and normal distribution could not be assumed. The Wilcoxon-Mann-Whitney test ranks observations from

both datasets and calculates the probabilities of the distributions being identical or significantly different. Results from the test were mapped across the landscape at both 0.07 ha and 1 ha scales.

Predicted biomass changes (0.07 ha and 1 ha)

The selection of field plots in 2008 was to cover the range of biomass for DESDynI rather than biomass change. The lack of sufficient re-measured field plots made it difficult to derive an equation to predict changes across the landscape as in Dubayah et al. (2010). We therefore estimated biomass individually for each year using a lidar regression equation derived over 126 0.07 ha field plots measured in 2000/2001. A similar approach was tested over 215 subplots (9 subplots in each 1 ha plot) measured in 2008. Predictions from both years were similar but the 1999 regression had lower residual errors. Field allometry was also more robust in 2000/2001 because biomass was calculated with both dbh and height as against only dbh in 2008 using Jenkins et al., 2003. We therefore used the 1999 lidar equation to predict biomass for both years and subtracted the two products to map changes across the landscape. Statistically significant changes were identified by selecting areas where 95% prediction intervals for both years did not overlap. The same equation was applied to average RH metrics at 1 ha scale.

Comparisons with optical imagery/landuse maps

Spatial maps of change from LVIS at 1 ha scale were compared with optical imagery as an alternate form of validation since plot level changes could not be

validated with field measurements as described. A visual metric was used to validate height losses with high resolution using aerial photos (Google Earth©). A random sample of 100 1 ha cells with significant canopy height loss estimated from LVIS was overlaid with temporal aerial photos and visually assessed for canopy cover change. We also compared canopy height changes with Land Management maps (USDA Forest Service, 2006). More canopy height increases were expected in protected areas than in areas under timber management. Lastly, we compared height change maps with stress maps (Swatantran et al., in press) generated in a previous study with lidar and hyperspectral data. The high stress areas had biomass > 200 Mg/ha but low moisture and chlorophyll content while low stress areas had both high biomass and high moisture/chlorophyll content. We hypothesized that areas predicted as high stress from hyperspectral data would show more canopy height losses and vice versa.

3.4 Results

3.4.1 LVIS and field measured changes

We used RH metrics for validating lidar data for individual years and RH_E metrics for change comparisons. Field heights were strongly correlated with RH100 for both 1999 and 2008 with similar root mean squared errors (R^2= 0.84, RMSE = 4.41 m for 1999 and R^2 = 0.85, RMSE = 4.07 m for 2008) (Fig. 3-6). RH50 was a good predictor of biomass for both years and prediction errors were similar (R^2= 0.70, RMSE = 186.4 Mg/ha for 1999 and R^2 = 0.79, RSE = 167.5 Mg/ha for 2008) (Fig. 3-7). Although LVIS metrics and field attributes were in good agreement for individual years, changes in RH_E metrics did not show any clear relationship with changes in average height, Lorey's height or biomass from field data. We discussed possible

reasons for this; however, the most important one was that changes in re-measured field plots were too small (Fig. 3-8) to be detected by LVIS.

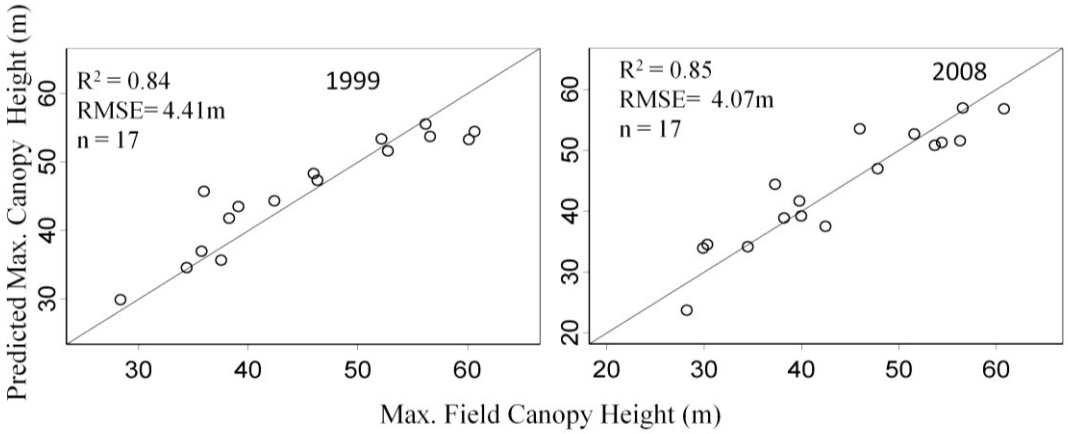

Fig. 3-6 Maximum field height was strongly correlated with LVIS RH100 for both years with similar errors.

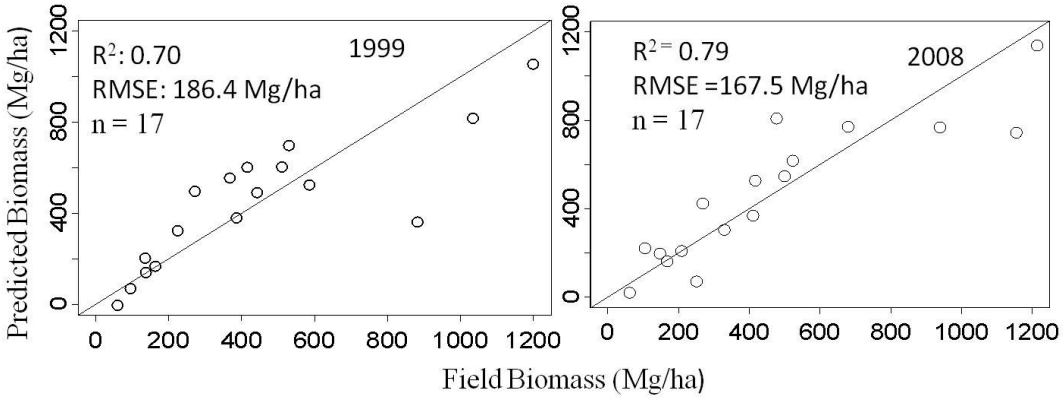

Fig. 3-7 Field biomass was strongly correlated with LVIS RH50 for individual years with similar errors. Note that these equations were not used to predict biomass for 1999 and 2008 across landscape.

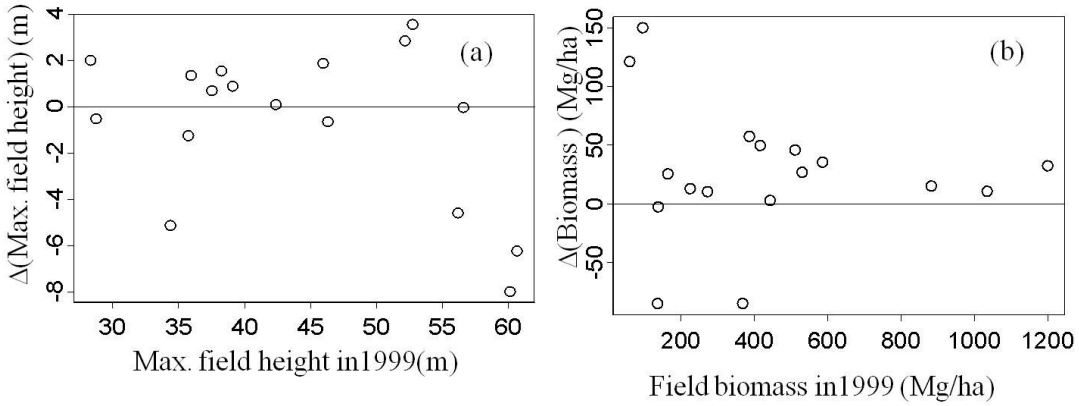

Fig. 3-8 Distribution of maximum field height changes relative to 1999 heights (a) 4 plots showed height decreases greater than 4m. The remaining show small changes, with increases slightly exceeding decreases. Biomass change relative to 1999 biomass (b)Two plots showed more than 50Mg/ha biomass gain and two less than 50Mg/ha loss. Both height and biomass changes were too small to be detected by LVIS metrics.

3.4.2 Coincident footprint dynamics

Height Changes

LVIS footprints between 2 and 2.5 m of each other were selected with 15% trimmed means to reduce geolocation and elevation errors. This resulted in a comparison of over 100,000 coincident footprints. Since elevation errors were reduced to a minimum, there was practically no difference between ΔRH and ΔRH_E metrics for these footprints, e.g., average ΔRH100 was 0.69 m +/- 7.94m and ΔRH_E100 was 0.69 +/ 7.91 m SD. Paired t-tests showed overall small but statistically significant positive changes in all LVIS RH metrics (Table 3-1).

Table 3-1 Changes in LVIS RH metrics from nearly co-incident footprints.

Scale	ΔRH100 (m)		ΔRH75 (m)		ΔRH50 (m)		ΔRH25 (m)	
Co-incident Footprint N = 113471	0.69	7.94	0.47	5.56	0.26	3.88	0.16	2.12

Height transition matrix

In general, the probabilities along the diagonal of the transition matrix or the likelihood of canopies staying in the same class were higher than transitions to other height classes (Table 3-2). Shorter canopies (< 40 m) showed more transitions to the immediate higher 5 m height class. Transitions to higher and lower classes were nearly equal between 35-40 m. In canopies taller than 40 m, shifts to lower height classes were more likely than shifts to higher height classes. Beyond 60 m, there was an abrupt drop in the probability of canopies staying in the same class while transitions to lower height classes increased exponentially (Fig.3- 9). There was a 5 -9 % probability of transitions to higher classes showing growth of more than 10 m. Height distributions from coincident footprints were bimodal in both 1999 and 2008 and did not match the projected steady state (Fig.3-10) although 2008 heights were closer to equilibrium than 1999. Projections from the transition matrix showed that canopy heights would gradually shift towards a normal distribution over successive nine-year time intervals and would hypothetically reach steady state in about 350 years under the current disturbance regime (Fig. 3-10).

Table 3-2 Transition matrix for LVIS canopy heights (RH100) co-incident footprints in 1999 and 2008. Reading down columns gives prospective transition probabilities. Reading across rows gives the retrospective transition probabilities. The diagonal shows the probability of staying in the same class.

	1999 RH100 (m) N = 113471															
2008 RH100 (m)	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	65-70	>70	N
0-5	**0.67**	0.08	0.01	0.01	0.01	0.01	0.01	0	0	0	0	0.01	0.01	0.03	0.08	6383
5-10	0.23	**0.57**	0.09	0.03	0.02	0.01	0.01	0.01	0.01	0.01	0	0.01	0.01	0.06	0.13	13880
10-15	0.03	0.25	**0.46**	0.1	0.03	0.02	0.01	0.01	0.01	0	0	0.01	0.01	0.1	0.11	12162
15-20	0.02	0.05	0.3	**0.43**	0.11	0.04	0.02	0.01	0.01	0.01	0.01	0.01	0.02	0.06	0.09	11699
20-25	0.01	0.02	0.08	0.29	**0.43**	0.13	0.04	0.03	0.02	0.01	0.01	0.01	0.02	0.06	0.14	12355
25-30	0.01	0.01	0.03	0.08	0.28	**0.42**	0.14	0.06	0.03	0.02	0.02	0.02	0.04	0.02	0.1	12509
30-35	0.01	0.01	0.01	0.03	0.08	0.26	**0.43**	0.15	0.06	0.04	0.03	0.02	0.02	0.03	0.1	12046
35-40	0.01	0	0.01	0.01	0.03	0.07	0.25	**0.45**	0.17	0.07	0.04	0.03	0.03	0.03	0.09	11046
40-45	0	0	0.01	0.01	0.01	0.03	0.06	0.22	**0.46**	0.18	0.08	0.04	0.04	0.03	0.05	8889
45-50	0	0	0	0	0.01	0.01	0.02	0.05	0.19	**0.45**	0.2	0.1	0.07	0.05	0.04	6223
50-55	0	0	0	0	0	0.01	0.01	0.02	0.04	0.17	**0.44**	0.2	0.08	0.05	0.02	3657
55-60	0	0	0	0	0	0	0	0	0.01	0.03	0.13	**0.42**	0.19	0.06	0.01	1728
60-65	0	0	0	0	0	0	0	0	0	0.01	0.02	0.1	**0.4**	0.16	0.01	693
65-70	0	0	0	0	0	0	0	0	0	0	0	0.01	0.06	**0.22**	0.01	144
>70	0	0	0	0	0	0	0	0	0	0	0	0.01	0.01	0.06	**0.03**	57
N	6430	18042	11577	11191	11835	12082	11360	10108	8440	5839	3536	1742	662	217	410	**113471**

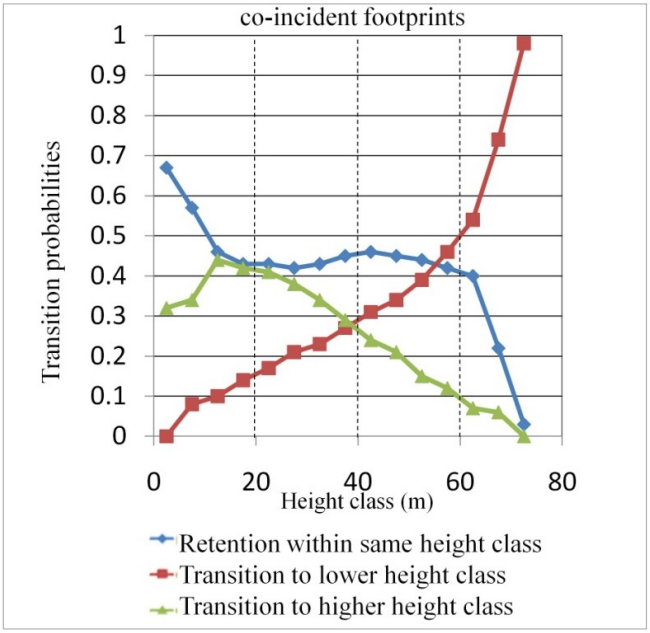

Fig. 3-9 Probability of transition to a higher height class is greater for shorter canopies, < 40m. In taller height classes (> 60m), losses increase exponentially and probability of staying in the same height class is very low.

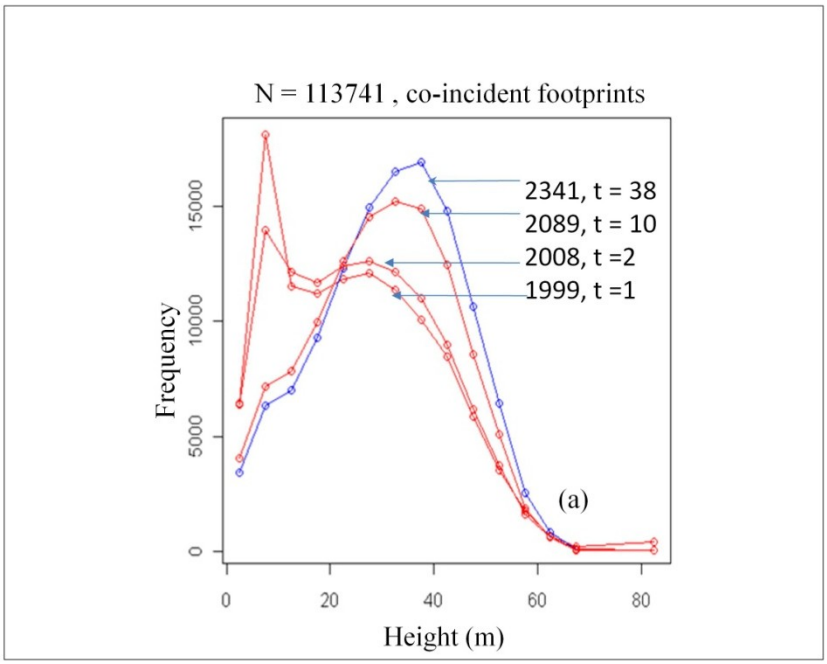

Fig. 3-10 Canopy Height (RH100) projections from the transition matrix at successive nine year time intervals (t) from 1999 to steady state in 2431 (t = 38).

3.4.3 Landscape level canopy height and biomass changes

Direct canopy height changes (0.07 ha and 1 ha)

Landscape maps of ΔRH_E100 showed many areas with losses greater than 10 m and also areas with more than 5 m growth (Fig. 3-11). The Wilcoxon-Mann-Whitney test identified significant changes in 7% of the grids at 0.07 ha scale and in 28% at hectare scale (Fig. 3-12). Spatial patterns of significant changes were similar at both scales with more than 65% being positive.

Predicted biomass changes (0.07 ha and 1 ha)

The 1999 lidar regression equation (Biomass = 0.50 * max $RH75^2$ –19.71* sdRH75 + 40) had a lower coefficient of determination (R^2 = 0.67, RMSE = 235

Mg/ha, Fig. 3-13a.), but also lower residual errors than the 2008 regression equation (Biomass = 0.81* mean $RH75^2$ +88.8* sdRH25 -58.52* mean RH25 -70.5, $R^2 = 0.72$, RMSE = 258 Mg/ha, Fig. 3-13b) and was used to predict biomass for both years. Errors were high with this equation because biomass range was as high as 3000 Mg/ha. Note that we did not use the equation derived over 17 field plots (Fig. 3-14) for this analysis because it does not cover the entire biomass range. When 95% prediction intervals were applied to predicted biomass for both years, around 33% of intervals did not overlap or were significant at 0.07 ha scale and 23% were significant at 1 ha (Fig. 3-15). At both scales, more than 65% of the significant changes were positive. Only 25% of the areas with significant height changes at 1 ha also had significant biomass changes (Fig. 3-13 & Fig. 3-15). For better clarity, 0.07 ha changes are shown only around the Teakettle region (North et al., 2002), an experimental forest of interest in the study area.

