

5-2012

Accuracy of Biomass and Structure Estimates from Radar and Lidar

Razi Uddin Ahmed
University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/open_access_dissertations



Part of the [Electrical and Computer Engineering Commons](#)

Recommended Citation

Ahmed, Razi Uddin, "Accuracy of Biomass and Structure Estimates from Radar and Lidar" (2012). *Open Access Dissertations*. 539.

<https://doi.org/10.7275/pabh-zz59> https://scholarworks.umass.edu/open_access_dissertations/539

This Open Access Dissertation is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

ACCURACY OF BIOMASS AND STRUCTURE ESTIMATES FROM RADAR AND LIDAR

A Dissertation Presented

by

RAZI AHMED

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2012

Electrical and Computer Engineering

© Copyright by Razi Ahmed 2012

All Rights Reserved

ACCURACY OF BIOMASS AND STRUCTURE ESTIMATES FROM RADAR AND LIDAR

A Dissertation Presented

by

RAZI AHMED

Approved as to style and content by:

Paul R. Siqueira , Chair

Stephen J. Frasier, Member

Yu Qian, Member

Scott Hensley, Member

C. V. Hollot, Department Head
Electrical and Computer Engineering

To my family

EPIGRAPH

“Any intelligent fool can make things bigger, more complex, and more violent. It takes a touch of genius, and a lot of courage, to move in the opposite direction.”

Albert Einstein

ACKNOWLEDGMENTS

I have been fortunate to have the support of a lot of people through my years in graduate school, years that I would have found impossible to navigate alone. First and foremost, I would like to thank Dr. Paul Siqueira, my advisor and mentor, for giving me the opportunity to work on so many interesting and challenging research questions. I have found his intellect, instinct and infinite amount of patience thoroughly inspiring. Needless to say, he has had a profound impact on me professionally and personally. I am deeply indebted to him.

I would also like to thank Dr. Scott Hensley for devoting so much of his precious time to help me with my dissertation. Looking back I find it hard to imagine doing all the work that I did without his help. This dissertation would simply not have been possible without his incredible insight and relentless push for higher standards. I am extremely grateful for his help and very excited to have the opportunity to work with him in the future.

I am also very thankful to Dr. Stephen Frasier for helping me out on various occasions during my graduate career as I grappled with issues in radars and remote sensing. I would also like to thank him and Dr. Qian Yu for agreeing to be on my dissertation committee, overseeing my work and for various pertinent, helpful suggestions.

There were many instances while working on my dissertation that I needed help on topics other than radars and remote sensing. I was fortunate enough to be able to seek out experts and have their ear. I would like to acknowledge and thank Dr. John Buonaccorsi from the Department of Mathematics and Statistics at the University of Massachusetts for his expert insight into complex statistics, and for taking time out to help an electrical engineer navigate murky statistical waters. I would like to

thank Dr. Kathleen Bergen from the School of Natural Resources at the University of Michigan for her help in various issues related to forest ecology and for sharing her broad insight into matters related to the remote sensing of forests, insight that led to some unique results. Even though my interaction with Dr. Michael Ter-Mikaelian of the Ontario Forest Research Institute was brief, it was very helpful and crucial to my understanding of forest mensuration. I am very grateful to him for helping out someone completely unfamiliar with his subject matter.

Any electrical engineer who has ever wandered into a forest for a field experiment would readily acknowledge the folly in underestimating the sheer amount of work. I was one such foolish individual, luckily I had the help of many people, without whom I would still be wandering the forests. I would like to acknowledge and thank Tim, Jeff, Mandy, Lili, Shanka, Tony, Caitlin, Tom among others who helped me hug trees, endlessly, one summer.

I have been fortunate to have developed everlasting friendships during my time at UMass. My friends have enriched my life, and are in no small part responsible for the person that I am now. For that I am very grateful to each and every one of them. To my girlfriend, Ale, for bringing so much joy and happiness in a very tough year, to Nino for teaching me everything about radar hardware, to Mauricio for being my closest confidant, to Katya, Lusya, Adri and Dragana, my second family, for always being there. I am also indebted to my colleagues at MIRSL, Edin, Brian, Harry, Harish, Vijay, Sumo, Salazar, Medina, Trabal and Linda for the many, many invaluable interactions. I am grateful to each and everyone for teaching me more than I ever let on.

Last, but not least, I would like to thank my family, my mother for making me the person that I am, my father for inspiring me to be better and my brother for daring to be everything we dreamt we would be as kids.

ABSTRACT

ACCURACY OF BIOMASS AND STRUCTURE ESTIMATES FROM RADAR AND LIDAR

MAY 2012

RAZI AHMED

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Paul R. Siqueira

A better understanding of ecosystem processes requires accurate estimates of forest biomass and structure on global scales. Recently, there have been demonstrations of the ability of remote sensing instruments, such as radar and lidar, for the estimation of forest parameters from spaceborne platforms in a consistent manner. These advances can be exploited for global forest biomass accounting and structure characterization, leading to a better understanding of the global carbon cycle. The popular techniques for estimation of forest parameters from radar instruments in particular, use backscatter intensity, interferometry and polarimetric interferometry. This dissertation analyzes the accuracy of biomass and structure estimates over temperate forests of the North-Eastern United States. An empirical approach is adopted, relying on ground truth data collected during field campaigns over the Harvard and Howland Forests in 2009. The accuracy of field biomass estimates, including the impact of the diameter-biomass allometry is characterized for the field sites. Full waveform lidar data from two LVIS field campaigns of 2009 over the Harvard and Howland forests

is analyzed to assess the accuracy of various lidar-biomass relationships. Radar data from NASA JPL's UAVSAR is analyzed to assess the accuracy of the backscatter-biomass relationships with a theoretical radar error model. The relationship between field biomass and InSAR heights is explored using SRTM elevation and LVIS derived ground topography. Temporal decorrelation, a major factor affecting the accuracy of repeat-pass InSAR observations of forests is analyzed using the SIR-C single-day repeat data from 1994. Finally, PolInSAR inversion of heights over the Harvard and Howland forests is explored using UAVSAR repeat-pass data from the 2009 campaign. These heights are compared with LVIS height estimates and the impact of temporal decorrelation is assessed.

TABLE OF CONTENTS

	Page
EPIGRAPH	v
ACKNOWLEDGMENTS	vi
ABSTRACT	viii
LIST OF TABLES	xiv
LIST OF FIGURES.....	xvii
 CHAPTER	
1. INTRODUCTION	1
1.1 Lidar Profiles	3
1.2 Backscatter Intensity.....	4
1.3 SAR Interferometry	6
1.4 Polarimetric Interferometry	10
2. FOREST MENSURATION, ERRORS AND FIELD CAMPAIGNS	15
2.1 Allometry	15
2.1.1 Methods for developing allometric relationships	17
2.1.2 Using existing allometric equations	19
2.1.2.1 Single site allometry: Ter-Mikaelian [118]	19
2.1.2.2 Ensemble allometry: Jenkins et. al. [65]	20
2.1.2.3 BLUE allometry: Lambert et. al. [71].....	21
2.2 Error Propagation Analysis	22
2.2.1 Errors in single site allometry	23
2.2.2 Error in ensemble equations	26

2.2.3	Error in BLUE equations	26
2.3	The Harvard Forest	29
2.3.1	Harvard Forest field campaign	29
2.3.2	Allometric equations for the Harvard Forest	31
2.3.3	Comparison of biomass estimates and errors	38
2.4	Howland Forest, Maine	42
2.4.1	Howland Forest field campaign	42
2.4.2	Allometric equations for the Howland Forest	42
2.4.3	Comparison of biomass estimates and errors	47
3.	LIDAR MEASUREMENT AND ERRORS	51
3.1	Lidar Data Over the Harvard Forest	51
3.1.1	LVIS shot selection methodology	52
3.1.2	LIVS RH metrics and field biomass	53
3.1.3	Discussion on LVIS RH metrics	56
3.1.4	Performance of other lidar metrics	58
3.1.5	Extending the range of field biomass values	67
3.2	Lidar Data Over the Howland Forest	72
3.2.1	Low biomass sites	75
3.2.2	Species composition	78
3.3	Combining Howland and Harvard Forest Datasets	81
3.3.1	Effects of various allometric equations	82
3.3.2	Saturation in RH metrics	85
4.	RADAR BACKSCATTER, ERROR MODEL AND OBSERVATIONS	87
4.1	Backscatter to Biomass	88
4.1.1	Error in radar estimates of biomass	89
4.1.2	Including parametric measurement error	90
4.1.2.1	Parametric bootstrapping	92
4.1.2.2	SIMEX: Simulation Extrapolation	94
4.2	Backscatter Error Model	95

4.2.1	Image speckle, thermal and multiplicative noise	96
4.2.2	Temporal variation	98
4.2.3	Error in backscatter calibration	99
4.2.4	Errors due to pixel area normalization	100
4.3	Radar Measurements	101
4.3.1	Backscatter-biomass relationships	106
4.3.1.1	Harvard Forest	106
4.3.1.2	Howland Forest	109
4.3.1.3	Combining Harvard and Howland Forest datasets	111
4.3.2	Accounting for measurement error	114
4.3.2.1	Analyzing temporal variability	114
4.3.3	Analyzing model error	122
5.	RADAR INTERFEROMETRY	127
5.1	Interferometric Correlation	129
5.1.1	Components of Interferometric Correlation	129
5.2	Interferometric Heights and Biomass	132
5.3	Observations of Temporal Decorrelation	135
5.4	Analysis of Temporal Decorrelation	138
5.4.1	Temporal decorrelation as a function of land-cover	139
5.4.2	Extent of atmospheric effects on interferometric coherence	145
6.	POLARIMETRIC INTERFEROMETRY	151
6.1	Mathematical Formulation	151
6.1.1	The coherence region	154
6.2	PolInSAR Parameter Inversion	158
6.2.1	Surface topography estimation	158
6.2.2	Height inversion	161
6.3	UAVSAR PolInSAR Data	163
6.3.1	Estimating the vertical wavenumber	165

6.3.2	Polarization synthesis and coherence optimization	166
6.3.3	Multi-polarization SNR correction	168
6.3.4	Estimation of topographic phase	173
6.4	PolInSAR Initial Results	173
6.4.1	Temporal stability of PolInSAR height estimates	178
6.4.2	Preliminary comparison with lidar heights	180
7.	CONCLUSIONS AND FUTURE WORK	183
7.1	Summary	183
7.2	Conclusions	184
7.3	Unique Contributions	187
7.4	Future Work	188
APPENDIX: VARIANCE-COVARIANCE MATRICES		190
BIBLIOGRAPHY		197

LIST OF TABLES

Table	Page
2.1 Summary of diameter data for the 23 species catalogued at the Harvard Forest and the single-site allometric equations chosen from Ter-Mikaelian and Korzukhin[118]	32
2.2 The ensemble equation coefficients from Jenkins et. al [65] with the twenty three species at the Harvard Forest grouped into eight categories	35
2.3 Summary of coefficients from the BLUE allometric equations of Lambert et. al. [71] for the twenty three species catalogued at the Harvard Forest	37
2.4 Summary of diameter data for the eleven species unique to the Howland Forest dataset and the corresponding single-site allometric equations chosen from Ter-Mikaelian [118]	45
2.5 Summary of coefficients from the BLUE allometric equations of Lambert et. al. [71] for the eleven species catalogued only at the Howland Forest	46
2.6 Comparison of biomass errors from the three allometries at the two study sites.	50
3.1 Regression statistics for LVIS RH metrics and field biomass for 5 Percent overlap.	54
3.2 Regression statistics for LVIS RH metrics and field biomass for at least 75 percent overlap.	55
3.3 Regression statistics for various Lidar Metrics and field biomass for an LVIS overlap of at least 5 percent. The RMSE values are in units of tons/hectare.	63
3.4 Regression statistics for various Lidar Metrics and field biomass for an LVIS overlap of at least 75 percent. The RMSE values are in units of tons/hectare.	64