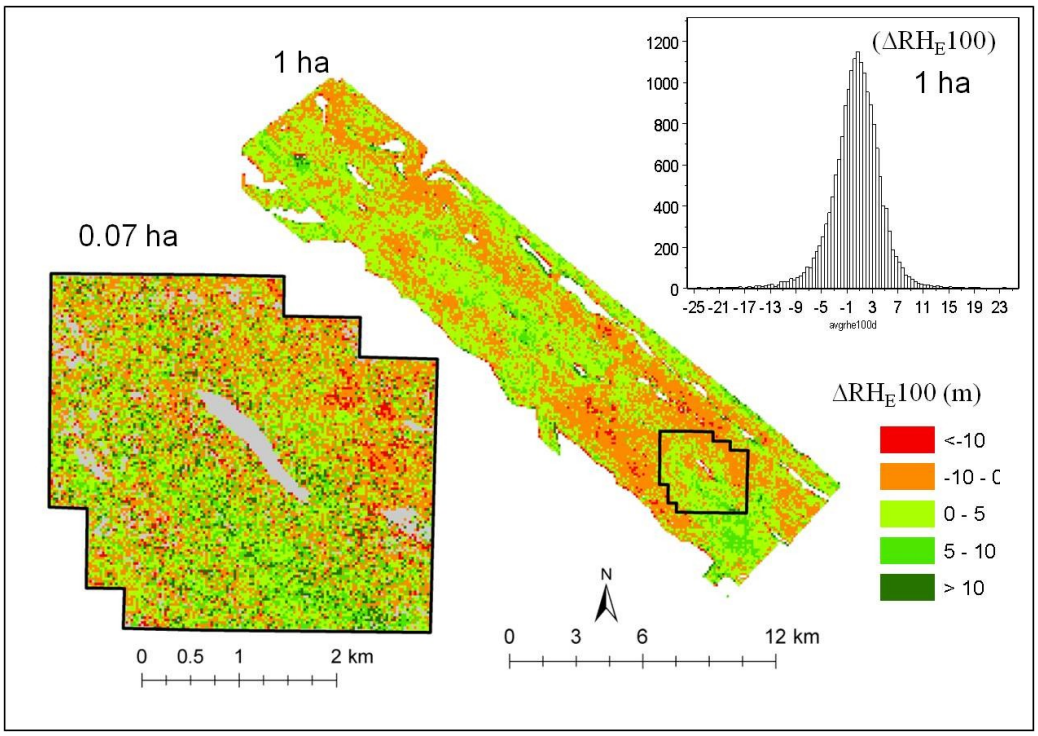

Fig. 3-11 Canopy height changes were mapped across the landscape by directly calculating average ΔRH_E100 from 1999 and 2008 LVIS data. Inset shows losses and gains at 0.07 ha scale over the Teakettle Experimental Forest

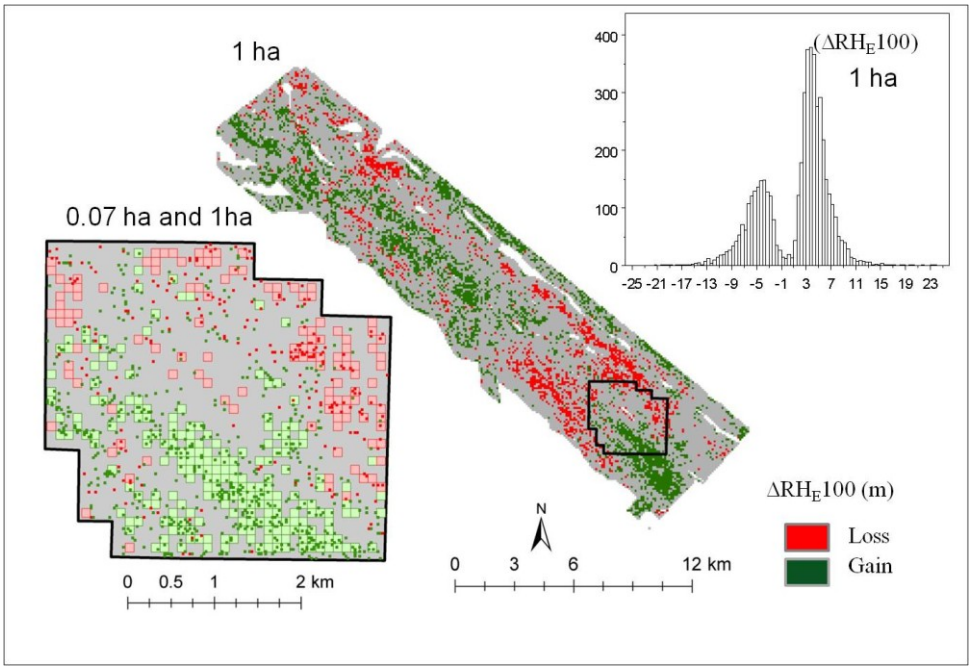

Fig. 3-12 Statistically significant RH_E100 changes were mapped across the landscape using non parametric tests. Inset shows losses and gains at both scales over the Teakettle Experimental Forest. Histogram shows net positive change.

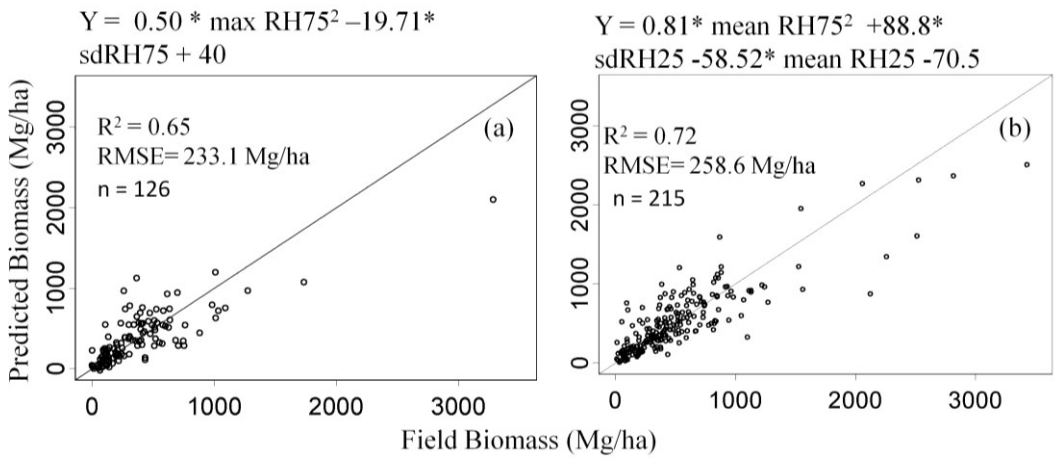

Fig. 3-13 Regression equations derived for predicting biomass from LVIS and 2000/2001 field data (a) and all 2008 subplots (b). The 1999 equation was used to predict biomass for both years.

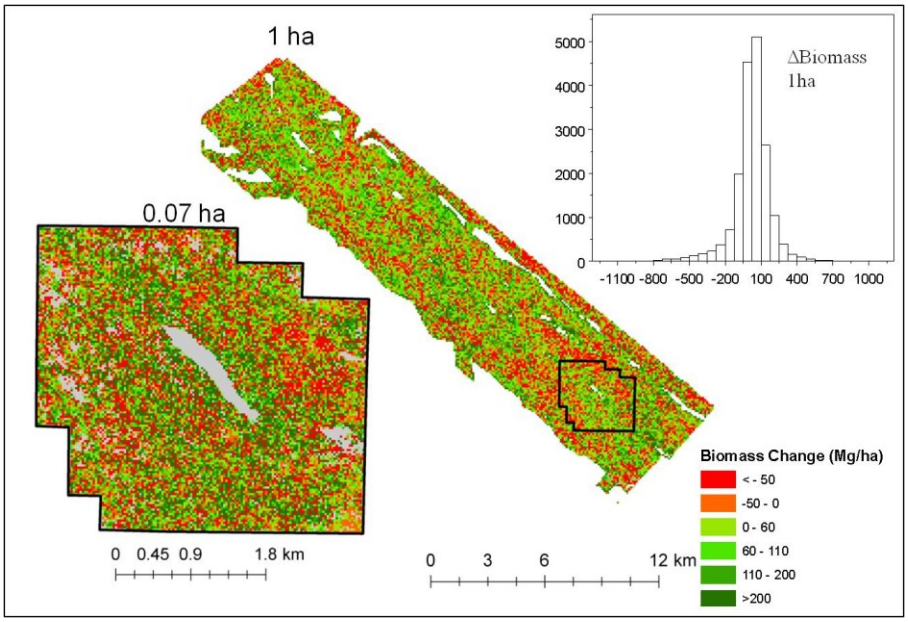

Fig. 3-14 Biomass changes predicted from LVIS (2008 – 1999) RH metrics at plot and hectare scales. Inset shows losses and gains at 0.07 ha scale over the Teakettle Experimental Forest.

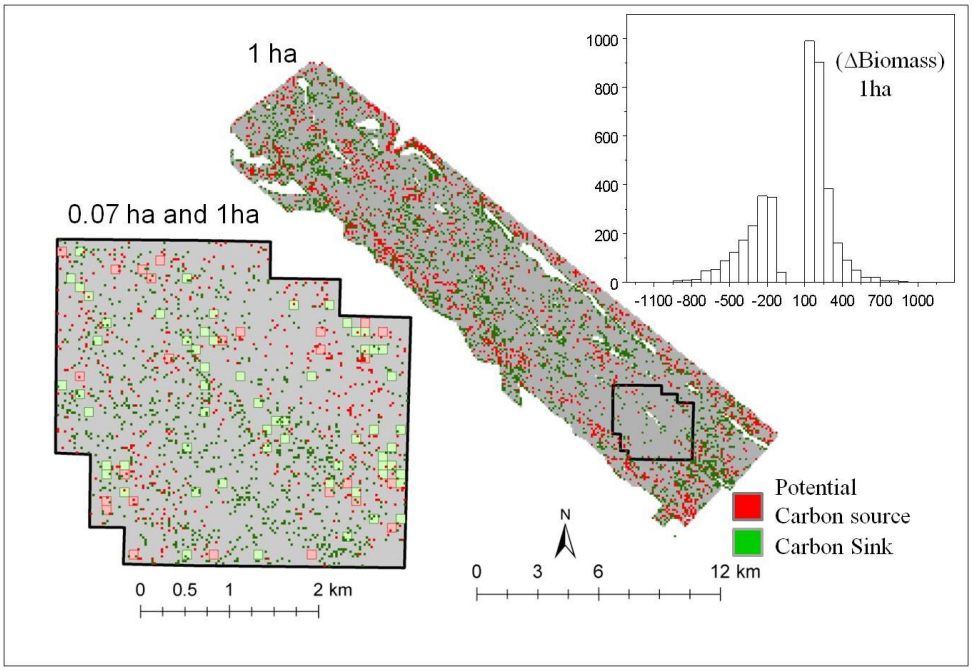

Fig. 3-15 Statistically significant biomass changes were mapped across the landscape. Inset shows losses and gains at both scales over the Teakettle Experimental Forest. Histogram shows net positive change.

Comparisons with optical imagery/landuse maps

Canopy height and biomass changes at 0.07 ha and 1 ha matched expected landuse patterns but height changes were more comparable with optical imagery and landuse maps (Fig. 3-16). A random sample of 100 cells with significant height losses at 1ha was visually compared with high resolution aerial photos for canopy cover change. Around 10% of the sample had more than 50% visually determined cover loss; 70% of the sample had visible but less than 50% cover loss and 20% did not show any visually identifiable change. In other words, 80% of the sample cells with height losses also showed canopy cover losses on aerial photos. Note that these visual estimates of changes in cover were approximate and more reliable in detecting large changes than small. Average height increases were higher in forests sensitive for

wildlife habitats (1.3 m) and slightly lower in areas under uneven aged management (0.8 m). Areas under high stress showed an overall canopy height loss (- 3.6 m) while those with low stress had an average height increase (1.17 m) (Fig. 3-17). A two sample unequal variance T-test showed that these changes were statistically significant.

3.5 Discussion

In this study, we tested the capabilities of waveform lidar in measuring and mapping forest structural changes at various spatial scales. Our major findings can be summarized as follows:

1. LVIS metrics were strongly correlated with field height and biomass for individual years, but regression equations relating changes in each could not be derived.
2. Direct comparisons between nearly coincident LVIS footprints showed overall positive changes in height after accounting for geolocation and elevation errors.
3. Canopy height projections from the footprint level transition matrix showed that the landscape was not in steady state in both 1999 and 2008.
4. Canopy height and biomass changes at 0.07 ha and 1 ha scale showed more gains than losses.
5. Statistically significant canopy height changes from LVIS were consistent with landuse and aerial photo patterns in the Sierra Nevada.

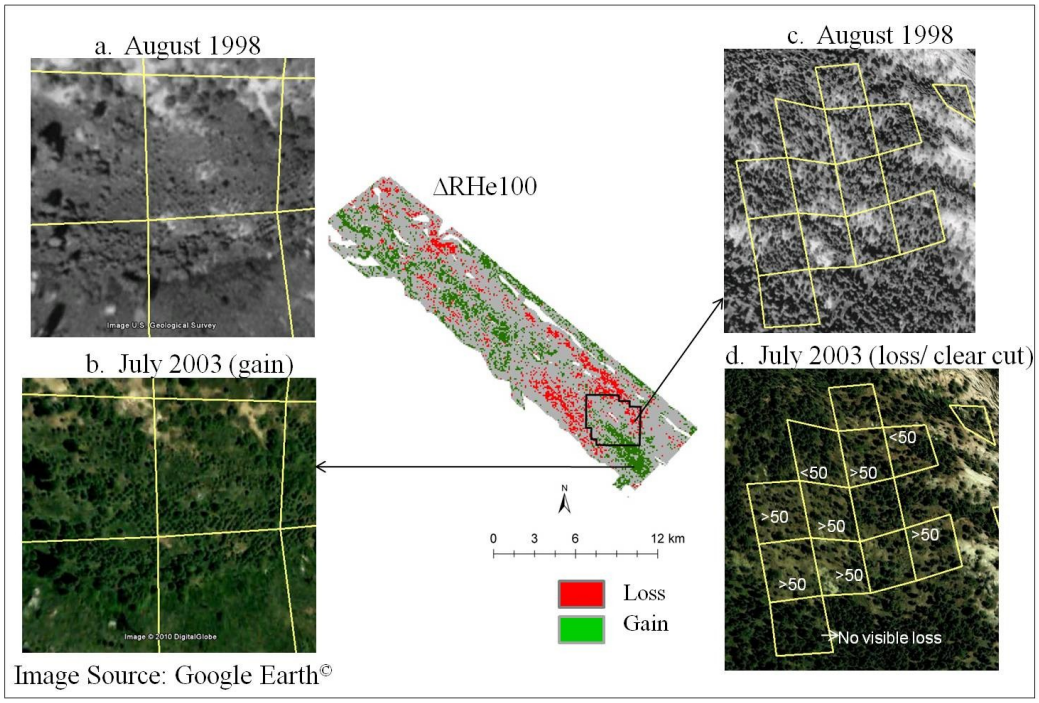

Fig. 3-16 Spatial overlay of average canopy top elevation (ΔRH_E100) change map with aerial photos showing areas with significant gain (a -> b) and significant loss (c-> d). Note that visual change in percent canopy cover was only an approximate estimate and more reliable for large changes than small.

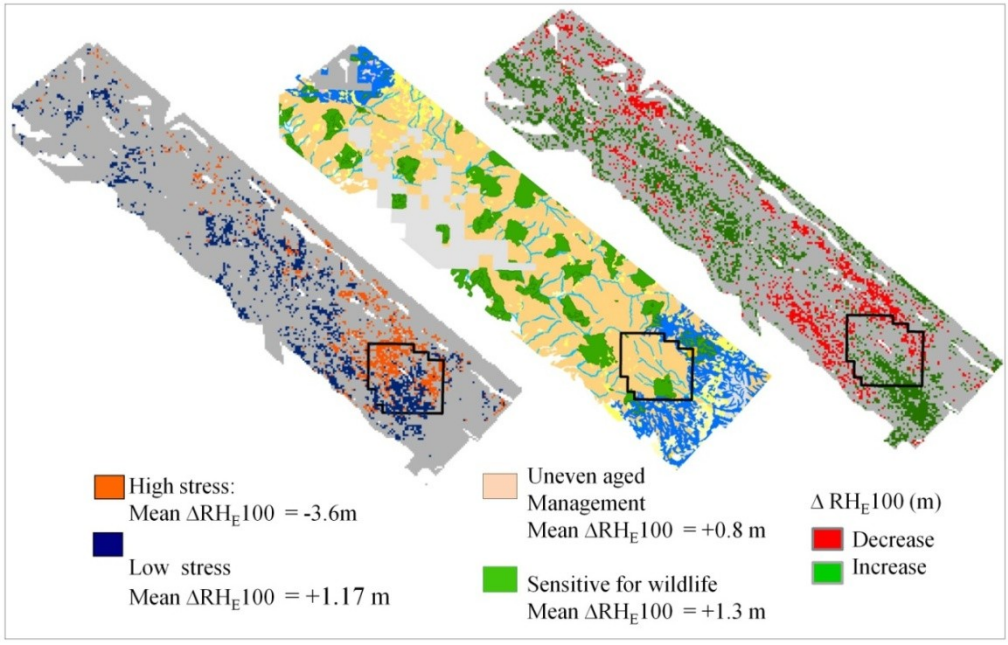

Fig. 3-17 Areas sensitive for wildlife showed overall increase in height. Forests under uneven aged management also had net height increase but less than protected areas. Low stress areas showed increase in height while high stress areas showed an overall decrease. Changes were statistically significant.

3.5.1 LVIS and field-measured changes

Despite strong correlations between LVIS and field metrics for individual years, canopy height and biomass changes observed in the field could not be predicted successfully with lidar data. This was primarily because these changes in the 17 plots did not show enough variability, since we had re-measured plots to cover the range of biomass and not biomass changes. Average changes in field heights in most plots were less than 4 m while heights predicted from lidar data had a root mean squared error of more than 4 m (Fig. 3-8a.). Similarly, biomass changes in most field plots were less than 100 Mg/ha, while prediction errors from lidar for both years were higher (Fig. 3-8b). This means that either large losses or rapid growth should have occurred for lidar to detect them or prediction errors should be reduced to less than 2 m (e.g., with small footprint data).

Large losses and gains did occur across the landscape but their range was not fully captured by the re-measured field plots. These plots were generally in experimental or old-growth forests where large changes did not occur. Almost none of them were in areas designated for commercial timber harvests. Although there were occasional signs of logging and fire, only two plots showed extensive losses (Fig. 3-8b). Increments in dbh and height in younger trees compensated for losses from mortality resulting in an overall positive change too small to be detected by lidar. One way of addressing this problem would be directly mapping large losses and gains from LVIS at landscape level and re-sampling field plots to cover the larger range of disturbance and growth.

Another factor influencing field validation might be the collocation differences between LVIS and field plots (Fig. 3-4). Since none of the lidar shots in 1999 or 2008 were coincident with plot centers, errors increased when lidar footprints were near the edge of a plot with tall trees outside it as shown previously by Hyde et al. (2005) and Anderson et al., (2006). A simulation of LVIS geolocation and canopy height errors with small footprint lidar showed that a shift of 2 m leads to an average canopy height error of 1.8 m (Tang & Dubayah, AGU 2010). Reprocessing of the 1999 LVIS data led to a geolocation shift of 3-5 m between field plot centers and lidar footprints (Swatantran et al., in press). Combining this with the relative shift of 2-2.5 m between 1999 and 2008 data could have reduced the predictive power of LVIS in plot level comparisons. Field and lidar change relationships may be stronger at the hectare scale, but they could not be tested in this study because only large trees were measured in 1999.