3.5	Regression statistics for linear combinations of lidar variables to biomass data. The RMSEs are in units of tons/hectare.	67
3.6	Regression statistics for various Lidar Metrics and field biomass for an LVIS overlap of at least 35 percent over hectares formed by aggregating subplots from a sorted list. The RMSE values are in units of tons/hectare.	69
3.7	Regression statistics for LVIS RH metrics and field biomass for 5 percent overlap.	73
3.8	Regression statistics for LVIS RH metrics and field biomass for 75 percent overlap.	74
3.9	Regression statistics for LVIS RH metrics and field biomass for 5 and 75 percent overlaps excluding low biomass sites.	76
3.10	Regression statistics for LVIS RH metrics and field biomass over Harvard and Howland Forests estimated using the ‘coniferous-deciduous’ allometry for 5 percent overlap.	79
3.11	Regression statistics for LVIS RH metrics and field biomass over Harvard and Howland Forests estimated using coniferous and deciduous equations for 75 percent overlap.	80
3.12	Regression statistics for LVIS RH metrics over both the Howland and Harvard Forest with biomass values estimated using ‘coniferous-deciduous’ allometry and 5 percent overlap.	82
3.13	Regression statistics for LVIS RH metrics over both the Howland and Harvard Forest with biomass values estimated using ‘coniferous-deciduous’ allometry and 75 percent overlap.	83
3.14	Regression statistics for LVIS RH metrics using 50% overlap over both the Howland and Harvard Forest with biomass values estimated using the three allometric equations; Single-Site (Ter-Mikaelian [118], Ensemble (Jenkins [65]), BLUE (Lambert [71]).	84
4.1	UAVSAR instrument parameters.	102
4.2	List of UAVSAR scenes chosen over the Harvard and Howland Forests.	105
4.3	UAVSAR hardware and processing specific variables needed for computing error in backscatter measurements.	115

4.4	Observation specific variables needed for computing error in backscatter measurements.....	116
5.1	Regression statistics for linear fits between SRTM derived heights and biomass data from the Harvard and Howland Forest sites. The RMS errors are in units of tons/ha.	134
6.1	Summary for the thirteen UAVSAR lines collected over a period of eleven days in August 2009 over the Harvard Forest that have been processed as a stack of interferograms. The reference track (track number 6) is highlighted in yellow.	164
6.2	Cross-track baselines (in meters) between all pairs for the thirteen UAVSAR lines.....	164
6.3	Temporal separation (in hours) between all pairs for the thirteen UAVSAR lines.....	165
A.1	Variance-Covariance matrices Σ_{ab} used for calculating the error in tree biomass estimates for the species summarized in [71]	190

LIST OF FIGURES

Figure	Page
1.1 Lidar systems sample echoes of a downward transmitted pulse from various parts of a forest and creates a profile based on the time sequence of these echoes.	3
1.2 Two schematic representations of the InSAR observation geometry.	6
2.1 Variation between mean predicted values from the various allometric equations summarized in [118] for Red Maple (<i>Acer rubrum</i>) and Paper Birch (<i>Betula papyrifera</i>).	25
2.2 Harvard Forest tracts and plots. The inset shows a 1-hectare plot with its sixteen 25m by 25m subplots numbered one through sixteen.	30
2.3 Estimated tree weights with 95% confidence intervals for the four major tree species at the Harvard Forest.	38
2.4 Comparison of subplot-level biomass estimated using the Site-specific (Ter-Mikaelian [118]), Ensemble (Jenkins [65]) and BLUE (Lambert [71]) allometries.	39
2.5 Hectare level biomass estimates using Ensemble (Jenkins [65]), Single-site (Ter-Mikaelian [118]) and BLUE (Lambert [71]) allometries.	40
2.6 Subplot and plot level biomass error as a function of mean biomass from the Ensemble (Jenkins [65]), Single-site (Ter-Mikaelian [118]) and BLUE (Lambert [71]) allometries.	41
2.7 Howland Forest Research facility includes two sites, one near the town of Howland and the other near Penobscott in central Maine.	43
2.8 Comparison of subplot-level biomass estimated using the Site-specific (Ter-Mikaelian [118]), Ensemble (Jenkins [65]) and BLUE (Lambert [71]) allometries over the Howland Forest.	47

2.9	Comparison of hectare-level biomass estimated using the Site-specific (Ter-Mikaelian [118]), Ensemble (Jenkins [65]) and BLUE (Lambert [71]) allometries over the Howland Forest.	48
2.10	Subplot and plot level biomass error as a function of mean biomass from the Ensemble (Jenkins [65]), Single-site (Ter-Mikaelian and Korzhukhin [118]) and BLUE (Lambert [71]) allometries over the Howland Forest.	49
3.1	The LVIS nadir track is highlighted by green lines while boundaries of the Harvard forest are shown by the yellow lines. LVIS first imaged this area in 2003 and then in 2009.	52
3.2	LVIS RH metrics plotted against the biomass values from the 208 subplots and the 13 one-hectare plots for an overlap of at least 5 percent or more.....	53
3.3	LVIS RH metrics plotted against the biomass at subplot and hectare scales for an overlap of 75 percent or more.....	54
3.4	An example of errors in ground detection and LVIS RH100 corrected for badly detected ground values.	56
3.5	Comparison of LVIS RH metrics distributed by GSFC and RH metrics calculated at UMass.	57
3.6	Illustration of the steps involved in calculating the canopy height profile.	60
3.7	Correlation matrices for the 24 lidar variables for hectares and subplots with 5 and 75% shot overlap percentage.	66
3.8	LVIS RH metrics plotted against biomass data over hectares created by aggregating subplots from a sorted list.	68
3.9	Performance of models developed by regressing lidar variables to biomass data averaged over hectares created from sorted and unsorted data.	70
3.10	Sorted model is applied to both the sorted data (regression curve, shown by the solid lines) and the unsorted data (prediction curve, shown by the dashed lines) for varying number of subplots aggregated.	71