3.5.2 Coincident footprint dynamics

We analyzed direct canopy height changes over more than 100,000 coincident LVIS footprints by reducing geolocation and elevation errors. The large number of footprints increased the power of the paired t-test in detecting height change. Even though changes in RH metrics were small, they were statistically significant and positive, indicating overall growth.

Canopy height transitions between coincident LVIS footprints were analyzed, similar to Dubayah et al. (2010) and Kellner et al. (2009). High probability values along the diagonal suggest that a major portion of the landscape did not undergo changes of more than 5 m. Most height increases occurred in canopies less than 40 m

tall. Canopies taller than 40 m are were more likely to shift to lower height classes. Losses in taller trees were consistent with evidence of clear-cut logging on optical imagery. Beyond 60 m, canopies had very little probability of staying in the same class. Some of these changes could be waveform errors or mortality in large trees. Transitions showing height increases of more than 10 m are more likely to be waveform errors or canopy filling from adjacent areas rather than actual growth.

Canopy height distributions in 1999 were not in steady state, suggesting higher rates of disturbance prior to time period studied, consistent with Raumman and Soulard (2007). On the other hand, height distributions in 2008 were closer to equilibrium showing that the landscape was recovering from past disturbances. Although predictions of reaching canopy height equilibrium in 350 years may be unrealistic, given the frequency of catastrophic events in the Sierra Nevada, our results demonstrated the possibility of modeling future transition pathways under different forest management scenarios using temporal lidar data. Matrix models used in this study can be further modified to account for harvest, growth and mortality (Buongiorno, 1980). Other transitions in canopy cover and biomass can also be modeled.

3.5.3 Landscape changes in canopy height and biomass

The third objective of this study was to analyze spatial patterns of canopy dynamics from lidar across the landscape. Canopy height changes were calculated directly from LVIS while biomass changes were predicted using lidar and field allometry. Although the non-parametric Wilcoxon-Mann-Whitney test was more robust than a two-sample t-test in detecting significant height differences at 0.07 ha

scale, it is possible that more grid cells at this scale had Type II errors because of small sample sizes. At the hectare scale, the test had greater power because sample sizes were larger. Changes as small as 0.25 m were detected as significant in the canopy height change map at 1 ha explaining why only 25% of the significant height changes showed significant biomass changes.

Our results showed that direct estimation of biomass from temporal lidar is difficult without adequate field data, concurring with Dubayah et al. (2010). This is partly because of the high prediction errors in biomass for this study area and also because of insufficient field plots. It may be more effective in areas where the prediction errors are not as high and number of field plots is large. Integration with other remotely sensed data disturbance maps from Landsat or repeat observations from radar data may also improve estimates. Despite these drawbacks, areas that did show biomass gains and losses are likely to be sources/sinks because changes had to be more than +/-100 Mg/ha (histogram, Fig. 3-15) to be recorded as significant. The biomass change map can therefore be used to detect large potential carbon sources and sinks but not smaller changes.

Canopy height and biomass changes over more than 75% of the study area were not statistically significant, if we assume that these areas did not have actual large changes. These results are consistent with previous change detection studies with Landsat imagery over the Sierra Nevada (Raumman and Soulard, 2007). In areas where significant changes did occur, increases in height and biomass outweighed decreases suggesting overall growth.

Although our results show overall growth, spatial patterns of canopy height losses showed interesting patterns when compared with temporal aerial photos. Around 80% of the lidar height losses at 1 ha were identified as visible canopy cover losses, although these were approximate estimates. Large patches of height loss were visually identified as selective logging or clear cuts on lands outside protected areas (Fig. 3-16c, d). Smaller changes could indicate increased pest infestations (e.g. fire engraver beetle) in areas where canopy heights reduced by only a few meters (Schwilk et al., 2006). Evidence of extensive crown damage was also observed in the study area during field data collection in 2008. Forests that were reserved for habitat conservation or unsuitable for timber generally showed growth as also did patches regenerating from past disturbances (Fig. 3-16a, b).

We could not analyze small tree mortality from moisture stress (van Mantgem et al., 2007) with temporal lidar data. Footprint scale was the closest we could get to individual trees but even at this scale changes in trees with dbh less than 10 cm could not be easily detected. We did, however, detect greater height losses in high biomass stands with moisture/chlorophyll stress and growth in low stress forests. Stress maps generated from lidar and hyperspectral data in 2003 could have been early indicators of canopy height loss detected in 2008. Another interesting observation was that the low stress areas generally coincided with protected forests, which also showed canopy height increases. More research is needed to determine whether an empirical relationship between stress, protected areas and height change exists and if such areas can be mapped with lidar and hyperspectral data.

3.6 Conclusion

This study highlights some of the challenges of analysis with temporal data and field observations, each of which has inherent errors. Despite several limitations, such as inadequate field data for validation, differences in LVIS instruments between 1999 and 2008, random geolocation, elevation and sampling errors, we showed that it was possible to quantify large changes in height and biomass. It may also be possible to detect small canopy height changes with LVIS, as shown by statistical tests, but the accuracies can be known only with more validation experiments. Although large gains and losses in biomass were detected with lidar data more research is needed to improve the accuracies for mapping smaller changes at finer scales.

An important requirement for robust validation of lidar with field changes is to cover the entire range of disturbance and growth in field measurements. Maps of direct canopy height and biomass changes from lidar can be stratified into disturbance ranges, which can further be used to resample field plots and improve validation. These results could be further improved with better calibration of lidar data and more field plots for validation.

There is considerable potential in using temporal lidar data by itself or in combination with other remote sensing data for ecological applications. We demonstrated the use of transition matrices predicting future canopy height distributions with lidar data. We also combined lidar data with landuse maps and hyperspectral data to understand canopy dynamics and stress. Maps of forest structural changes can be of considerable value to forest managers in understanding how losses and gains are spatially distributed and can be extremely useful for making decisions regarding conservation and forest management.

Chapter 4 Mapping Bird Habitat Quality in New Hampshire using Radar, Lidar and Multispectral Fusion

4.1 Introduction

Habitat management efforts for wildlife species are often hindered by lack of accurate maps of both forest structural characteristics as well as species distribution. Habitat selection in wildlife species, particularly birds, is strongly influenced by vegetation structure and composition, among other factors (MacArthur & MacArthur 1961; Robinson & Holmes, 1984; and Degraaf et al., 1998) but cannot be quantified from one remote sensing system alone. With advances in remote sensing technology, newer data with complementary attributes are becoming increasingly available (Bergen et al., 2009). Multi-sensor fusion is therefore a promising approach for optimizing capabilities of different remote sensing data to improve forest structure and habitat mapping.

Optical remote sensing data have already been used extensively to map habitat preferences by relating species presence/abundance to spatial distribution of vegetation across landscapes. Yet, while vegetation characteristics such as land cover (Franklin & Wulder, 2003), phenology (Moody & Johnson, 2001) patch size, and fragmentation (Gustafen, 1998; and McDermid et al., 2005) have been mapped from optical data, measurements of vertical structure are not easily available. Light detection and ranging (Lidar) is an active remote sensing system that can provide

accurate measurements of vertical vegetation structure and is increasingly being used in ecological applications (Vierling et al., 2008).

Lidar instruments essentially record the time taken by an infrared laser pulse to reach the earth's surface or canopy top from an airplane/spacecraft and return, in order to measure ground elevation and canopy height. Depending on the area covered by the laser beam on the ground, it can be classified as a small- (< 2m), medium- (10 -30 m) or large- footprint (> 70m). Discrete return lasers record two or more returns, namely, one from the ground, one from the top of the canopy and some in between (Lefsky et al., 2002). Full waveform digitizing lidar instruments record the entire outgoing and return signal to provide a waveform with amplitudes proportional to the vertical distribution of canopy material within a footprint (Blair et al., 1999). Both small-footprint discrete-return lidar and medium-footprint waveform lidar data have been used to map forest structural characteristics such as canopy height, canopy cover and aboveground biomass, in addition to sub-canopy topography (Drake et al, 2002a; Clark et al., 2004; Popescu et al., 2004; Hyde et al., 2005; and Anderson et al., 2005). Many recent studies have also shown the potential of small- and medium-footprint lidar in mapping wildlife habitat characteristics (Bradbury et al., 2005; Nelson et al., 2005; Hinsley et al., 2006; Goetz et al., 2007; Martinuzzi et al., 2009; and Goetz et al., 2010) However, few studies have compared the relative capabilities of small- and medium-footprint lidar in mapping habitat characteristics.

Radio Detection and Ranging (radar) is also an active remote sensing system that records backscattered radiation in the microwave region of the electromagnetic spectrum. Radar sensors are more sensitive to structural and dielectric characteristics

of vegetation than optical remote sensing data, but do not provide direct measurements of structure unlike lidar. Radar data have been used to derive structural attributes (Imhoff, 1995; and Saatchi et al., 2007) and map bird habitat characteristics (Imhoff et al., 1997; and Bergen et al., 2007). Adding lidar to radar can increase accuracies for mapping vegetation structure. Fusion of the two sensors can also increase spatio-temporal coverage of forest structure, which is lower with lidar alone. There is considerable interest in combining measurements of vertical vegetation structure from lidar samples with structural attributes from radar and spectral attributes from optical remote sensing data to improve habitat mapping at larger spatial scales. However, questions remain on the efficacy of metrics that can be derived from these data, their accuracies and their combined use for mapping species habitats (Bergen et al., 2009).

Bird population trends and species preferences for vegetation structure and type have long been studied with field data in the Hubbard Brook Experimental Forest in New Hampshire. The availability of radar, lidar and multispectral remote sensing data in addition to bird data provided an opportunity to explore fusion applications for habitat mapping in the Hubbard. Our objective was to test how well each dataset predicted multi-year bird presence (hereafter, "prevalence") individually and in combination for 8 songbird species. We further analyzed the importance of predictor variables to determine which remote sensing metrics were more useful in describing bird habitat characteristics. Finally, we mapped prevalence as a measure of habitat quality across the landscape and compared spatial patterns with known habitat preferences for each species. By exploring vertical and spatial canopy structure in

detail through fusion, this research may potentially add new technical and ecological insights to the already well-studied bird species habitats in this study area.

4.2 Study Area and Data

4.2.1 Study Area

This study was conducted in the Hubbard Brook Experimental Forest (HBEF) in the White Mountains of New Hampshire, USA (Fig. 4-1). The HBEF is a bowl-shaped watershed covering an area of 3,160 ha with elevations ranging from 220 m to 1,015 m (Schwarz et al., 2001). Slopes are predominantly north-south facing with an average gradient of 16% and as high as 70%. Dominant deciduous tree species at lower elevations include beech (Fagus grandiflora) and sugar maple (Acer saccharum). At higher elevations, forests are dominated by deciduous birch (Betula sp) and conifer species such as balsam fir (Abies balsamica) and red spruce (Picea rubens). Understory vegetation includes saplings of dominant tree species as well as striped maple (Acer pensylvanicum), mountain maple (Acer spicatum), hobblebush (Viburnum alnifolium) and many herbs. The study area is a long-term ecological research (LTER) site and is representative of northern hardwood forests. Detailed site characteristics can be found in Holmes et al. (1979), Schwarz et al. (2001) and Solomonoff (2007).

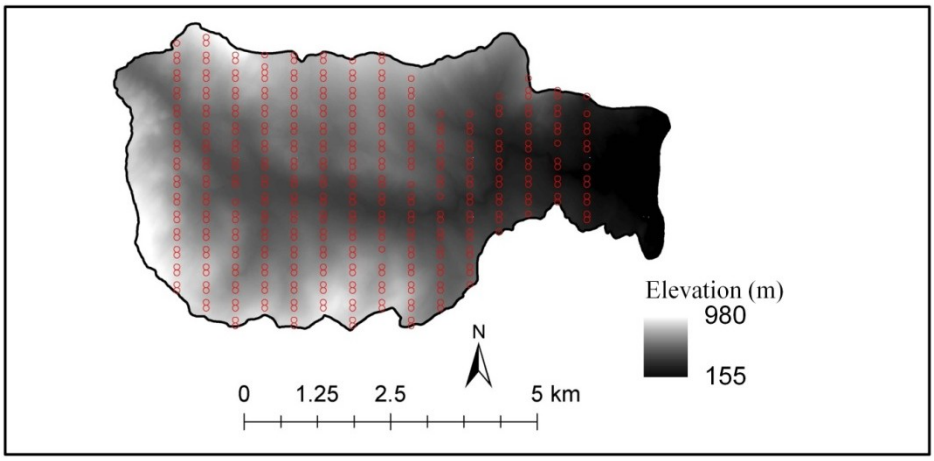

Fig. 4-1 Study area in the Hubbard Brook Experimental Forest showing plot transects laid out by Schwarz (2001). Bird point count data were collected over 371 plots by Doran et al., (2005).

4.2.2 Bird Data

Long-term bird population trends have been monitored since the 1960s at the HBEF for over 70 species. The data used in this study were collected over a period of 9 years between 1999 and 2008 (Doran et al., 2005) over a grid of 371 plots laid out by Schwarz et al. (2001). The plots are separated by 100 m or 200 m and run in north-south transects across the entire study area (Fig. 4-1). Bird sightings were recorded for 10 minutes within a radius of 50 m around each plot center (0.79 ha area) following the point count methods outlined by Ralph et al. (1995). Observations were recorded two or three times every year during the peak breeding season. Bird point counts were used to calculate presence, absence, multi-year presence (prevalence), total species richness and average abundance over the 9-year time interval. For this study, we focused on 8 migratory songbirds (Table 4-1) where preliminary analyses showed strong relationships (> 30%) between lidar metrics and variability in prevalence. These included the blackpoll warbler (BLPW), black-throated blue warbler (BTBW), magnolia warbler (MAWA), yellow-rumped warbler (MYWA),

ovenbird (OVEN), red-eyed vireo (REVI), dark-eyed junco (SCJU) and the yellow-bellied flycatcher (YBFL). A detailed description of habitat characteristics of these birds and data collection can be found in Doran et al. (2003). Most of the other species had very low values of prevalence and large number absences. We restricted the number of species modeled to 8 in this study to avoid errors because of zero – inflated data. These species have also been studied more extensively for their habitat preferences (Doran et al., 2003) than the birds that had very low prevalence allowing for a better comparison with remote sensing variables.

Table 4-1 Common and scientific names for songbird species

Common Name	Code	Scientific Name
Blackpoll Warbler	BLPW	*Dendroica striata*
Black Throated Blue Warbler	BTBW	*Dendroica caerulescens*
Magnolia Warbler	MAWA	*Dendroica magnolia*
Yellow Rumped Warbler	MYWA	*Dendroica coronata*
Ovenbird	OVEN	*Seiurus aurocapillus*
Red Eyed Vireo	REVI	*Vireo olivaceus*
Dark Eyed Junco	SCJU	*Junco hyemalis*
Yellow Bellied Flycatcher	YBFL	*Empidonax flaviventris*

4.2.3 Radar Data

The Uninhabited Aerial Vehicle Synthetic Aperture Radar (UAVSAR) is an L-band radar developed at the Jet Propulsion Laboratory (Rosen et al., 2006) that records backscattered energy from the earth's surface in the microwave region of the electromagnetic spectrum (23 cm). UAVSAR has full polarimetric capabilities, i.e., it can record four combinations of transmitted and received polarizations (Fig. 4-2). Co-polarized bands include horizontal transmitted and horizontal received polarization (HH) as well as vertical transmitted and vertical received (VV). Cross-polarized

bands include horizontal transmitted and vertical received polarization (HV) and vice versa. Studies have suggested that fully polarimetric data are useful for studying surface and volume scattering from vegetation and are more sensitive to structural properties such as Leaf Area Index (LAI), basal area and biomass than single polarization (Imhoff, 1995; Balzter, 2001; and Treuhaft et al., 2004). L-band has greater penetrating power and is more sensitive to tree trunks and vertical structure than smaller wavelengths (Imhoff, 1995).

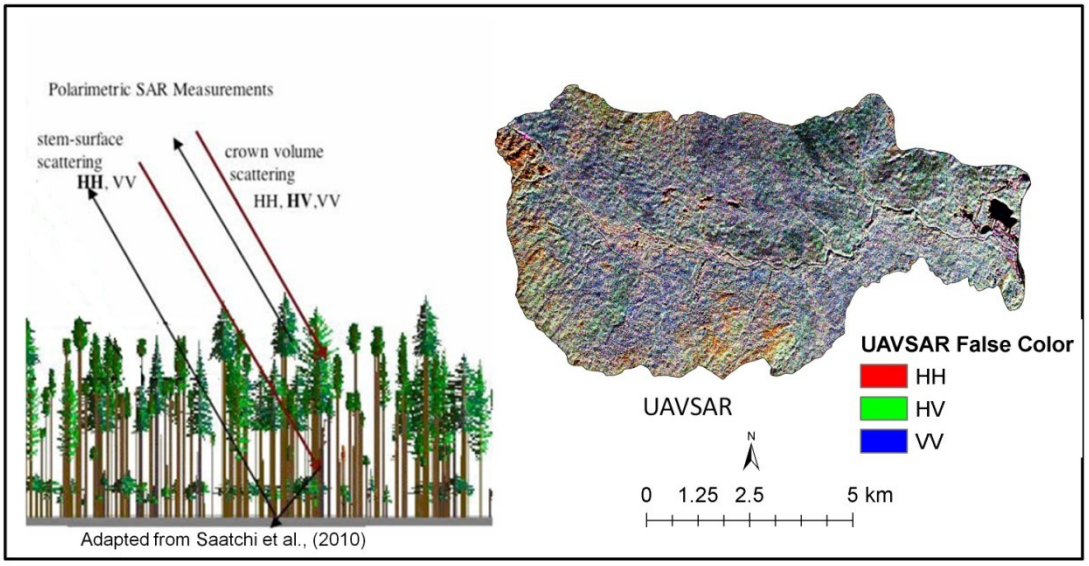

Fig. 4-2 UAVSAR backscatter images from horizontal transmitted horizontal return (HH), vertical transmitted vertical return (VV) and Horizontal transmitted vertical return (HV) polarizations.

In this study, we used co-polarized HH, VV and cross-polarized HV data from the L-band radar (Fig. 4-2). Raw data were processed into backscatter images with 5 m nominal spatial resolution, orthorectified with digital elevation models and corrected for slope at the Jet Propulsion Laboratory. We converted backscatter values into power, applied a (3*3) gamma filer to reduce speckle and calculated average

statistics for HH, VV and HV bands within bird plots. In addition, band ratios HH/VV, HV/VV HV/HH and normalized difference band ratios [HH-VV/HH+VV] (Simental et al., 2005), [VV-HV/VV+HV] and [HH-HV/HH+HV] (Saatchi et al., 2010) were also calculated to increase sensitivity to structure. Backscatter ratios and indices have been found to be sensitive to biomass (Saatchi et al., 2010), canopy height, basal area (Imhoff et al., 1997) and other canopy biophysical variables. In this study, we did not derive biomass from radar because of difficulties in biomass estimation from lidar and radar in the HBEF. Instead, we used the ratios directly because of their sensitivity to vegetation structure.

4.2.4 Small-Footprint lidar Data

Small-footprint discrete return lidar (DRL) data were collected over the study area in September 2009 by the Canaan Valley Institute using the Optech ALTM 3100 instrument with at least one laser shot per square meter (but on average > 5 shots) and four vertical returns. The laser ranging data were geolocated with GPS and inertial navigation units and interpolated to create a digital surface model with a resolution of 0.5 m. A three dimensional canopy height model (CHM) with a horizontal resolution of 0.5 m and vertical resolution of 0.15 m was derived by subtracting ground elevation from the digital surface model. The high spatial resolution of the CHM made it possible to delineate dominant and co-dominant tree crowns. We used an adaptive 'local maxima' filtering algorithm (TreeVaW) developed by Popescu et al. (2004) to identify individual treetops from the CHM and obtain crown radii and height (Fig. 4-3). The algorithm was calibrated using field measurements of canopy height and crown radii collected over the HBEF in 2009. Details of the TreeVaW

algorithm are available in Popescu et al. (2004) and Popescu et al. (2007). We calculated summary statistics for crown diameter and individual tree heights within bird plots and also crown weighted height for each plot using the following equation:

$$Cwght = \frac{\Sigma\,(\text{crown diameter} * \text{height})}{\Sigma\,(\text{crown diameter})}$$, where Cwght = crown weighted height. This metric is similar to Lorey's (basal area weighted) height from field data.

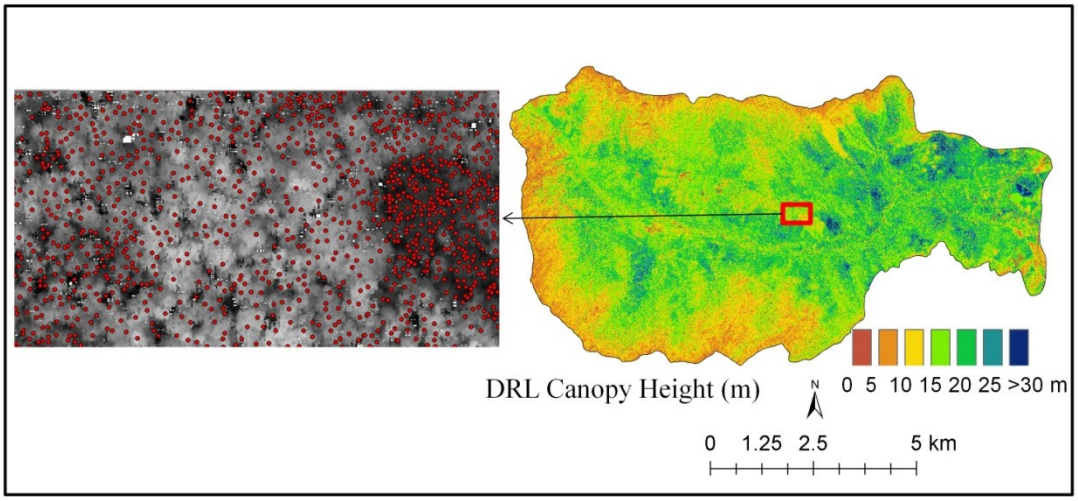

Fig. 4-3 Discrete return lidar canopy height map showing individual tree locations (inset) detected by TreeVaW. Crown radius and height are calculated for each tree by the algorithm.

4.2.5 Medium-footprint lidar data

The Laser Vegetation Imaging sensor (LVIS) is a medium-footprint (25 m diameter), full-waveform digitizing lidar designed and developed at NASA's Goddard Space Flight Center. (Blair et al., 1999) LVIS data were acquired over New Hampshire in the summer of 2009 with trees in leaf-on condition. The waveforms were geolocated with respect to the WGS 84 ellipsoid. Canopy top was detected by finding the lidar return greater than the noise threshold at the top of the waveform (Fig. 4-4). Because of occasional large ground-finding errors in this study area where

LVIS ground elevations were determined incorrectly, we substituted LVIS elevations with those from discrete return lidar data. Canopy height (RH100) was calculated by subtracting the average DRL ground elevation within each LVIS footprint from the canopy top. Other metrics such as heights of 25% (RH25), 50% (RH50) and 75% (RH75) energy returns were calculated in a similar manner (Swatantran et al., in press). Total canopy cover was calculated from the normalized cumulative energy return following methods in Ni Meister et al. (2001). Additionally, canopy cover was calculated from the cumulative profile at every 5 m interval between ground and 40 m, resulting in 8 metrics. Metrics showing variations in amount of foliage at different levels within the canopy possibly could explain bird occurrence/prevalence better than total canopy cover alone. Summary statistics for canopy cover metrics and LVIS RH metrics were calculated within bird plots.

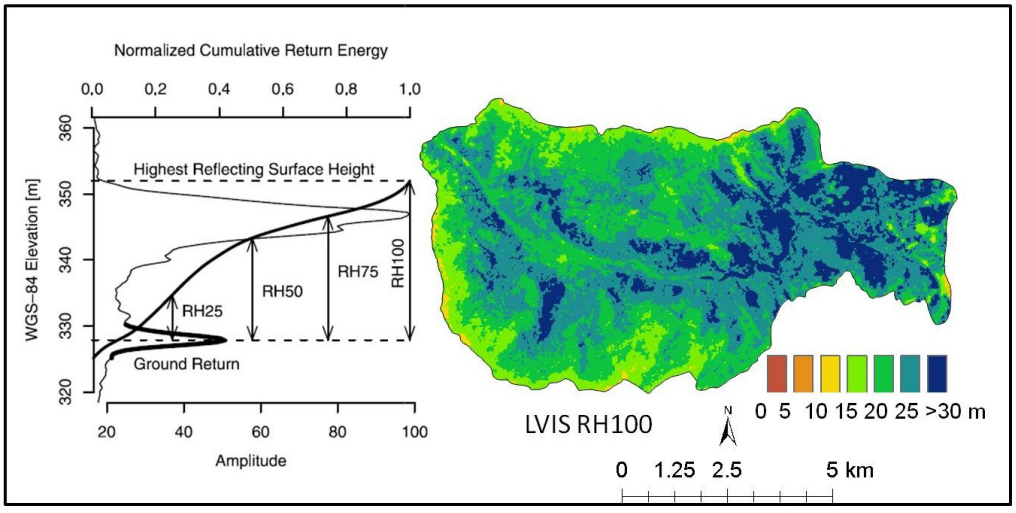

Fig. 4-4 Components of an LVIS waveform (left) and canopy top height map showing vertical and spatial distribution of heights. Heights of 25%, 50%, and 75% lidar energy returns (RH metrics) are calculated from the waveform. The bold solid line shows the normalized cumulative energy return from which canopy cover at every 5m interval was calculated.

4.2.6 Landsat Data

Landsat ETM images acquired in August 1999 and late October 2000 were corrected for Earth-Sun distances and solar zenith angle variations, converted into top-of-atmosphere reflectance and georeferenced (Goetz et al., 1997). The Normalized Difference Vegetation Index (NDVI) was calculated for both images. In this study, we used the NDVI as a measure of greenness and the difference between NDVI from leaf-on and leaf-off seasons as a measure of deciduousness (Fig. 4-5). Goetz et al. (2010) showed the importance of these variables in determining habitat quality for the Black-throated Blue Warbler. In this study, we tested the use of these metrics for 8 species including the BTBW.

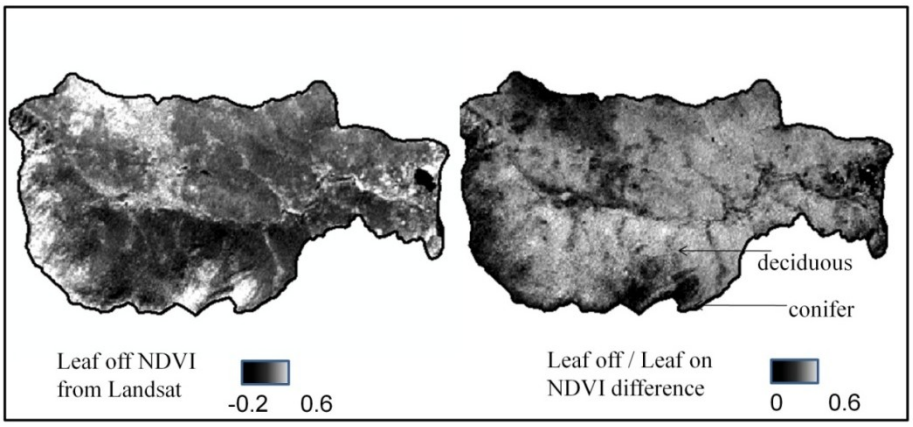

Fig. 4-5 Normalized difference vegetation index from Landsat data (leaf off) and seasonal NDVI change between from leaf off and leaf on data. High NDVI difference values show deciduous cover, low values show conifers.

4.3 Methods

All datasets were brought into a common frame of reference using the UTM 19N projection and the WGS 84 ellipsoid. Summary statistics (Table 4-2) were calculated for all metrics from radar, lidar and Landsat data within bird plots. Bird

prevalence was calculated as the total number of years a bird was sighted in a plot out of the 9 years observed, i.e., the lowest prevalence was zero and maximum prevalence was 9. Since bird observations were recorded many times a year, prevalence was treated as a continuous variable rather than categorical. Statistical models for predicting prevalence were tested for all bird species using radar, LVIS, DRL and Landsat metrics individually and in combination with each other. IDL was used for processing LVIS waveforms, ArcGIS v 9.3 for spatial analyses and packages in R (R Development Core Team, 2009) for statistical analyses.

Table 4-2 Metrics calculated from the different datasets.

Metrics used	(Min,Max,Mean,Std.deviation)
Radar (UAVSAR)	HV backscatter
	HH backscatter
	VV backscatter
	HH/VV ratio
	HV/VV ratio
	HV/HH ratio
	HH-VV/HH+VV index
	HH-HV/HH+HV index
	VV-HV/VV+HV index
Multispectral (Landsat)	NDVI (only Mean,Std.deviation)
	NDVI change (only Mean,Std.deviation)
Medium footprint lidar (LVIS)	Elevation (only Mean,Std.deviation)
	RH metrics (RH25, RH50,RH75,RH100)
	Total Canopy Cover
	Canopy cover at 5 m intervals between 0 – 35 m
Small footprint lidar (DRL)	Height of individual trees
	Crown diameter s of individual trees,
	Sum of crown diameters of trees identified within bird plots
	Crown weighted height
	Product of height and crown diameter

4.3.1 Predicting prevalence

Many models have been used to derive empirical relationships between field measures of habitat characteristics and remote sensing observations (Guisan & Zimmerman, 2000). Commonly used statistical methods include least square regression for Gaussian distribution, logistic regression for binary data and Generalized Linear Models (GLM), or Generalized Additive Models (GAM) for point count data. Recent studies have shown that machine learning algorithms, such as Classification and Regression Trees (CART), have more advantages over other methods because they do not make any assumptions about the relationships between the explanatory and response variables (De'ath and Fabricius, 2000). Regression trees are constructed by partitioning the data into two homogenous sets based on the best explanatory variable. The binary tree is further subdivided using decision rules and the terminal node provides a mean value for a response variable.

Random Forests, RF (Breiman, 2001; Breiman and Cutler, 2003) is a data mining method in which a large number of such regression trees are fit to a dataset (~ 800). Bootstrap samples are used from the data to construct each tree and at each node, a random subset of predictors are tested, hence, the name 'Random Forests' (Prasad et al., 2006). Response values from all trees are averaged to provide accurate predictions. Around 37% of the data are retained and "out-of-bag" error estimates (Breiman, 2001; Berk, 2008) are calculated using each regression tree, thus avoiding over fitting and eliminating the need for cross validation. Predictions from RF regression are more accurate than other methods and can be used to model linear/non-linear relationships using a large number of predictor variables (Cutler et al., 2007).

Recent studies have demonstrated applications of Random Forests in mapping presence, absence (Cutler et al., 2007; and Magness et al., 2008) and habitat quality (Goetz et al., 2010).

We constructed RF models using radar, lidar, and Landsat metrics individually and in combination to predict prevalence for 8 species, resulting in a total of 40 models. We compared the decrease in mean residual error with increasing the number of trees from 100 to 8,000 and found that 800 trees gave the best predictions. Growing more than 800 trees did not improve predictive power for any species. By default, the RF algorithm sampled one-third of the total variables at each node split, which was not modified in this study. Accuracies were assessed by percent variance in bird prevalence explained by each model. Variance explained, also known as 'pseudo r-squared' was calculated by the RF algorithm from out-of-bag estimates for each tree and averaged across all trees to give a cross-validated value, similar to 'r-squared' from classical regression models (Breiman, 2001).

4.3.2 Variable importance

In addition to pseudo r-squared, RF also gives a measure of variable importance calculated as decrease in accuracies on removing a predictor variable. A large decrease in model accuracy on removal of a variable indicates high importance (Breiman, 2001). We compared important variables from RF models to known habitat preferences for the birds from previous studies (Doran et al., 2003). The frequency of occurrence of a radar, lidar or Landsat metric within the 10 most important predictors for the 8 bird species was recorded to determine which variables were selected more often than others.

4.3.3 Prevalence at plot and landscape scales

Histograms of prevalence predicted from fusion were compared to observed values at plot locations. We then mapped prevalence across the landscape for the 8 bird species using the models with the lowest accuracy (radar variables alone) and the ones with the highest accuracy (from fusion). Prevalence values were classified into 4 indicators of habitat quality. Prevalence of less than 2 years was considered as low quality, between 2 and 4 years was medium, 4 to 6 years was considered as good habitat and more than 6 years' prevalence was classified as excellent habitat quality. Spatial distributions of habitat quality were compared with known habitat preferences of the individual bird species.

RF models only retain the mean observations at each node in a regression tree and do not take into account the entire distribution of the predicted values. In contrast, quantile regression forests (QRF) keep the values of all observations allowing for the construction of prediction intervals for the observed values (Meinshausen, 2006). We predicted the best possible (90%) and worst possible (10%) habitat quality for the Black-throated Blue Warbler as a test case using QRF in addition to mean predictions from fusion and compared the three maps spatially.

4.4 Results

4.4.1 Predictive capabilities of different datasets for prevalence

Predictive power for each dataset individually and with all metrics combined for 8 species are summarized in Fig 4-6. Radar metrics alone explained more than 30% variability in prevalence for all the 8 species studied. The highest variance explained was for the Magnolia Warbler (50%) while lowest was for the Black-

throated Blue Warbler (33%). Landsat data performed better than radar for all species. In the case of the Magnolia Warbler, Landsat metrics alone accounted for 64% variability, similar to the results from lidar. For the other species, Landsat metrics explained between 35% and 50% variance in prevalence. LVIS metrics predicted more than 50% variance in prevalence for all species, except the Black-throated Blue Warbler. Fusion improved the predictive power of radar metrics by 25% (Fig 4-7), Landsat metrics by 15% and lidar metrics by 5% on an average, although increase in predictive power varied between species. Results from DRL metrics were similar to LVIS for most species. In general, predictive power was lowest for the Black-throated Blue Warbler.

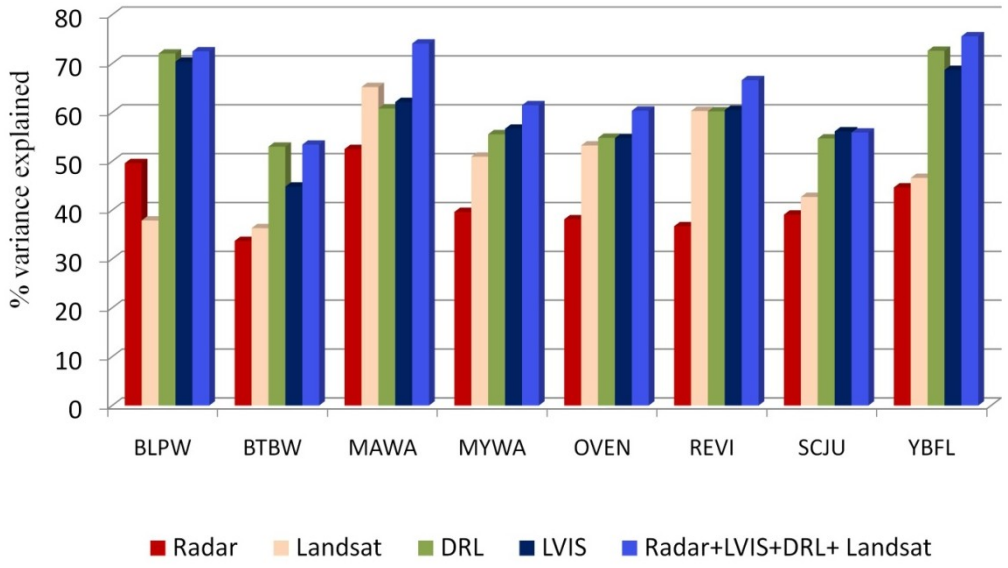

Fig. 4-6 Random Forests regression results for bird prevalence using radar, lidar and Landsat data individually and in combination with each other.