3.11 LVIS RH metrics plotted against the biomass values from the 364 subplots and the 23 hectare plots for an overlap of 5 percent or more.	73
3.12 LVIS RH metrics plotted against the biomass values from 193 of 364 subplots and the 23 hectare plots for an overlap of 75 percent or more.	74
3.13 LVIS RH metrics plotted against the biomass estimates from subplots and hectares excluding low-biomass datapoints for an overlap of 5 percent or more.	76
3.14 LVIS RH metrics plotted against data from both Harvard and Howland Forests using ‘coniferous-deciduous’ allometry and 5% overlap.	81
3.15 LVIS RH metrics plotted against data from both Harvard and Howland Forests using ‘coniferous-deciduous’ allometry and 75% overlap.	82
3.16 LVIS RH metrics averaged over 16-subplots chosen from a list of subplots sorted by biomass.	85
4.1 Polarimetric composite (HH-red, HV-green, VV-blue) of three stages of the UAVSAR backscatter correction, the left plane shows the calibrated image distributed by JPL, the center image is corrected for pixel area variations while the image on the right is corrected for both the pixel area variations and the backscatter law.	104
4.2 Cross-pol backscatter from UAVSAR plotted against field biomass estimates from the Ter-Mikaelian allometry at Harvard.	107
4.3 Cross-pol backscatter from UAVSAR plotted against field biomass estimates from the Jenkins allometry at Harvard.	107
4.4 Cross-pol backscatter from UAVSAR plotted against field biomass estimates from the Lambert allometry at Harvard.	108
4.5 Cross-pol backscatter from UAVSAR plotted against field biomass estimates from the Ter-Mikaelian allometry at Howland.	109
4.6 Cross-pol backscatter from UAVSAR plotted against field biomass estimates from the Jenkins allometry at Howland.	110

4.7	Cross-pol backscatter from UAVSAR plotted against field biomass estimates from the Lambert allometry at Howland.	110
4.8	Backscatter from UAVSAR plotted against field biomass estimates from the Ter-Mikaelian allometry for data combined from the two sites.	111
4.9	Backscatter from UAVSAR plotted against field biomass estimates from the Jenkins allometry for data combined from the two sites.	112
4.10	Backscatter from UAVSAR plotted against field biomass estimates from the Lambert allometry for data combined from the two sites.	112
4.11	Confidence intervals on biomass estimates using data from both Harvard and Howland Forests.	113
4.12	Cross-polarized backscatter UAVSAR data from twenty ascending lines over two regions of mostly evergreens and deciduous trees near the Harvard Forest.	117
4.13	Subplot and hectare scale errors in backscatter for the combined dataset over Harvard and Howland Forests. Also plotted, are model errors and allometric errors (in field biomass using the Single-site [118] equations) projected into units of radar backscatter using fitted coefficients.	118
4.14	Estimation of bias introduced by measurement error on the fit coefficients using SIMEX. Fits using biased and unbiased coefficients are plotted as well.	120
4.15	Cross-pol backscatter from UAVSAR plotted against field biomass estimates from the BLUE allometry [71] for data combined from the two sites. Confidence intervals from the naive fits as well as the ones corrected for modeled measurement error are shown.	121
4.16	Impact of species on the relationship between backscatter and field biomass over the Harvard and Howland Forests.	122
4.17	Measurement error corrected confidence intervals using the BLUE [71] allometry.	123

4.18 Measurement error corrected confidence intervals on biomass estimates from radar backscatter for the three allometries; single-site [118], ensemble [65] and BLUE [71] allometric equations.	124
4.19 Width of the confidence intervals on biomass estimates obtained from radar backscatter using the three allometries; single-site [118], ensemble [65] and BLUE [71] allometric equations.	125
5.1 Typical configuration for a repeat orbit interferometric SAR (InSAR).	128
5.2 SRTM derived topography and LVIS heights.	133
5.3 Heights derived from SRTM DEM and LVIS ground plotted against biomass from Harvard and Howland Forests at the subplot and hectare scales.	134
5.4 The SIR-C swath over eastern US.	136
5.5 Probability density functions of correlation magnitude are shown for four different land cover types, water, agricultural land, evergreen and deciduous forests.	140
5.6 Cumulative distribution functions of correlation magnitude derived from the histograms shown in Figure 5.5.	141
5.7 Coherence statistics, density (left) and distribution (right) functions of the four land-classes, plotted for two polarizations: HH-HH, and VV-VV.	142
5.8 Probability density functions of coherence obtained from pixels from the entire SIR-C swath are shown for four different land-classes and polarizations.	143
5.9 Distribution functions of coherence are shown for four land-classes and four polarizations.	144
5.10 Correlation magnitude of agricultural and forested land is plotted as a function of latitude.	144
5.11 Hourly precipitation leading up to pass 2 with zero indicating the actual time of SIR-C overpass (3am on the 10th).	147

5.12 Wind data from weather stations for the two shuttle passes over the eastern US is plotted.	149
6.1 Parameterization of the Poincaré sphere using the Deschamps parameters.	153
6.2 Illustration of the RVoG model.	157
6.3 Illustration of methods for resolution of topographic phase ambiguity.	161
6.4 Polarization synthesis and coherence optimization for UAVSAR data.	167
6.5 Examples of SNR correction with varying NESZ.	169
6.6 Example of SNR and MNR correction for a <i>dark</i> pixel.	172
6.7 Examples of straight line fits to polarimetric coherences (using points from the boundary region only) and selection of the pair of ground and volume coherence channels from four forest pixels with high SNR.	174
6.8 Aerial photograph and radar image of a region measuring roughly 1km×0.5km near Petersham, MA chosen for analysis of PolInSAR performance.	175
6.9 PolInSAR height estimates obtained from six baselines varying between 10m and 55m over a region near the Harvard Forest site.	176
6.10 Histogram of PolInSAR height estimates from six UAVSAR baselines.	177
6.11 Effect of temporal decorrelation on UAVSAR PolInSAR height estimates.	179
6.12 Comparison of PolInSAR and lidar heights.	180
6.13 Comparison of PolInSAR heights from 25, 30, 45 and 55m baselines and the LVIS RH50 metric that seems to behave most similar to the PolInSAR heights.	181