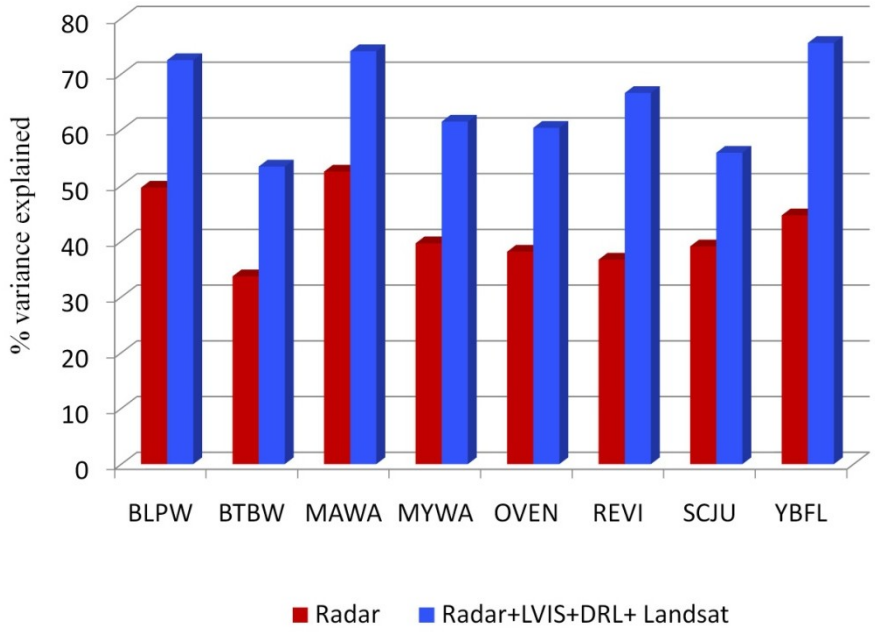

Fig. 4-7 Comparison of variance explained from models using radar metrics alone with those using all metrics.

4.4.2 Variable importance in predictions

Radar metrics were rarely selected within the 10 most important predictors when used in combination with LVIS and DRL (Fig. 4-8). When selected, co-polarized backscatter ratios, particularly the HH/VV ratio, the (HH-VV/HH+VV) index and the HV/VV metric were more important for predictions than other radar metrics. The NDVI difference or the measure of deciduousness from Landsat data was an important predictor for 7 out of the 8 species studied. Total crown diameter, crown weighted height and other crown diameter metrics from discrete return lidar were also selected in almost every model. Ground elevation was an important variable for 6 out of 8 species. LVIS RH metrics were found to be more useful than most canopy cover variables. Canopy cover at 20-25 m was selected for 5 out of the 8 species. Canopy cover metrics at lower heights were never within the 10 important variables for any species.

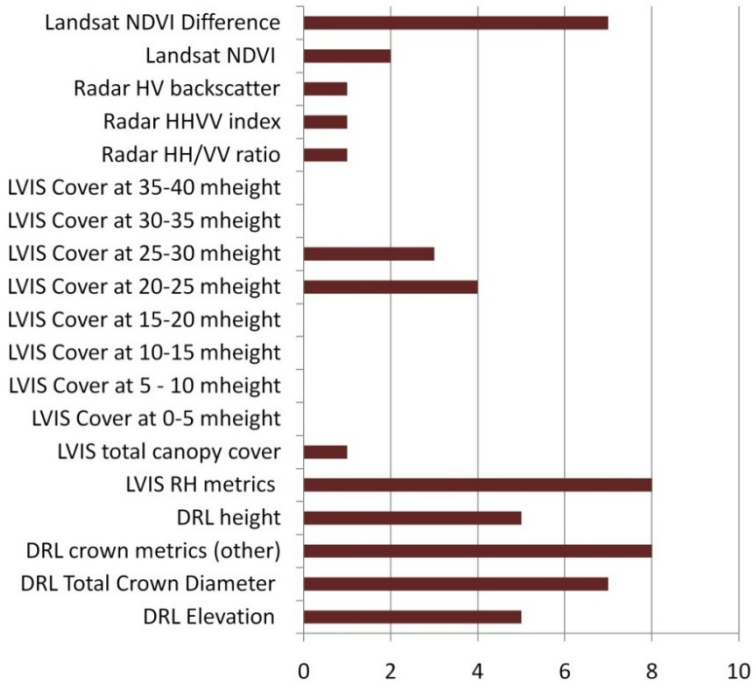

Fig. 4-8 Variable selection within 10 most important predictors for all species taken together. Model used : UAVSAR +LVIS +DRL + Landsat.

4.4.3 Predictions at plot and landscape scale

At plot scale, histograms of actual and predicted prevalence showed many similarities and differences (Fig.4-9 & Fig. 4-10). There were no negative values or predictions greater than 9 years, however, RF models overestimated predictions in some ranges of prevalence and underestimated it in others. Predictions were underestimated when actual prevalence was more than 7 years. Absences were more accurately predicted than prevalence.

Landscape patterns of prevalence values from both radar data and fusion were similar but showed wide inter-specific spatial variations (Fig. 4-11. & Fig. 4-12). There was consistency between known habitat preferences of birds and the spatial distribution of predicted habitat quality across the landscape. For example, the

Blackpoll Warbler is known to prefer conifer vegetation at high elevation with lower canopy heights and cover (Doran, 2003). The spatial pattern of predicted prevalence from both radar and fusion showed highest habitat quality for the BLPW at the corners of the study areas that occur at high elevations. On the other hand, the Black-throated Blue Warbler prefers deciduous vegetation at lower elevations with high canopy cover and well-developed understory (Doran et al., 2005; and Goetz et al., 2010). Spatial patterns of prevalence show that best habitat quality for the BTBW was at lower elevations near the valley where canopy height, cover and deciduousness were high. Although spatial patterns of predictions from radar variables were similar to those from fusion (Fig 4), all radar models consistently underestimated prevalence and had more false presences in areas of actual absence.

The 90% and 10% predictions from quantile regression forests were very different from the 50% predictions for the Black-throated Blue Warbler. The 90% prediction map or optimistic estimate showed excellent habitat quality or prevalence greater than 6 years over almost the entire study area. On the other hand, the 10% quantile or the conservative estimate showed little or no areas as excellent habitat and more medium- or good-quality habitat, i.e., between 2 and 6 years prevalence (Fig. 4-13).

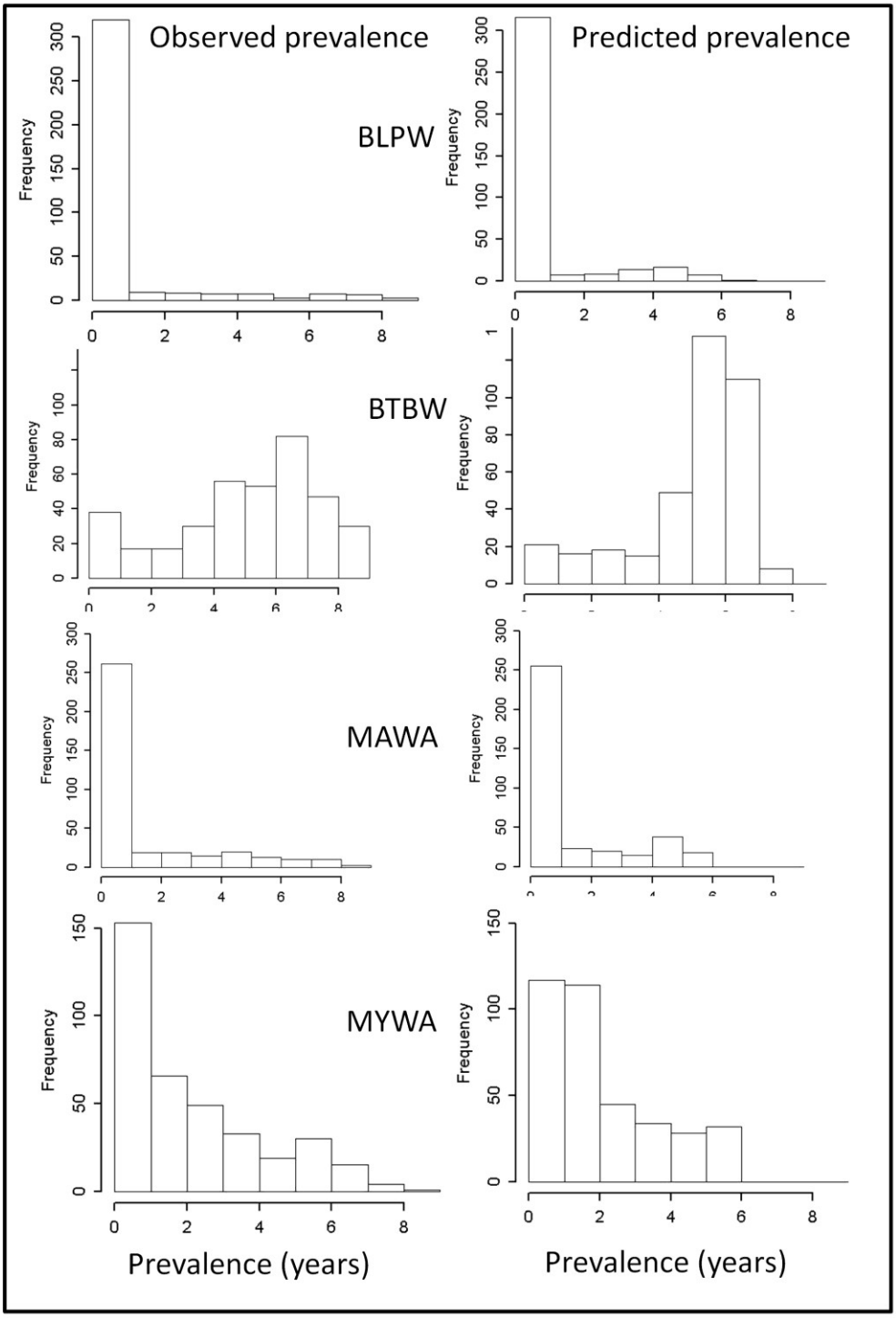

Fig. 4-9 Histograms of actual (left) versus predicted prevalence from models with all metrics combined.(right) . For species codes refer to Table 4-1.

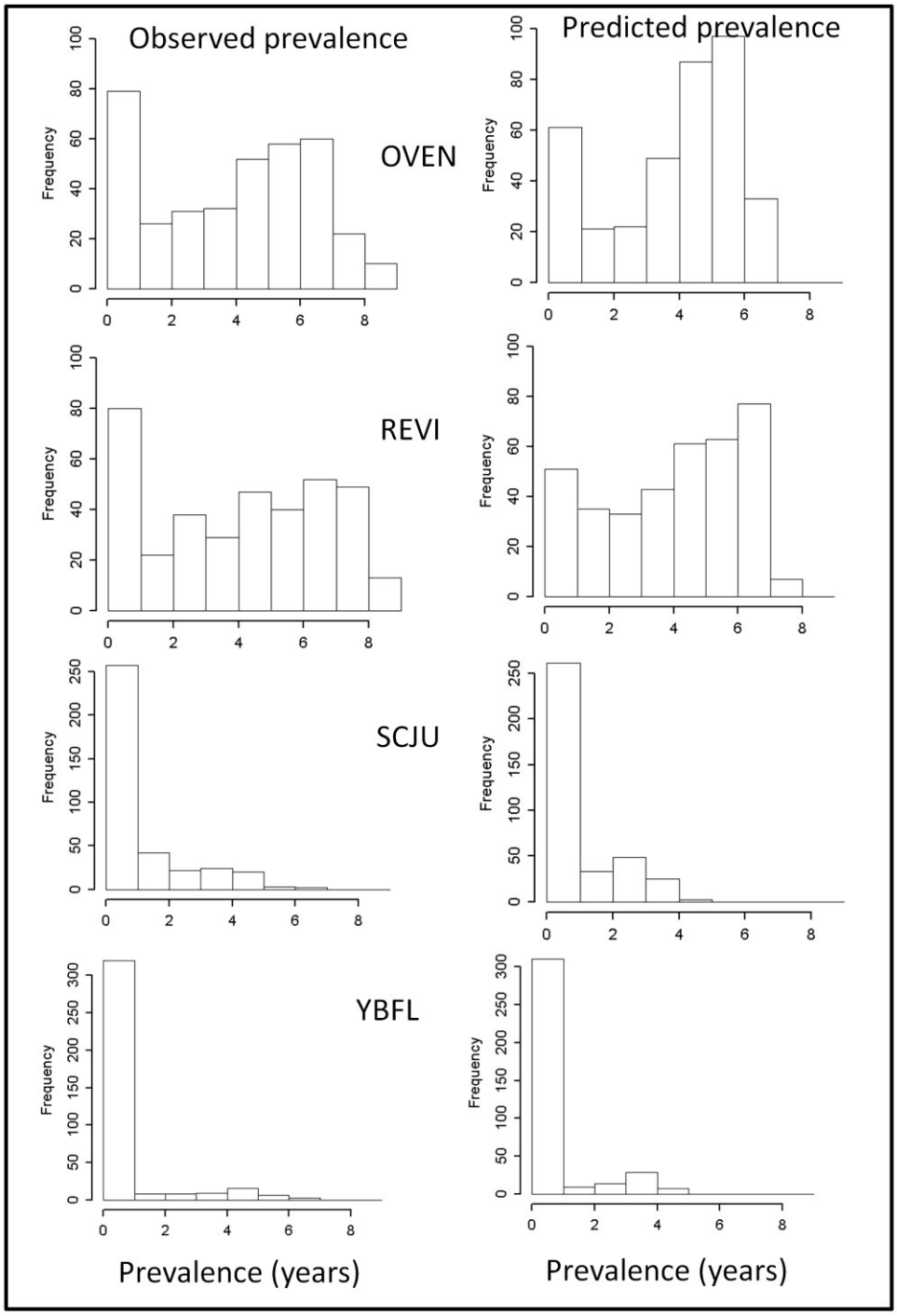

Fig. 4-10 Histograms of actual (left) versus predicted prevalence from models with all metrics combined (right). For species codes refer to Table 4-1.

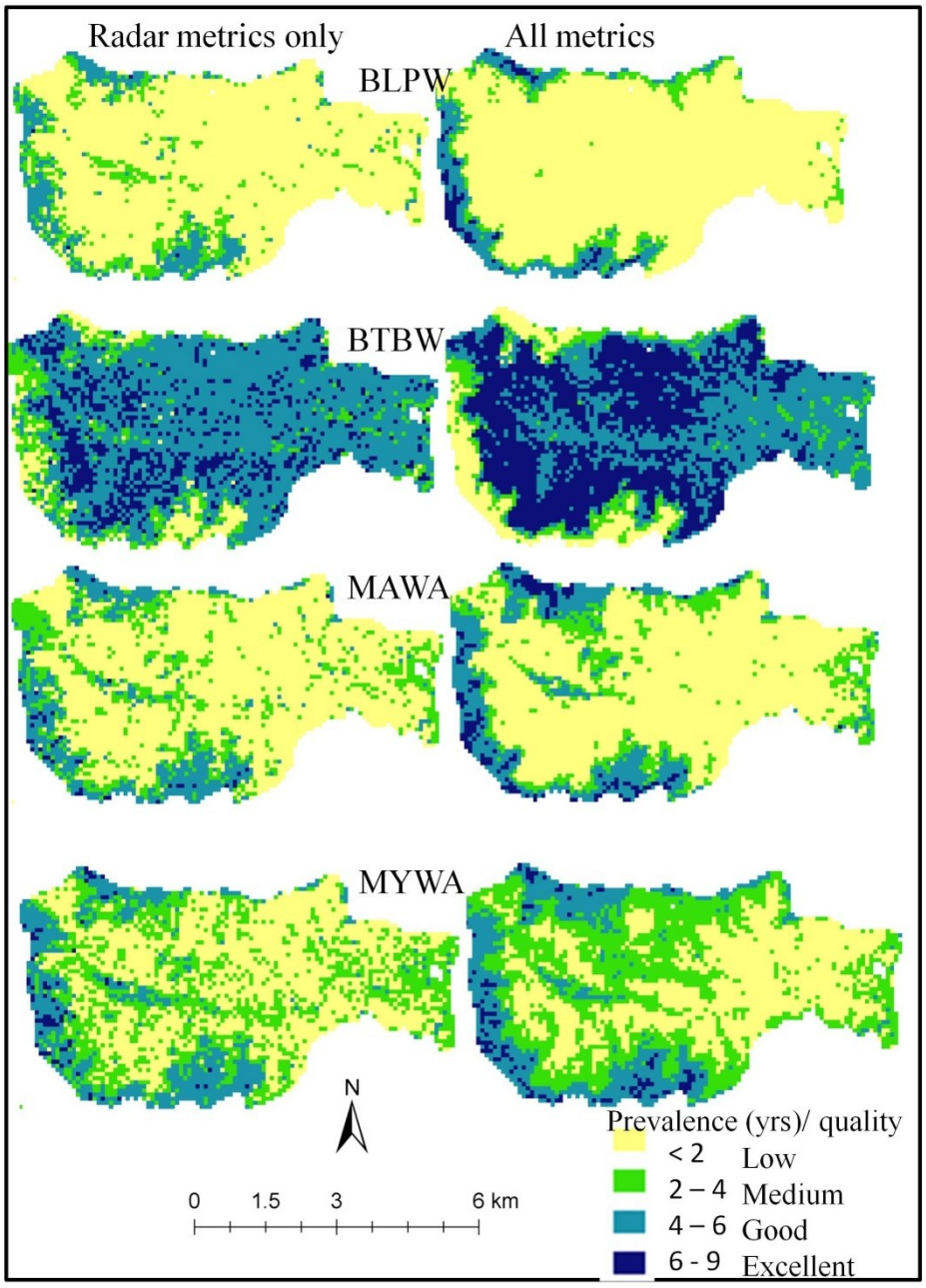

Fig. 4-11 Predicted prevalence from radar metrics alone (left) and all metrics together (right) . For species codes refer to Table 4-1.

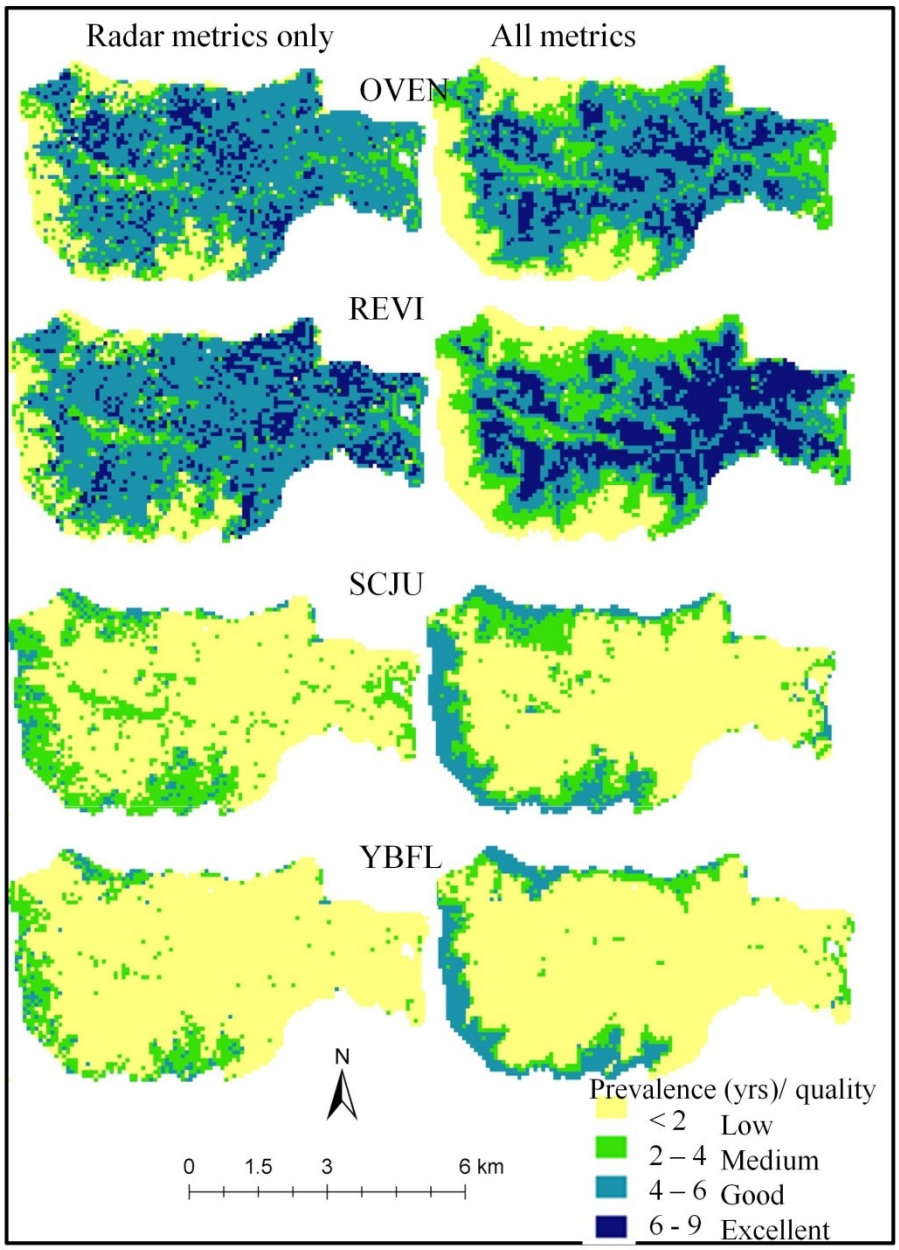

Fig. 4-12 Predicted prevalence from radar metrics alone (left) and all metrics together (right). For species codes refer to Table 4-1.

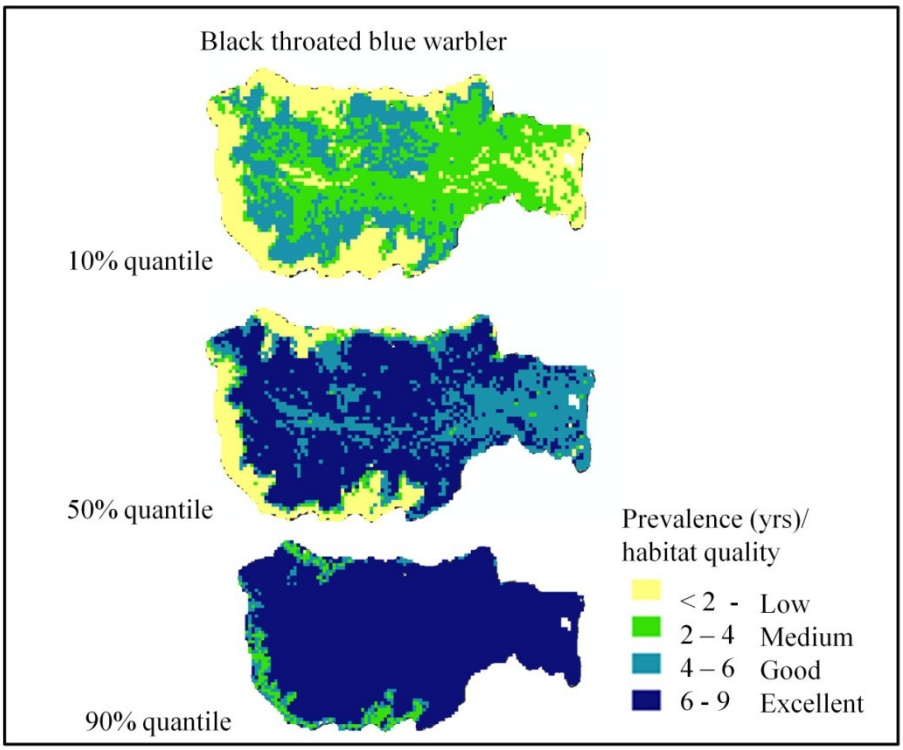

Fig. 4-13 Quantile predictions of habitat quality using Random Forest Regression.

4.5 Discussion

4.5.1 Predictive capabilities of different datasets for prevalence

Previous studies by Imhoff et al. (1997) and Bergen et al. (2007) have shown the usefulness of radar data with different wavelengths and polarizations in mapping bird habitat characteristics. Our results also showed that radar backscatter data could explain more than 30% variability in bird prevalence and in some cases as high as 50%. This was because of the sensitivity of the metrics to vegetation structure, since radar HH/VV ratios were also moderately correlated with lidar heights, canopy cover and crown diameter.

Landsat metrics performed better than radar data for all species, even though radar is known to be more sensitive to structure than optical data (Imhoff et al., 1995; and Treuhaft et al., 2004). One reason for this could be the use of the NDVI seasonal change, which not only differentiated between conifer and deciduous cover, but also correlated canopy height. Interestingly, Landsat metrics alone were as good as lidar for the Ovenbird and even better than lidar for the Magnolia Warbler. This implies that deciduous cover was a more important explanatory variable than structure for these birds.

Our results showed that both LVIS and DRL metrics performed equally for most species, explaining more than 50% in prevalence because LVIS RH metrics (particularly RH75) were strongly correlated with crown diameter/crown weighted height from DRL data. Although LVIS and DRL had similar predictive power, fewer variables from DRL data were sufficient to explain the same variation than with more variables from LVIS. In addition, the finer spatial scales of DRL data could complement the larger scale LVIS data if information such as patch characteristics, understory shrub density, and other variables are derived from DRL. Conversely, the broader mapping capabilities of waveform lidar such as LVIS as compared to DRL are a large advantage for landscape scale mapping.

Bergen et al., (2007) derived biomass from radar data and found that it improved predictions of bird species distributions when added to Landsat data. In this study, biomass was not used as a predictor; instead a suite of crown, height and canopy cover metrics from lidar were combined with Landsat and radar data to improve predictive power. Fusion improved accuracies by an average of 25% for

radar and 15 % for Landsat. These results show that multi-sensor fusion could improve results from either sensor alone. It is furthermore likely that wall-to-wall lidar mapping is not required; limited samples or transects of lidar may be sufficient to increase the predictive power of radar and/or Landsat, but this needs more research.

Goetz et al. (2010) tested the combined potential of lidar and Landsat data in mapping prevalence for the Black-throated Blue Warbler. Our results show a 10% improvement in variance explained with radar, LVIS, DRL and Landsat combined. We speculate the increase is because of addition of small-footprint lidar metrics. Despite this improvement, the variance explained for the BTBW by all models was the lowest among the 8 species studied. It is possible that the predictive power for other species was inflated because of more zero values than were in the BTBW. It is also likely that predictions were lower for the BTBW because of factors other than structure and vegetation type (e.g., social information) influencing prevalence for this species' information (Betts et al., 2008).

4.5.2 Variable importance in predictions

Radar metrics were not selected within the 10 most important when lidar data were used because of the correlation between radar and lidar metrics and the stronger explanatory power of lidar. When selected, the HH/VV or the ratio or surface-to-volume scattering was more effective than other ratios using cross-polarized bands because they were more correlated with structural metrics than any other radar metric. The HV/VV metric, which is also strongly sensitive to volume scattering, was useful in models using radar metrics alone but had lower explanatory power than the ratios

of HH and VV bands. In this study, we only used band ratios and indices from radar data. Another useful approach could be to derive characteristics such as patch characteristics and edges similar to Imhoff et al. (1997) or structure from interferometric SAR (Treuhaft et al., 2004).

The NDVI difference or deciduousness metric from Landsat was extremely useful in predicting bird prevalence individually and in combination with lidar metrics because it accounted for differences in conifer and deciduous species. It was not selected as an important metric only for the Blackpoll Warbler because elevation and DRL crown metrics were more important for this species. Deciduousness was the most important predictor for the Ovenbird and the Magnolia Warbler. As previously discussed, the difference is that the Ovenbird prefers high deciduousness or broadleaved species while the Magnolia Warbler prefers low deciduousness or conifer species, consistent with Doran et al. (2003).

Elevation from lidar data was an important variable for 5 out of the 8 species studied. Total crown diameter for trees identified within each plot from DRL data was the best predictor for 5 out of the 8 bird species. This metric was very similar to stem density from DRL data. Some species were positively correlated with total crown diameter (e.g., Blackpoll Warbler, Magnolia Warbler) showing preference for greater stem density and small trees while others (e.g., Black-throated Blue Warbler, Red-eyed Vireo) were negatively correlated showing preference for lower stem density and large trees. In addition to total crown diameter, average crown diameter and crown weighted height were also important metrics in many models.

DRL crown metrics were strongly correlated with LVIS RH metrics and LVIS canopy cover metrics at different levels within the canopy but had more explanatory power for bird prevalence than LVIS metrics for most species. We expected cover metrics at lower levels within the canopy to be selected as important for species that had a higher preference for shrubs (Ovenbird) and well-developed understory (Black-throated Blue Warbler). However, the LVIS canopy cover between 20-25 m or overstory cover were better predictors for most species than lower-canopy cover metrics. This does not mean that lower-cover metrics were not important. For some species (e.g., Yellow-rumped Warbler), the relationship with cover at 20-25 m was negative, meaning bird prevalence was higher in areas with low cover in the overstory, which is more consistent with expected patterns. Limited analysis showed variations in prevalence at different levels within the canopy for the same species, consistent with findings from field data (Robinson & Holmes, 1984). More research is needed to model these relationships explicitly with lidar data and understand their influence on habitat selection.

4.5.3 Predictions at plot and landscape scale

Our results showed that RF models overestimated prevalence in some cases and underestimated them in others. This was probably because the regression was weighted by the most frequently occurring prevalence values. When absences were more than presences, they were predicted accurately or overestimated. In birds that were more abundant, the class with the highest frequency was overestimated. For example, in the case of the Black-throated Blue Warbler, the highest prevalence was between 4 and 5 years. RF Forest predictions overestimated the 4-5 year prevalence

for BTBW by nearly 40%. On the other hand, prevalence was always underestimated beyond 7 years in all species. This shows that the RF models were weaker in predicting beyond 6 years' prevalence. More work is needed to compare Random Forests regression models with other statistical approaches like Boosted Regression Trees (Elith et al., 2006) or ensemble modeling (BIOMOD, Lomba et al., 2010) for fusion applications.

Spatial patterns of predictions at landscape level from both radar and fusion were consistent with bird abundance maps generated by Doran et al. (2003) despite some over- and underestimation. Although there was more mixing between low and good habitat predicted from radar alone, areas with low- and medium-habitat quality were well identified. Radar data per se may be more useful in classifying presence and absence even if it is weaker for prevalence. Radar data can also be used to stratify areas at larger scales to identify species presence and further detailed mapping of prevalence or abundance can be done with fusion.

Our results from quantile regression forests suggest that predictions from Random Forests and other methods may not be sufficient for decision making if they only provide mean values. The 10% or worst-case predictions for the Black-throated Blue Warbler suggest that there may be little or no habitat with high quality in the areas predicted as excellent from mean values. If resources for habitat management were limited, a 10% quantile map would show areas that necessarily need conservation. On the other hand, the 90% quantile map predicts that practically the entire study area is suitable habitat for the Black-throated Blue Warbler. This is useful when maximum resources can be applied to conserve a particular species. The

absence of uncertainty maps in most studies of statistically-predicted habitat has been a major limitation and must be given much more attention.

4.6 Conclusion

This research showed that combining metrics from radar, lidar and Landsat data can improve predictions from either dataset alone. The improvement in predictive power was highest for radar variables. With space-borne radar data becoming increasingly available, both spatially and temporally, there is considerable potential in mapping habitat characteristics directly from radar alone or in combination with Landsat and other multispectral data such as ASTER. In areas where lidar data are available, the addition of structural metrics could further improve maps of habitat characteristics.

Random forest regression is a powerful machine learning method that can be applied in ecological and habitat studies in many ways. We showed the application of variable importance measures from RF models in determining useful metrics for fusion. These studies can be extended to include other metrics and other datasets as well. More research is needed to test the accuracy of random forest regression and compare it with methods to model zero-inflated data for rare bird species.

Finally, this study shows that multi-sensor fusion is powerful for mapping multi-dimensional habitat attributes and can provide much more information than any one sensor alone. We found several similarities and differences between small- and medium-footprint lidar data at plot scale in this study. More research is needed to test whether these similarities can be used to derive attributes of one dataset from the other using limited spatial sampling. There is also considerable potential in using

other metrics from lidar such as understory shrub density, canopy layers and patch characteristics as well as other landscape metrics from canopy 3D structure that is potentially available from radar interferometry (Treuhaft et al., 2004).

Chapter 5 Discussion and Conclusion

This dissertation explored multi-sensor fusion for bridging the gap between science requirements for carbon/biodiversity studies and lack of comprehensive forest structure maps using remote sensing. My research resulted in the generation of a wide range of quantitative maps including canopy height, biomass, canopy cover at 5 m vertical intervals, high resolution maps of vegetation type/genera, biochemical status, canopy height and biomass change. The methods used in this dissertation are not site-specific and can be applied to other study areas where data are available. My research has added insights into remote sensing fusion approaches and also provided further evidence that multi-sensor fusion with lidar is powerful over using lidar, hyperspectral or radar alone.

Chapter 2 focused on combining lidar and hyperspectral data for mapping biomass in the Sierra Nevada. Here, I used fusion at two levels: the first one was a direct statistical fusion of lidar height metrics and hyperspectral band ratios using stepwise regression, a simple yet useful and widely used method for biomass estimation. My results showed that narrow band derivative indices and water band ratios from hyperspectral data added little value to biomass estimates from lidar because they were also moderately correlated with lidar metrics. In addition, these indices suffered saturation effects in the high biomass forests of this study area, even though they are known to be more sensitive to canopy biophysical properties than multispectral data. Nevertheless, the correlation of these indices with lidar heights

suggests that hyperspectral data could be more effective in low biomass density (< 100 Mg/ha) ecosystems than in high biomass forests.

The other approach involves stratification of hyperspectral data into vegetation types/genera using sub-pixel level image processing and integration with lidar equations for biomass estimation. Even though stratification before lidar estimation of biomass did not improve results significantly at field plot scale, confidence intervals were narrower and spatial predictions with hyperspectral and lidar data had lower errors for some genera, such as hardwoods and pines. This method is applicable to multispectral data while hyperspectral data have the added advantage in discriminating species level differences in vegetation. Species stratification with hyperspectral/multispectral data could therefore be a viable strategy for biomass estimation with sparse lidar data.

While the integration of lidar and hyperspectral data has advantages for the carbon cycle, fusion could be of greater value for mapping ecosystem and habitat characteristics. I demonstrated one application by mapping areas with relatively high and low stress in the Sierra National Forest. Many high biomass density stands showed low moisture and chlorophyll content in the study area. Spectra from hyperspectral images in these sites also showed signatures of chlorophyll stress, senescence or dead vegetation (non photosynthetic vegetation).Areas with a combination of 'low-chlorophyll', low-moisture' and higher abundance of NPV spectra in high biomass, high canopy cover forests were identified as likely to be under physiological/structural stress. While these areas could be early signs of canopy loss or mortality from various factors, they could be caused by more woody debris,

leaf litter or simply mature forests with less green understory. These results are noteworthy and could potentially be linked with recent findings on increased tree mortality in the Sierra Nevada from temperature induced moisture stress (van Mantgem et al., 2009).

The next step in taking this research forward would be to define 'stress' more robustly. One way to do this is by differentiating between mature woody vegetation and physiological stress using changes in chemical composition from hyperspectral data (e.g. xanthophylls using the Photochemical Reflectance Index). Structural stress can also be analyzed better by mapping abundance of NPV within individual tree crowns. Combining small footprint lidar with hyperspectral data and stem maps from field data can be another way of mapping stress related changes.

I compared stress and canopy structural changes further as one part of the analyses in Chapter 3, where I mapped canopy height and biomass changes in the Sierra Nevada with temporal lidar data. The larger goal of Chapter 3 was to evaluate the efficacy of multi-date lidar data in mapping canopy dynamics, as a test bed for space borne lidar sensors such as the Deformation, Ecosystem Structure and Dynamics of Ice (DESDynI). The availability of two sets of field and lidar measurements, one in 1999 and again in 2008, allowed the quantification of changes over a decade. However, field plots were measured in 2008 for biomass validation experiments for DESDynI and not biomass change and did not cover the range of disturbance and growth. Therefore, field measured height and biomass changes did not have enough variability to be detected with lidar measurements from 1999 and 2008.

Next, I tested whether changes could directly be measured from waveform lidar data. This approach was used because lidar metrics had been validated for quantifying height and biomass in previous studies. My analysis showed that after accounting for geolocation and elevation errors between nearly co-incident lidar footprints there was a small positive change in canopy height significant over more than 100,000 lidar footprints in the nine year time interval. By further analyzing height changes in these footprints with transition probability matrices, I was able to project height distributions at equilibrium/steady state. Comparing these with lidar canopy height distributions from 1999 and 2008 showed that the landscape was not in steady state in both years, but was recovering from past disturbances (e.g. clear cut logging, fire), with more growth in smaller trees. Results also showed that the landscape would reach steady state after around 300 years under the current disturbance regime.

I further mapped height and biomass changes across the landscape to detect the areas that had statistically significant change. Only about 20 % of the 22,000 ha study area showed significant height and biomass changes suggesting most of the landscape did not undergo catastrophic changes large enough to be detected by lidar. Biomass changes that were significant were also likely to be potential carbon sources and sinks because they had to be larger than 100 Mg/ha to be detected. Results from footprint, plot and hectare scales taken together suggest overall positive changes and a small net carbon sink in these forests. These results are important because currently no maps quantifying height and biomass changes at such fine scales used in this study.