CHAPTER 1

INTRODUCTION

Climate change is widely believed to have a significant impact on Earth's ecosystems. However, we do not completely understand some critical issues in global ecosystem science. Among these issues, the most pressing include better understanding of the global carbon cycle and its influences on atmospheric green house gases, the sustainability and health of global ecosystems, and the effect of global warming and land-use change on biodiversity. The major source of uncertainty in our understanding of the global carbon cycle comes from large errors in estimates of global carbon storage, especially in vegetation. The primary source of this error is the lack of consistent, homogeneous and spatially resolved data on carbon stocks in Earth's forests. Factors, such as land-use change, further complicate the quantification of carbon budgets. While forest height and vertical structure information can be used to estimate carbon stocks, this information is also necessary for conservation efforts in biodiversity, characterization of habitats, and assessment of ecosystem health. Studies have shown forest structure to be strongly related to biodiversity and a key component in habitat selection. Similarly, forest structure metrics, such as canopy height, cover, and vertical biomass profiles, are necessary for better modeling of processes related to the physiological use of light, nutrients and water that are key indicators of ecosystem health. The development of globally consistent and spatially resolved estimates of above ground carbon biomass and vegetation vertical structure, quantification of changes in carbon budgets due to land-use change, and characterization of habitat through measurements of forest structure are key to better understanding and quan-

tification of climate change. Remote sensing instruments, such as lidar and radar, are crucial for observations on global scales with high spatial resolutions, however, such observations of carbon stocks and forest structure require highly accurate observations of vegetation. It is expected that the accuracies required are best achieved through a combination of the high precision and directness of lidar observations and the large spatial extent of radar measurements [28]. Recently much work has gone into developing algorithms for estimation of forest parameters using various radar and lidar techniques. No one technique is accepted as the most ideal means for forest structure or biomass estimation. The objective of the proposed research is to characterize the accuracy of estimates of above ground biomass from various observation methods using existing radar and lidar data with field measurements.

An imaging instrument, such as a synthetic aperture radar, is able to provide rapid and extensive spatial coverage but lacks the capability to profile vertical structures, whereas a profiling instrument, such as a lidar, can provide accurate vertical profiles but is unable to image extensive areas on short time scales. An InSAR builds on the imaging capabilities of a traditional SAR by using phase difference between two spatially displaced SAR antennas to provide an estimate of the height of the imaged area. The accuracy of height estimates derived from interferometry scales with the spatial separation (baseline) between the two antennas. An L-band interferometer is commonly regarded as the most practical space-borne mission scenario for estimating biomass and structure on global scales [28, 42]. The baselines needed for high accuracies at L-band, however, are large enough (several hundreds of meters) that a single-platform instrument becomes impractical. It is possible however to use two passes of a satellite as two ends of a baseline, this is commonly referred to as repeat-pass interferometry. Such a configuration introduces additional errors whose impact on biomass estimates must be characterized. In addition to a phase difference measurement, an InSAR system can also measure some aspects of vegeta-

tion structure from other radar variables such as polarimetric backscatter, coherence and polarimetric phase. A discussion of algorithms based on these variables is given below.

1.1 Lidar Profiles

Lidar systems obtain profiles of vertical structure by sampling the echoes of a laser pulse from various elements of a forest [88, 93]. Figure 1.1 shows the return signal from a lidar beam that contains reflections from various scatterers such as leaves of the canopy, low lying vegetation and the ground. Lidar returns are often

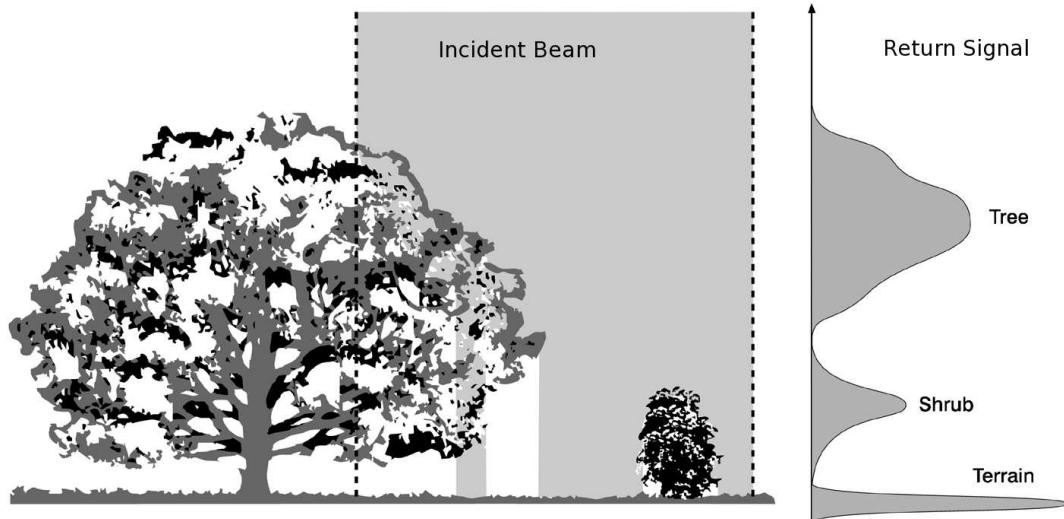


Figure 1.1: Lidar systems sample echoes of a downward transmitted pulse from various parts of a forest and creates a profile based on the time sequence of these echoes.

sampled as full waveforms (such as seen in Figure 1.1) that are a time sequence of the reflections. Some aspects of the forest structure, such as tree heights can be estimated directly from the waveforms using detection algorithms that identify ground returns and canopy tops [60, 74]. Forest biomass is not a quantity that can directly be measured by lidars, however, moments of lidar waveforms, such as mean height, canopy height, canopy extent, etc., are most often used to derive empirical regression curves with field biomass data [93, 88, 29, 61, 60, 74]. These regression

curves are dependent on biomes and predict above ground biomass with varying degrees of accuracy. In structurally homogeneous biomes, canopy height tends to be highly correlated to biomass [74, 75] and therefore tends to be a good predictor of biomass on its own, however in more complex forests other moments, such as canopy extent, canopy cover, variation in canopy metric among others are also needed [86, 29, 77, 78]. Sources of error in lidar estimates of structure and biomass include problems in the determination of ground elevation due to canopy attenuation, ground-slope effects and pointing errors among others.