Analyzing forest structural changes across the landscape with two dates of lidar data revealed many aspects relevant to land use and ecosystem processes in the study area. Comparisons of height change maps with aerial photos showed signs of re-growth in areas with height increases and clear-cut logging/other losses in places where heights decreased. The Sierra National Forest has areas managed for timber as well as those protected for the California Spotted Owl habitat. My analysis showed that both protected and managed forests had an average height increase in the time period studied but growth in protected areas was significantly higher. Going back to results from Chapter 2, I compared height changes in the low and high stress areas detected with hyperspectral data, acquired in 2003. I found that there were significant canopy height losses in the stressed areas while the forests with lower stress had significant canopy height increase. This exercise showed that stress maps from fusion may be indicators of areas likely to undergo changes.

While Chapters 2 and 3 focused on deriving forest structural characteristics from lidar and fusion, Chapter 4 was an application of different sensors in habitat mapping. I tested fusion approaches in the Hubbard Brook Experimental Forest in New Hampshire, where abundant bird data, radar, lidar and Landsat data were available. The relationships between forest structure, vegetation composition and bird habitat preferences have been studied at the HBEF since the 1960s. More recently, Goetz et al (2010) suggested that novel metrics can be derived from waveform lidar for mapping bird habitat quality for the black-throated blue warbler. In this study, I tested the relative and combined explanatory power of radar, lidar and multispectral

data in predicting bird habitat quality for 8 bird species including the black-throated blue warbler using data mining methods.

Results from this study showed that radar data alone explained more than 30% variance in bird prevalence for the eight bird species studied. Fusion improved results from either sensor alone; by 25% for radar, 15% for Landsat and 5% for lidar on an average. These results suggest that radar and multispectral data could be used to predict and mask forests where a given species was not likely to be present. Areas that show presence can further be mapped more accurately with fusion approaches or scattered lidar.

Analysis of predictor variables from each dataset further showed that complementary attributes such as crown characteristics, canopy cover, phenology improved results over using any one sensor alone. Although bird habitat quality/ prevalence maps from fusion were more accurate than from radar data alone, spatial distributions from both were consistent with known habitat preferences for the birds. One interesting aspect of this study was the comparison of relative efficacies of small- and medium-footprint lidar. This study showed that individual tree crown diameter metrics from small-footprint lidar was among the most important predictor variables suggesting the need to derive more high resolution metrics from this dataset. However, there were also strong correlations between from the small– and medium-footprint lidar metrics suggesting that medium-footprint lidar would be just as effective for larger landscape scale habitat mapping.

The habitat study at Hubbard Brook was a combination of better data processing methods and more advanced machine learning methods than shown in the

other studies. The results from this study can further form a feedback loop to refine methodological approaches used in Chapters 2 and 3. Even though Chapter 4 had a different study area, the methods and analyses are common to both study sites and can be applied to other study areas as well with site-specific modifications.

Common limitations in all the studies were geolocation errors, and time intervals between acquisition dates of different datasets. I made an effort to reduce them to a minimum by correcting errors when possible or using spatial scales that minimized errors (such as geolocation) for my studies. However, some geolocation shifts between lidar and field plots were not easy to detect and correct. Lack of field data for validation is a common problem in many remote sensing studies and this study was no exception. Although research has suggested that lidar data could be used as ground truth because of high validation accuracies, my results suggest the need for more rigorous field validation for mapping canopy dynamics from lidar. Lidar and hyperspectral fusion studies require more field spectra for validating stress and species composition maps. In addition to data for forest structure, bird and animal data are also required to develop useful fusion applications for habitat studies.

The limitations described above point to some fundamental problems encountered in multisensor fusion applications. It is very likely that if the input datasets are not processed robustly for atmospheric corrections, terrain, geolocation and other issues, fusion results could be poorer than using either sensor by itself. Since the number of datasets used in this research was small, it was possible to analyze some of the errors at fine spatial scales. Scaling these studies to regional or landscape level will require more rigorous error analysis on input datasets, more

innovation and automated methods for integrating large databases using data mining approaches.

Remote sensing fusion is still in its infancy and far more research is need to develop and assess algorithms for combining data from different sensors in ways that are germane to the task at hand. The largest limitation of lidar data is that it is expensive to obtain from airborne sensors and has reduced spatial sampling from existing and planned space borne sensors. This almost demands that fusion techniques be developed for applied uses of these data to forest and habitat management. This research has shown great advantages in combining different remote sensing data with lidar. Lidar and hyperspectral data can form a powerful combination for habitat mapping by proving structure, species composition and stress. Radar data is also useful for large scale mapping and accuracies of forest structure and habitat characteristics from radar can be further improved with lidar samples. It is unquestionable that the coming decades will see an explosion of work on fusion. The many new missions coming online, from SMAP (Soil Moisture Active and Passive) to DESDynI to HyspIRI will provide an unprecedented wealth of global remote sensing data. The availability of spaceborne lidar, radar and hyperspectral data can make it possible to extend the fusion applications developed in this study to larger spatial scales, for carbon monitoring in tropical forests, mapping pine beetle infestations in temperate forests and mapping habitat for rare and endangered species in other parts of the world.

From a habitat management perspective, lidar, radar, hyperspectral and multispectral data in synergy can provide information on forests and habitats that

were previously unavailable. Armed with such multi-dimensional information and improved computational capabilities, ecologists can explore biodiversity and ecological interactions in many novel ways. With more synthesis between the world views of ecologists, remote sensing scientists and engineers, fusion applications can become more powerful and truly applicable to conservation in the real world. This research was one such attempt to integrate technical aspects of remote sensing with practical applications in carbon science and biodiversity. Towards this end, I hope my dissertation has provided some forward progress

Bibliography

Adams, J.B., Smith M.O. and Johnson P.E. (1986). Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander 1 site. *Journal of Geophysical Research 91(B8),* 8090–8112.

Anderson, J.E. Plourde, L.C., Martin,M.E., Braswell, B.H., Smith, M.L., Dubayah, R.O., Hofton, M.A., Blair, J.B. (2008). Integrating waveform lidar with hyperspectral imagery for inventory of a northern temperate forest. *Remote Sensing of Environment*: *Remote Sensing Data Assimilation Special Issue, 112(4),* 1856-1870.

Anderson, J.E., Martin, M.E., Smith, M.L., Dubayah, R.O., Hofton, M., Hyde, P., Peterson, B., Blair, J.B., and Knox, R. (2005). The use of waveform lidar to measure northern temperate mixed conifer and deciduous forest structure in New Hampshire. *Remote Sensing of Environment, 105(3),* 248-261.

Asner, G. P. (1998). Biophysical and Biochemical Sources of Variability in Canopy Reflectance. *Remote Sensing of Environment*, *64(3),* 234-253.

Asner,G.P., Knapp, D.E., Ty Kennedy-Bowdoin, Jones, M.O., Martin R.E., Boardman J., Hughes F. (2008). Invasive species detection in Hawaiian rainforests using airborne imaging spectroscopy and LiDAR. *Remote Sensing of Environment 112(5),* 1942-1955.

Balzter, H. (2001). Forest mapping and monitoring with interferometric synthetic aperture radar (InSAR). Progress in Physical Geography 25:159-177.

Bergen, K. M., Goetz, S. J., Dubayah, R. O., Henebry, G. M., Hunsaker, C. T., Imhoff, M. L., et al. (2009). Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. *Journal of Geophysical Research, 114*, 1-13. doi: 10.1029/2008JG000883.

Bergen, K.M., Gilboy, A.M., and Brown, D.G. (2007). Multi-dimensional vegetation structure in modeling avian habitat. *Ecological Informatics*, *2(1),* 9-22.

Bergen, K.M., Knox R.G., and Saatchi, S. (Eds.) (2006). Multi-Dimensional Forested Ecosystem Structure: Requirements for Remote Sensing Observations, *NASA/CP-2005-212778, NASA GSFC, Washington, D.C., 36 pp*.

Berk R.A.(2008). Statistical Learning from a Regression Perspective,DOI: 10.1007/978-0-387-77501-2 1, *Springer Science+Business Media*, LLC 2008

Betts, M. G., Hadley, A. S., Rodenhouse, N., & Nocera, J. J. (2008). Social information trumps vegetation structure in breeding-site selection by a migrant songbird. *Proceedings Biological sciences / The Royal Society, 275(1648)*, 2257-63. doi: 10.1098/rspb.2008.0217.

Biondini, M., & Kandus, P. (2006). Transition matrix analysis of land-cover change in the ´ river accretion area of the lower delta of the parana (argentina) reveals two succession pathways. *Wetlands*, *26*(4), 981-991.

Blair J.B., Hofton M.A, Rabine D.L. (2006). Processing of NASA LVIS elevation and canopy (LGE, LCE, and LGW) data products*, version 1.01. http://lvis.gsfc.nasa.gov, 2006.*

Blair, J.B., Rabine, D.L., & Hofton, M.A.(1999). The Laser Vegetation Imaging Sensor (LVIS): A medium- altitude, digitization-only, airborne laser altimeter for mapping vegetation and topography. *ISPRS Journal of Photogrammetry and Remote Sensing, 54,* 115-122.

Bradbury, R. B., Hill, R. A., Mason, D. C., Hinsley, S. A., Wilson, J. D., Balzter, H., et al. (2005). Modelling relationships between birds and vegetation structure using airborne lidar data: A review with case studies from agricultural and woodland environments. *IBIS, 147(3),* 443-452.

Breiman L, Cutler A. (2003). Setting up, using, and understanding Random Forests v4.0. [online] URL: http://www.stat.berke- ley.edu/users/breiman/rf.html.

Breiman, L. (2001). Random forests. *Machine Learning* 45:15–32.

Buongiorno, J. (1980). A Matrix Model of Uneven-Aged Forest Management. *Forest Science, 26(4)*, 609-625.

Canadell JG, *et al.* (2007). Contributions to accelerating atmospheric CO2 growth from economic activity, carbon intensity, and efficiency of natural sinks. *Proc Natl Acad SciUSA* 104:18866–18870.

Caprra, P., & Rivest, L. (1995). On the variance of the trimmed mean. *Statistics*, *22*, 79-85.

Caswell, H. (2000). Matrix Population Models: Construction, Analysis, and Interpretation, *1st edition, Sinauer Associates Inc., Sunderland, MA, USA*.

Chambers J.Q., Asner G.P., Morton D.C., Anderson L.O., Saatchi S.S., Fernando D.B. Espirito-Santo, Palace M., Souza C. Jr.(2007). Regional ecosystem structure and function: ecological insights from remote sensing of tropical forests. *Trends in Ecology & Evolution*, *22(8),* 414-423.

Clark M.L., Clark D.B. & Roberts D.A. (2004). Small-footprint lidar estimation of sub-canopy elevation and tree height in a tropical rain forest landscape. *Remote Sensing of Environment, 91,* 68-89

Clark, M.L., Roberts, D.A., Clark, D.B.(2005).Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sensing of Environment, 96(3-4),* 375-398.

Collins, B. M., Kelly, M., Wagtendonk, J. W., & Stephens, S. L. (2006). Spatial patterns of large natural fires in Sierra Nevada wilderness areas. *Landscape Ecology*, *22*(4), 545-557.

Coppin, P., I. Jonckheere, K. Nackaerts, B. Muys, Lambin E..(2004). Digital change detection methods in ecosystem monitoring; A review. *International Journal of Remote Sensing 25*,1565–1596.

Curran, P.J., Dungan, J.L., and Peterson, D.L. (2001). Estimating the foliar biochemical concentration of leaves with reflectance spectrometry testing the Kokaly and Clark methodologies. *Remote Sensing of Environment, 76 (3),* 349–359.

Cutler, D. R., T. C. Edwards, K. H. Beard, A. Cutler, K. T.Hess, J. Gibson, and J. J. Lawler (2007). Random forests for classification in ecology. *Ecology* 88:2783–2792.

Das, A., Battles, J., van Mantgem P.J., Stephenson. N. (2008). Spatial elements of mortality risk in old-growth forests. *America*, *89*(6), 1744-1756.

De'ath, G., and K. E. Fabricius. (2000). Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology 81*, 3178–3192

Degraaf, R. (1998). Associations between breeding bird abundance and stand structure in the White Mountains, New Hampshire and Maine, USA. *Forest Ecology and Management*, *103(2-3)*, 217-233.

Dennison, P. E., Roberts, D. A. (2003). Endmember selection for multiple endmember spectral mixture analysis using endmember average RMSE. *Remote Sensing of Environment, 87(2-3),* 123–135.

Dennison, P.E, Halligan, K.Q., Roberts, D.A. (2004). A comparison of error metrics and constraints for multiple endmember spectral mixture analysis and spectral angle mapper. *Remote Sensing of Environment, 93(3),* 359-367.

Doran, P. J. (2003). Intraspecific spatial variation in bird abundance: patterns and processes. *Dissertation*. Dartmouth College, Hanover, New Hampshire, USA.

Doran, P. J. and Holmes R.T. (2005). Habitat occupancy patterns of a forest dwelling songbird: causes and consequences. *Canadian Journal of Zoology 83*,1297-1305.

Drake, J.B, Dubayah, R., Knox, R.G., Clark, D.B., Blair, J.B. (2002b). Sensitivity of large-footprint lidar to canopy structure and biomass in a neotropical rainforest. *Remote Sensing of Environment,81(2-3),* 378– 392.

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. (2002a). Estimation of tropical forest structural characteristics using large-footprint lidar. *Remote Sensing of Environment, 79(2-3),* 305-319.

Dubayah, R. O., & Drake, J. B. (2000). Lidar remote sensing for forestry. *J. Forestry 98(6),* 44-46.

Dubayah, R. O., Sheldon, S. L., Clark, D. B., Hofton, M.A., Blair, J.B., Hurtt, G.C., Chazdon, R.L.(2010). Estimation of Tropical Forest Height and Biomass Dynamics Using Lidar Remote Sensing at La Selva , Costa Rica. *Journal of Geophysical Research, 115,* G00E09, doi:10.1029/2009JG000933

Elith J., Graham C.H., Anderson R.P, Dudík M., Ferrier S., Guisan A., Hijmans R.J, Huettmann F., Leathwick J.R., Lehmann A., Li J., Lohmann L.G., Loiselle B.A, Manion G., Moritz C, Nakamura M., Nakazawa Y., Overton J.M.M., Peterson A.T., Phillips S.J., Richardson K., Scachetti-Pereira R., Schapire R.E., Soberón J., Williams S., Wisz M.S. Zimmermann N.E. (2006). Novel methods improve prediction of species' distributions from occurrence data, *Ecography* 29 pp. 129–151.

Elvidge, C.D., and Chen, Z. (1995). Comparison of broad-band and narrow-band red and near-infrared vegetation indices. *Remote Sensing of Environment, 54(1),* 38-48.

Franke J., Roberts D.A., Halligan K., Menz G. (2009). Hierarchical Multiple Endmember Spectral Mixture Analysis (MESMA) of hyperspectral imagery for urban environments. *Remote Sensing of Environment*, *113(8),*1712-1723.

Franklin, S.E. and Wulder, M.A. (2003): Remote sensing methods in high spatial resolution satellite data land cover classification of large area. *Progress in Physical Geography 26*, 173–205.

Frolking, S., Palace, M. W., Clark, D. B., Chambers, J. Q., Shugart, H. H., Hurtt, G. C., et al. (2009). Forest disturbance and recovery: A general review in the context of spaceborne remote sensing of impacts on aboveground biomass and canopy structure. *Journal of Geophysical Research*, *114*, G00E02, doi:10.1029/2008JG000911

Gao, B.C. (1996). NDWI-A normalized difference water index for remote sensing of vegetation liquid water from space, *Remote Sensing of Environment 58(3),* 257-266.

Gitelson, A.A., Kaufman, Y.J., Stark, R., Rundquist, D. (2002). Novel algorithms for remote sensing estimation of vegetation fraction, *Remote Sensing of Environment, 80(1),* 76–87.

Goetz S.J., Steinberg D., Betts M.G., Holmes R.T, Doran P.J., Dubayah R., Hofton M. (2010). Lidar remote sensing variables predict breeding habitat of a Neotropical migrant bird. *Ecology, 91(6),* 1569–1576

Goetz, S. J. (1997). Multi-sensor analysis of NDVI, surface temperature, and biophysical variables at a mixed grassland site. *International Journal of Remote Sensing 18*:71–94.

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(3),* 254-263.

Green, R. O. Conel, J. E., and Roberts, D. A. (1993). Estimation of aerosol optica1 depth and additional atmospheric parameters for the calculation of apparent reflectance from radiance measured by the Airborne Visible/Infrared Imaging Spectrometer. In *Summaries of the Fourth Annual JPL Airborne Geoscience Workshop, Oct 25-29.1993, Vol.1 AVIRIS Workshop, Washington, DC. pp.* 73-76.

Green, R. O., Eastwood, M. L., Sarture, C. M., Chrien, T. G., Aronsson, M., Chippendale, B. J., Faust, J. A., Pavri, B. E., Chovit, C. J., Solis, M., Olah, M. R., & Williams, O. (1998). Imaging spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). *Remote Sensing of Environment, 65(3),* 227–248.

Guisan, A. & Zimmerman, N.E. (2000). Predictive habitat distribution models in ecology. *Ecological Modeling*. *135*,147-186.

Gustafsen, E.J. (1998). Quantifying landscape spatial pattern: what is the state of the art? *Ecosystems 1*, 143–56.

H. Tang, R. Dubayah (2010). Sensitivity of LIDAR Canopy Height Estimate to Geolocation Error, proceedings of the American Geophysical Union.

Hall, D.L., 1992, Mathematical Techniques in Multi-sensor Data Fusion, Artech House, Norwood, Massachusetts.

Hall D.L., Llinas J.(1997). An introduction to multisensor data fusion, *Proceeding,IEEE, vol. 85*, 6–23

Hall, F. G., Botkin, D. B., Strebel, D. E., Woods, K. D., & Goetz, S. J. (1991). Large-Scale Patterns of Forest Succession as Determined by Remote Sensing. *America*, *72*(2), 628-640.

Hese S., Lucht, W., Schmullius, C., Barnsley, M., Dubayah, R., Knorr, D., Neumann, K., Riedel, T., and Schröter, K. (2005). Global biomass mapping for an improved understanding of the CO2 balance? The Earth observation mission Carbon-3D. *Remote Sensing of Environment*, *94(1),* 94-104.