1.2 Backscatter Intensity

Radar backscatter is a function of instrument parameters and properties of the imaged targets. SAR parameters such as incidence angle, wavelength and polarization determine variations in backscatter. At incidence angles close to nadir the backscatter coefficient is dominated by the specular scattering component whereas at higher incidence angles volumetric scattering is the dominant source of scattered energy. Similarly backscatter intensity of oriented structures depends on the choice of polarizations. The choice of radar wavelength affects the source of scattering and wave penetration into a volume. Target properties such as extent of the scattering volume, orientation, structure, moisture content also affect the amount of scattered energy. In the case of vegetation such as crops, at frequencies such as L-Band where the wavelength of is commensurate with the physical dimensions of the target, there is a combination of volumetric scattering from the trees or crops, surface scattering from ground and multiple bounce scattering from the interaction of soil and vegetation. The contribution of these sources to backscatter intensity is primarily determined by the extinction of the signal through the vegetation layer, extent of the vegetation and the incidence angle. Various approaches for using backscatter intensity for inferring vegetation characteristics (such as [6, 125]) model the net backscatter coefficient as a

sum of the contributions from various scattering processes, such as (simplification of [27])

$$\sigma^0 = \sigma_v^0 + \tau_v^2 (\sigma_s^0 + \sigma_{int}^0) \quad (1.1)$$

where $\sigma_v^0, \sigma_s^0, \sigma_{int}^0$ are the backscattering coefficients of the vegetation layer, the surface and the interaction of the two respectively, while τ_v is the one way transmissivity of the vegetation layer. This basic model is developed further in studies for a number of different forest types [106, 33, 124, 72, 27] and compared with biomass data collected on the ground using linear regression techniques. Similarly, studies such as [106] use (1.1) to model the vegetation layer as a cloud of water particles [125], and compare SIR-B L-band data over the Mount Shasta region of northern California. Other studies, such as [134], compare L-Band SIR-A and airborne SAR data over a pine plantation in Baldwin county, Alabama, while Le Toan et. al. [72] derive biomass estimates from airborne L-band SAR over the Landes pine forest in southwest France and Dobston et. al. [27] derive statistical relationships between SIR-C L-band data over Raco, Michigan and measurements of above ground biomass.

These studies generally develop regression power-law type relationships between σ^0 and biomass that are specific to forest types and structures. Most studies report a stronger correlation of cross-polarized data to biomass than co-polarized returns. However, a simple inversion of biomass from cross-polarized radar echoes is complicated by issues of saturation [63, 27], where σ^0 does not increase proportionally with an increase in biomass above typical values of 50 tons/hectare at L-Band. Furthermore, incidence angle variations, due to topographic change, affect the scattering mechanism or the contribution of each scattering process in (1.1) complicating the inversion process further. Among others, [27, 12] have attempted to mitigate the effect of these two issues respectively using modeling based approaches with some degree of success.

1.3 SAR Interferometry

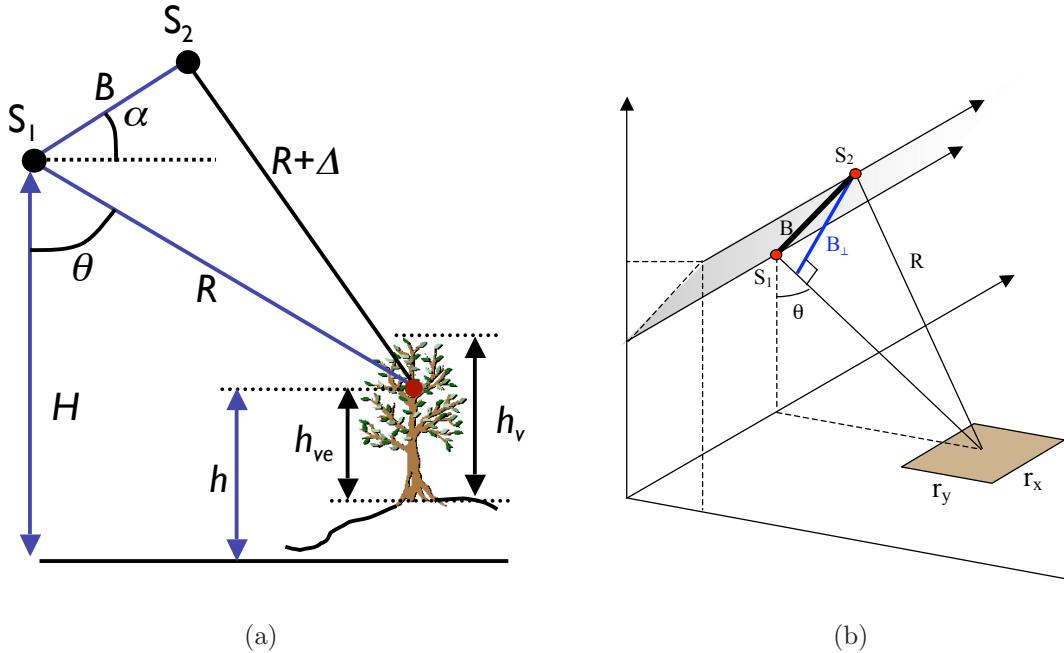


Figure 1.2: Two schematic representations of the InSAR observation geometry. The instrument is at a height H above the Earth's ellipsoid and the two antennas (S_1 and S_2) separated by a baseline of length B at an angle α from the horizon with a look angle of θ to an image pixel of dimension $r_x \times r_y$ at a slant range distance of R or $R + \Delta$.

While conventional SAR systems map the location of a target using ranging and Doppler information, they are unable to distinguish the scenery in the vertical dimension. SAR interferometry (or InSAR) builds on the imaging capability of conventional SAR systems and provides an estimate of the height of the object relative to the sensor location. In principle, this is achieved by relating the phase difference between scattered electric fields received at two antennas separated by some known distance in space, called the baseline. Figure 1.2 shows a schematic representation of the InSAR observation geometry. The first application of InSAR for Earth observations was shown using an airborne instrument over the San Juan area in Puerto-Rico [50]. This instrument mixed signals from two antennas separated by 40-wavelengths in space to form an image with fringes in amplitude that corresponded to the underlying topog-

raphy. This was improved upon through the sampling of complex signals and using the phase difference between the two channels to form a interferogram [137]. The first spaceborne demonstration of a repeat-pass interferometer (where the two passes of a satellite are used as two ends of a baseline), was reported using SeaSAT data over California [79].

The phase difference between the two channels of an interferometer, $\Delta\phi$ is related to topographic height through the following relationship

$$h = H - R \cos \left(\alpha + \sin^{-1} \left(\frac{\Delta\phi\lambda}{4\pi B} \right) \right) \quad (1.2)$$

where H is the sensor height, B the baseline, α the baseline angle from the horizon and R the range to target. As highlighted in Figure 1.2, the estimated height, h , from (1.2) is of a location above a reference ellipsoid. The vertical location of this height depends on the extinction coefficient of the scattering volume, the relative contributions of surface and volumetric scattering in a pixel and the underlying ground topography. If one were to invert this height for the that of a tree, h_v (see Figure (1.2)), one would need to subtract the topographic height of bare ground and account for the penetration of the waves into the the volume, determined by its extinction and transmissivity properties. The current standard for topographic height or a digital elevation model (DEM) was derived from a C-Band single pass interferometer from the shuttle radar topography mission (SRTM) [34]. Location of the SRTM phase center unfortunately includes height contributions of trees. If somehow bare ground topography was available, location of the phase center in the volume could be estimated as

$$h_{ve} = \frac{\Delta\phi - \phi_0}{\kappa_z} \quad (1.3)$$

where h_{ve} is height of the effective scattering center of the volume, ϕ_0 is the ground phase and κ_z , the vertical wavenumber is given by

$$\kappa_z = \frac{4\pi\Delta\theta}{\lambda \sin \theta} \approx \frac{4\pi B_{\perp}}{\lambda R \sin \theta} \quad (1.4)$$

where B_{\perp} is the component of the baseline perpendicular to the look direction (see Figure 1.2). To derive tree heights from this one would still need the knowledge of the extinction coefficient and transmissivity of the scattering volume or the trees. Furthermore, the accuracy of the phase estimates, and therefore the height estimates obtained from it, is inversely proportional to the correlation between the two channels [107]

$$\sigma_{\phi}^2 = \frac{1}{2N_L} \frac{1 - \gamma^2}{\gamma^2} \quad (1.5)$$

where N_L is the number of independent samples used to derive phase and γ is the measure of correlation between the two channels. Despite the limitations discussed above, interferometric phase has been used to estimate forest attributes. SRTM data over the mangrove forests in Florida was successfully used to generate biomass maps [113]. Forest heights, estimated by subtracting the national elevation dataset (NED) topography from SRTM elevations were related to forest biomass as well [68, 129]. Airborne interferometric heights, from the NASA/JPL (Jet Propulsion Laboratory) airborne Topographic SAR (TOPSAR), and lidar derived ground topography were used to estimate biomass successfully [114] while regression curves, developed from combinations of optical, lidar and interferometer systems have also proven useful in estimating forest biomass [61, 62].

In addition to a phase difference measurement, interferometric correlation magnitude also contains information on the structure of the target [120, 121]. This information can also be used to derive estimates of forest biomass [130, 4]. The interferometric correlation between the two channels of an InSAR system is given by [79, 108]

$$\gamma = \frac{\langle S_1 S_2^* \rangle}{\sqrt{\langle S_1 S_1^* \rangle \langle S_2 S_2^* \rangle}} \quad (1.6)$$

where $\langle \cdot \rangle$ is the expected value while S_1 and S_2 are the complex signals received at the two channels of an interferometer (see Figure 1.2). The correlation coefficient γ is affected by four different factors, namely thermal noise (γ_T), changes in viewing geometry (γ_G), volumetric scattering (γ_V) and temporal effects (γ_T) in the case of repeat-pass InSAR. These effects can be expressed according to [138, 107, 52] as

$$\gamma = \gamma_N \gamma_G \gamma_V \gamma_T \quad (1.7)$$

with the interferometric coherence, $|\gamma| \leq 1$. If the scattering volume is regarded as spatially homogeneous, the volumetric effects in correlation can be written as a Fourier transform of the radar backscatter as a function of height [107, 52, 120, 121]

$$\gamma_V = \frac{\int \sigma(z) \exp(-i\kappa_z z) dz}{\int \sigma(z) dz} \quad (1.8)$$

where γ_V , the volumetric correlation can be estimated by accounting for the other sources of decorrelation in (1.7). The vertical structure function, $\sigma(z)$, in (1.8), describes the effective radar backscatter cross section per unit height z . The particular form of $\sigma(z)$ depends on properties of the scattering medium, however, in general volumetric correlation decreases with increase in height of the scattering volume. The vertical structure function, $\sigma(z)$, can be modeled in a range of ways, the simplest of which is a constant scattering amplitude unaffected by extinction, i.e. $\sigma(z) = 1$. The modeled volumetric coherence is then simply a sinc function, given by

$$|\tilde{\gamma}_V| = \frac{\sin(\kappa_z h_v / 2)}{\kappa_z h_v / 2} \quad (1.9)$$

where h_v is the height of the volume and κ_z is the vertical wavenumber. From which height can be inverted using the approximation

$$\kappa_z h_v \ll 1 \Rightarrow |\tilde{\gamma}_V| = \frac{\sin(\kappa_z h_v / 2)}{\kappa_z h_v / 2} \approx 1 - \frac{(\kappa_z h_v / 2)^2}{6} \Rightarrow h_v \approx \sqrt{\frac{24(1 - |\tilde{\gamma}_V|)}{\kappa_z^2}}. \quad (1.10)$$

The uniform vertical structure model may however be an oversimplification. A more commonly used structure function is the so called ‘exponential model’, where the effect of extinction through the canopy is accounted for by weighing the contributions from the top of the canopy more strongly than those from the bottom. The inclusion of the extra extinction coefficient however makes the inversion problem more complicated. The Interferometric Water Cloud Model (IWCM), for example, proposed by [4, 110, 5] as an extension to the backscatter *water cloud model* [6] inverts for the stem volume using a combination of backscatter intensities and the exponential vertical structure model for interferometric coherence. This model has been used with C- and L-band repeat pass imagery over Hökmark and Kätböle, Sweden [4, 110, 5] and more recently over Siberia [111] with some success under ideal weather conditions. Algorithms that rely on InSAR coherence for forest height inversion from repeat-pass data invariably suffer from the presence of temporal decorrelation, $1 - |\gamma_T|$ [138, 4, 103, 1]. As stated earlier, while other factors such as γ_N and γ_G can be accounted for when estimating γ_V from (1.7), temporal decorrelation is almost impossible to model and eliminate from data. Presence of weather events between two acquisitions separated by even a single day can cause significant loss of coherence and therefore a misinterpretation of tree heights [1]. Stable weather conditions during the two passes of the radar become necessary for accurate estimation of forest parameters using coherence.