Hinsley SA, Hill RA, Bellamy PE, and Balzter H. (2006). The application of lidar in woodland bird ecology: climate, canopy structure, and habitat quality. *Photogrammetry and Remote Sensing 72: 1399–1406.*

Hofton, M. A., & Blair, J. B. (2002). Laser altimeter return pulse correlation: a method for detecting surface topographic change. *Journal of Geodynamics*, *34(3-4)*, 477-489.

Hofton, M. A., R. Dubayah, J. B. Blair, and D. Rabine (2006). Validation of SRTMelevations over vegetated and non- vegetated terrain using medium footprint lidar, Photogramm. Eng. Remote Sens., 72(3), 279–285.

Holmes, R. T., Bonney, R.E.Jr., Pacala S.W.(1979). Guild structure of the Hubbard Brook bird community: a multivariate approach. *Ecology 60(3)*, 512-520

Houghton, R. A., and S. J. Goetz (2008). New satellites help quantify carbon sources and sinks, *Eos Transactions, AGU, 89(43),* doi:10.1029/2008EO430001.

Houghton, R. A., Hall, F., Goetz, S. J. (2009). Importance of biomass in the global carbon cycle. *Journal of Geophysical Research*, *114*, 1-13

Houghton, R.A. (2005). Aboveground forest biomass and the global carbon balance. *Global Change Biology, 11,* 945-958.

Hudak, A.T., Lefsky, M.A, Cohen, W.B., & Berterretche, M. (2002). Integration of lidar and LANDSAT ETM+ data for estimating and mapping forest canopy height. *Remote Sensing of Environment, 82(2-3),* 397-416.

Huete, A.R. HuiQing Liu van Leeuwen, W.J.D. (1997). The Use of Vegetation Indices in Forested Regions: Issues of Linearity and Saturation *Geoscience and Remote Sensing, IGARSS '97*. Remote Sensing - A Scientific Vision for Sustainable Development 1997 IEEE International

Hunsaker C.T., Boroski B.B and Steger G.N. (2002). Relations between canopy cover and occurrence and productivity of California spotted owls, predicting species occurrences, issues of accuracy and scale*, Island Press, Covelo, CA (2002).*

Hyde P., Dubayah R., Peterson B., Blair J.B., Hofton M. and Hunsaker C. (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. (2007a). Mapping forest structure for wildlife habitat analysis using multi-sensor (Lidar, SAR/InSAR, ETM+, Quickbird) synergy, *Remote Sensing of Environment, 102(1-2),* 63-73.

Hyde, P.,Nelson, R.,Kimes, D.,Levine, E.(2007b).Exploring LiDAR-RaDAR synergy-predicting aboveground biomass in a southwestern ponderosa pine forest using LiDAR, SAR and InSAR, *Remote Sensing of Environment,106 (1),*28-38.

Imhoff M.L., Sisk T.D., Milne A., Morgan G., Orr T. (1997). Remotely sensed indicators of habitat heterogeneity: Use of synthetic aperture radar in mapping

vegetation structure and bird habitat. *Remote Sensing of Environment, 60(3)*, 217-227

Imhoff, M.L. (1995). Radar backscatter and biomass saturation: ramifications for global biomass inventory. *IEEE Transactions on Geoscience and Remote Sensing*, *33*, 510-518.

Jenkins, J. C., Chojnacky, D. C., Heath, L. S., & Birdsey, R. A. (2003). National-Scale Biomass Estimators for United States Tree Species. *Forest Science*, *49(1),*12-35

Jordan, C. F. (1969). Derivation of leaf area index from quality of light on the forest floor. *Ecology, 50*, 663-666

Keith, H., Mackey, B.G., Lindenmayer, D.B. (2009). Re-evaluation of forest biomass carbon stocks and lessons from the world's most carbon-dense forests. *PNAS, 106(28),* 11635-11640

Kellner, J. R., Clark, D. B., & Hubbell, S. P. (2009). Pervasive canopy dynamics produce short-term stability in a tropical rain forest landscape. *Ecology letters*, *12*(2), 155-64.

Kern, R. A., Gray, A. N., North, M., & Zald, H. S. (2008). Initial tree regeneration responses to fire and thinning treatments in a Sierra Nevada mixed-conifer forest, USA. *Forest Ecology and Management*, *256(1-2),* 168-179

Kim, Y., Yang, Z., Cohen, W. B., Pflugmacher, D., Lauver, C. L., Vankat, J. L., et al. (2009). Remote Sensing of Environment Distinguishing between live and dead standing tree biomass on the North Rim of Grand Canyon National Park , USA using small-footprint lidar data. *Remote Sensing of Environment*, *113(11),* 2499-2510

Kimes, D.S., Ranson, K.J., Sun, G., Blair, J.B, (2006). Predicting lidar measured forest vertical structure from multi-angle spectral data, *Remote Sensing of Environment, 100 (4),* 503 – 511.

Koetz, B., Morsdorf, F., Van der Linden, S., Curt, T., Allgöwer, B.(2008). Multi-source land cover classification for forest fire management based on imaging spectrometry and LiDAR data, *Forest Ecology and Management, 256(3),* 263-271.

Koetz, B., Sun, G., Morsdorf, F., Ranson, K.J., Kneubühler, M., Itten, K. Allgöwer, B. (2007). Fusion of imaging spectrometer and LIDAR data over combined radiative transfer models for forest canopy characterization *Remote Sensing of Environment Vol. 106(4),* 449-459.

Kokaly, R.F., Clark, R.N. (1999). Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. *Remote Sensing of Environment, 67(3),* 267–287.

Lefsky, M. A., Cohen, W. B., Harding, D. J., Parker, G. G., Acker, S. A., & Gower, S. T. (2002a.). Lidar remote sensing of above-ground biomass in three biomes. *Global Ecology & Biogeography*, *11(5),* 393-399.

Lefsky, M.A., Cohen, W.B., Parker, G.G., & Harding. D.J (2002) Lidar remote sensing for ecosystem studies. *Bioscience*, *52(1),* 19-30.

Lomba A., Pellissier L., Randin C., Vicente J., Moreira F., Honrado J., Guisan A.(2010). Overcoming the rare species modelling paradox: A novel

hierarchical framework applied to an Iberian endemic plant. *Biological Conservation, 143(11),* 2647-2657

Lutz, J.A., van Wagtendonk, J.W., and Franklin, J.F.(2009). Twentieth-century decline of large-diameter trees in Yosemite National Park, California, USA. *Forest Ecology and Management 257(11)* 2296-2307

MacArthur RH and MacArthur JW. (1961). On bird diversity. *Ecology* 42: 594–98.

Magness D.R., Huettmann, F., & Morton, J. M. (2008). Using Random Forests to Provide Predicted Species Distribution Maps as a Metric for Ecological Inventory & Monitoring Programs, *Studies in Computational Intelligence (SCI) 122,* 209–229

Maki, M., Ishiahra, M., and Tamura, M. (2004). Estimation of leaf water status to monitor the risk of forest fires by using remotely sensed data. *Remote Sensing of Environment*, *90*, 441 – 450

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

Martinuzzi, S., L. Vierling, W. Gould, and K. Vierling (2009). Improving the characterization and mapping of wildlife habitats with lidar data:

Mayer, K.E., and W.F. Laudenslayer, Jr. (Eds.). (1988). A Guide to the Wildlife Habitats of California. *California Dept of Forestry and Fire Protection. Sacramento, CA.*

McDermid, G.J., Franklin, S.E. and LeDrew, E.F.(2005). Remote sensing for large-area habitat mapping. *Progress in Physical Geography, 29,* 449–74.

Merton, R. N. (1998). Monitoring community hysteresis using spectral shift analysis and the red-edge vegetation stress index. *Proceedings of the Seventh JPL Airborne Earth Science Workshop, 275– 284. Pasadena, CA: Jet Propulsion Laboratory.*

Moody and Johnson (2001). A. Moody and D.M. Johnson, Land-surface phonologies from AVHRR using the discrete Fourier transform. *Remote Sensing of Environment 75(3),* 305–323.

Mundt, J., Streutker, D.R., Glenn, N.F. (2006). Mapping sagebrush distribution using fusion of hyperspectral and lidar classifications, *Photogrammetric Engineering and Remote Sensing (PE&RS), 72,* 47-54.

N. Meinshausen (2006). "Quantile Regression Forests", *Journal of Machine Learning Research 7,983-999*

National Research Council (NRC) (2007). Earth Science and Applications From Space: National Imperatives for the Next ecade and Beyond, Committee on Earth Science and Applications From Space: A Community Assessment and Strategy for the Future, 456 pp., Natl. Acad. Press, Washington, D. C.

Nelson R, Keller C, and Ratnaswamy M. (2005). Locating and estimating the extent of Delmarva fox squirrel habitat using an airborne LiDAR profiler. *Remote Sensing of Environment 96(3-4)*, 292–301.

Ni Meister, W., Dubayah, R., Jupp, D. (2001). Modeling lidar waveforms in heterogeneous and discrete canopies, *IEEE Transactions on Geoscience and Remote Sensing, 39(9),* 1943-1958.

North, M. P. (2002). The Teakettle Experiment 1 Teakettle's Role in the Kings River. *Sierra*, 47-54.

Numata, I., Roberts, D.A, Chadwick, O.A., Schimel, J.P., Galvao, L.S., Soares, J.V. (2008). Evaluation of hyperspectral data for pasture estimate in the Brazilian Amazon using field and imaging spectrometers, *Remote Sensing of Environment, 112(4),* 1569-1583.

Peñuelas, J., Filella, I., Biel, C., Serrano, L. & Savé, R. (1993). The reflectance at the 950–970 nm region as an indicator of plant water status. *International Journal of Remote Sensing, 14 (10),* 1887-1905.

Perry, G. L., & Millington, J. D. (2008). Spatial modelling of succession-disturbance dynamics in forest ecosystems: Concepts and examples. *Perspectives in Plant Ecology, Evolution and Systematics 9*,191-210

Pierce, L., Walker, W., Dobson, M., Hunsaker, C., Fites-Kauftnan, J., Dubayah, R., et al. (2002). Fusion of optical and SAR data for forestry applications in the Sierra Nevada of California. *IEEE International Geoscience and Remote Sensing Symposium*, 3(24-28), 1771-1773

Popescu S.C.(2007). Estimating biomass of individual pine trees using airborne lidar, *Biomass and Bioenergy*, *31(9),*646-655

Popescu S.C., Wynne R.H. (2004).Seeing the trees in the forest: using LIDAR and multispectral data fusion with local filtering and variable window size for estimating tree height. *Photogrammetric Engineering & Remote Sensing*, *70(5),* 589–604.

Prasad, A. M., L. R. Iverson, and A. Liaw (2006). Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems 9,* 181–199.

Pu, R., Gong, P., Biging, G.S., Larrieu, M.R.(2003) Extraction of red edge optical parameters from Hyperion data for estimation of forest leaf area index, *IEEE Transactions on Geoscience and Remote Sensing*, *41(4),*. 916-921.

R Development Core Team.(2009) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. hhttp://www. R-project.orgi

Ralph, C. J., J. R. Sauer, and S. Droege (1995) Monitoring bird populations by point counts. USDA Forest Service General Technical Report PSW-GTR-149, Albany, California, USA.

Rambo, T. R., & North, M. P. (2009). Forest Ecology and Management Canopy microclimate response to pattern and density of thinning in a Sierra Nevada forest. *Forest Ecology and Management*, *257*, 435-442.

Raumman, C. G., & Soulard, C. E. (2007). Land-Cover Trends in the Sierra Nevada Ecoregion, 1973-2000. *U.S. Geological Survey Scientific Investigations Report* 2007-5011 http://pubs.usgs.gov/sir/2007/5011/

Roberts, D. A., Gardner, M., Church, R., Ustin, S., Scheer, G., Green, R. O. (1998). Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models. *Remote Sensing of Environment, 65(3), 267– 279.*

Roberts, D. A., Green, R. O., Adams, J. B. (1997). Temporal and spatial patterns in vegetation and atmospheric properties from AVIRIS, *Remote Sensing of Environment, 62 (3),* 223-240.

Roberts, D., Halligan, K., and Dennison, P. (2007). VIPER Tools User Manual. *http://www.vipertools.org/, Santa Barbara.*

Roberts, D.A, Dennison, P.E. , Peterson, S., Sweeney, S. and Rechel, J. (2006). Evaluation of AVIRIS and MODIS Measures of Live Fuel Moisture and Fuel Condition in a Shrubland Ecosystem in Southern California*, J. Geophysical. Res. Biogeosciences*, *111,* G04S02, doi:10.1029/2005JG000113

Roberts, D.A. Dennison, P.E., Gardner, M., Hetzel, Y., Ustin, S.L., Lee, C. (2003) Evaluation of the potential of Hyperion for fire danger assessment by comparison to the Airborne Visible/Infrared Imaging Spectrometer, *IEEE Transactions on Geoscience and Remote Sensing 41 (6),*1297–1310.

Roberts, D.A., Ustin, S.L., Ogunjemiyo, S., Greenberg, J., Dobrowski, S.Z., Chen, J. and Hinckley, T.M. (2004). Spectral and structural measures of Northwest forest vegetation at leaf to landscape scales, *Ecosystems, 7*, 545-562.

Robinson, S. K. and R. T. Holmes. 1984. Effects of plant species and foliage structure on the foraging behavior of forest birds. *The Auk 101*,672-684

Rosen P.A., Hensley S., Wheeler K., Sadowy G., Miller T., Shaffer S., Muellerschoen R., Jones C., Zebker H., Madsen S.(2006) "UAVSAR: A New NASA Airborne SAR System for Science and Technology Research", 2006 *IEEE Conference on Radar, 24-27 April 2006, pp8.*

Rosenqvist, A., Milne, A., Lucas, R., Imhoff, M., Dobson, C. (2003). A review of remote sensing technology in support of the Kyoto protocol *Environmental Science & Policy, 6(5),* 441-455.

Roth, K. L. (2009). A combined lidar and hyperspectral remote sensing analysis for mapping forest biomass. Unpublished master's thesis. *University of California, Santa Barbara, Dept of Geography.*

Saatchi S., Dubayah R., Clark D., Chazdon R., Hollinger D. (2010).Estimation of Forest Biomass Change From Fusion of Radar and Lidar, *IEEE International Geoscience and Remote Sensing Symposium.*

Saatchi S., Houghton R., Avala R., Yu Y., Soares, J-V (2007).Spatial Distribution of Live Aboveground Biomass in Amazon Basin," *Global Change Biology, 13*, 816-837

Schwarz, P. A., T. J. Fahey, C. W. Martin, T. G. Siccama, and A. Bailey (2001). Structure and composition of three northern hardwood–conifer forests with differing disturbance histories. *Forest Ecology and Management, 144,* 201–212.

Serrano, L., Ustin, S.L., Roberts, D.A., Gamon, J.A., Penuelas, J. (2000). Deriving Water Content of Chaparral Vegetation from AVIRIS Data, *Remote Sensing of Environment, 74(3),* 570-581.

Simental E., Guthrie V., Blundell S.B. (2005). Polarimetry Band Ratios, Decompositions and Statistics for Terrain Characterization, *Pecora 16 "Global Priorities in Land Remote Sensing, Sioux Falls, South Dakota, Conference Proceedings*

Smith, K. (2004). Use of hyperspectral derivative ratios in the red-edge region to identify plant stress responses to gas leaks. *Remote Sensing of Environment*, *92*(2), 207-217.

Smith, T. F., Rizzo, D. M., North, M. (2005). Patterns of mortality in an old-growth mixed-conifer forest of the Southern Sierra Nevada, California. *Forest Science, Vol. 51(3)* 266-275.

Sonnetag, O., Chen, J.M., Roberts, D.A., Talbot, J., Halligan, K.Q., and Govind, A., (2007). Mapping tree and shrub leaf area indices in an ombrotrophic peatland through multiple endmember spectral unmixing, *Remote Sensing of Environment, 109,* 342-360

Swatantran A., Dubayah R., Roberts D., Hofton M., Blair J.B. (in press) Mapping Biomass and Stress in the Sierra Nevada using Lidar and Hyperspectral Data Fusion, *Remote Sensing of Environment, VEG3D Special Issue*.

Tanaka, H., & Nakashizuka, T. (1997). Fifteen Years of Canopy Dynamics Analyzed by Aerial Photographs in a Temperate Deciduous Forest, Japan. *Ecology*, *78*(2), 612-620.

Tanner, James T. (1942). The Ivory-billed Woodpecker. Research report No. 1. National Audubon Society, New York.

Treuhaft R., Law B.E., Asner G.P., (2004) Forest Attributes from Radar Interferometric Structure and Its Fusion with Optical Remote Sensing, *BioScience, 54(6),*561-571

Tucker, C.J. (1979) Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment 8,* 127-50.

USDA Forest Service, Remote Sensing Lab, Region 5 (2006) Land & Resource Management Plan (L&RMP) Land Suitability Class Map. http://www.fs.fed.us/r5/rsl/projects/frdb/layers/plsc.html

Usher, M. B. (1981). Modelling ecological succession, with particular reference to Markovian models. *Vegetatio*, *46-47*(1), 11-18.

Ustin, S. L., Roberts, D.A, Gamon, J.A., Asner, G.P., Green, R.O. (2004).Using imaging spectroscopy to study ecosystem processes and properties, *Bioscience 54(6),* 523-534.

van Mantgem, P. J.,. Stephenson N. L, Byrne J.C., Daniels L. D., Franklin J.F., Fulé P.Z., Harmon M.E., Larson A.E., Smith J.M., Taylor A.H., and Veblen T.T. (2009) "Widespread Increase of Tree Mortality Rates in the Western United States." *Science 323(5913):* 521-524.

Verner, J.; McKelvey, K.S.; Noon, B.R.; Gutiérrez, R.J.; Gould, G.I., Jr.; Beck, T.W.; [Technical Coordinators] (1992) The California Spotted Owl: A Technical Assessment of its Current Status Gen. *Tech. Rep. PSW-GTR-133*. Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture; 285 p

Vierling K.T., Vierling L.A., Gould W.A., Martinuzzi S., Clawges R.M. (2008) Lidar: shedding new light on habitat characterization and modeling. *Frontiers in Ecology and the Environment, 6(2),* 90-98

Waddell, K., & Hiserote, B. (2003). Technical documentation for the integrated database, version 1.0. *USDA Forest Service Pacific Northwest Research Station.*

Zald, H. S., Gray, A. N., North, M., & Kern, R. A. (2008). Initial tree regeneration responses to fire and thinning treatments in a Sierra Nevada mixed-conifer forest, USA. *Forest Ecology and Management*, *256(1-2),* 168-179.

Zhang J. (2010). Multi-source remote sensing data fusion: status and trends, *International Journal of Image and Data Fusion*, *1(1),*5 – 24.