1.4 Polarimetric Interferometry

Polarimetric Interferometry using synthetic aperture radars, or PolInSAR, combines the structure information contained in polarimetric SAR data with the phase information obtained from an interferometric system. The technique, first presented

by [15, 16], relies on the ability of fully-polarized SAR systems to synthesize an arbitrary polarization state through linear combinations of the four principally measured polarizations (HH, HV, VH, VV), and estimate the interferometric coherence for that particular polarization channel. The spread of all the polarimetric-interferometric coherences on the complex unit circle is related to physical parameters through models, allowing for the inversion of structural attributes through a polarimetric measurement. A fully polarimetric measurement from one end of an interferometric baseline is represented by a target scattering vector, written in the Pauli basis as

$$\vec{k}_i = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{hh} + S_{vv} & S_{hh} - S_{vv} & \sqrt{2}S_{hv} \end{bmatrix}^T \quad (1.11)$$

where $i = 1, 2$ for either channels of an interferometer and S_{xx} the received field at the i^{th} channel for a specific transmit-receive linear polarization pair. The full interferometric measurement of a single-baseline fully polarimetric instrument is encapsulated in the 6×6 correlation matrix

$$T_6 = \left\langle \begin{bmatrix} \vec{k}_1 \\ \vec{k}_2 \end{bmatrix} \begin{bmatrix} \vec{k}_1^{*T} & \vec{k}_2^{*T} \end{bmatrix} \right\rangle = \begin{bmatrix} T_{11} & \Omega_{12} \\ \Omega_{12}^{*T} & T_{22} \end{bmatrix} \quad (1.12)$$

where $\langle \rangle$ represents the expected value, $T_{ii} = \langle \vec{k}_i \vec{k}_i^{*T} \rangle$ are the complex hermitian correlation matrices that describe polarimetric properties of each channel while Ω_{12} is a non-hermitian complex matrix that contains the polarimetric and interferometric information of the scattering object. In order to synthesize other polarimetric channels, one must allow the target vector \vec{k} to be projected onto a different polarization basis. This is achieved by defining two unitary ‘scattering basis’ vectors, \vec{w}_i , one for each interferometric channel and the projection of \vec{k}_i onto those basis vectors as $\vec{\mu}_i$, such that

$$\mu_i = \vec{w}_i^{*T} \vec{k}_i. \quad (1.13)$$

Finally the complex interferometric correlation for an arbitrary choice of scattering mechanisms \vec{w}_1 and \vec{w}_2 can be written as

$$\gamma(\vec{w}_1, \vec{w}_2) = \frac{\mu_1 \mu_2^*}{\sqrt{\langle \mu_1 \mu_1^* \rangle \langle \mu_2 \mu_2^* \rangle}} = \frac{\langle \vec{w}_1^{*T} \Omega_{12} \vec{w}_2 \rangle}{\sqrt{\langle \vec{w}_1^{*T} T_{11} \vec{w}_1 \rangle \langle \vec{w}_2^{*T} T_{22} \vec{w}_2 \rangle}}. \quad (1.14)$$

The key to PolInSAR is relating these correlation values to physical parameters. This is generally attempted using model based approaches, for example [95, 17] which use the random volume over ground, or RVoG, model [120, 121] to relate forest parameters such as tree height, ground elevation and extinction coefficient to a select set of polarimetric-interferometric correlation. The estimation process is written as [95]

$$\min \left\| \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix} - [M] \begin{bmatrix} h_v \\ \exp\{i\phi_0\} \\ \sigma \\ m_1 \\ m_2 \\ m_3 \end{bmatrix}^T \right\| \quad (1.15)$$

where ϕ_0 is the topographic phase, h_v and σ are the height and extinction coefficients of the scattering volume and the model M is given by

$$\tilde{\gamma} = \exp(i\phi_0) \frac{\tilde{\gamma}_v + m(\vec{w})}{1 + m(\vec{w})} \quad (1.16)$$

where $\tilde{\gamma}_v$ is the volumetric correlation alone and $m(\vec{w})$ is the polarization dependent effective ground to volume return ratio given by

$$m(\vec{w}) = \frac{m_g(\vec{w})}{m_v(\vec{w})} \exp\left(-\frac{2\sigma h_v}{\cos \theta}\right) \quad (1.17)$$

where m_g and m_v are the ground and volume scattering amplitudes and θ is the angle of incidence. The volumetric correlation $\tilde{\gamma}_v$, is often modeled as

$$\tilde{\gamma}_v = \frac{I}{I_0} \begin{cases} I &= \int_0^{hv} \exp\left(\frac{2\sigma z'}{\cos \theta}\right) \exp(i\kappa_z z') dz' \\ I_0 &= \int_0^{hv} \exp\left(\frac{2\sigma z'}{\cos \theta}\right) dz' \end{cases} \quad (1.18)$$

which is an extension of (1.8) using the exponential structure function for $\sigma(z)$.

Equations (1.16) through (1.18) describe the RVoG model used in (1.15) for forest parameter inversion from PolInSAR data. The combination of polarimetric and interferometric information leads to six measurements (3 independent complex correlations), while the RVoG model is also a six parameter model, leading to the possibility of a critically determined inversion problem (i.e. the number of unknowns equal the number of observations). The height inversion capabilities of this technique have recently been demonstrated in Boreal forests using L-band data [95] and over tropical forests [53] using L and X-band data with promising results. It has also been used to derive biomass estimates using height-biomass allometry [87]. Among the limitations of this technique includes the presence of temporal decorrelation in repeat-pass PolInSAR data [17] that unless accounted for would result in an overestimation of tree heights. One such mechanism is proposed by [96] where a revised RVoG model is presented that incorporates the effects of temporal decorrelation.

The following chapters are devoted to the analysis of biomass estimates and their accuracy from each of the four techniques described above, using remote sensing data from lidar and radar instruments and comparing them to ground validation data. Chapter 2 discusses the field campaign to collect ground truth data in the form of diameter and species information over two research forests in the North-Eastern United States, the Harvard and Howland Forests. An attempt is made to characterize the accuracy of biomass estimates from diameter measurements. Chapter 3 uses full-waveform lidar data from a campaign of the NASA/GSFC LVIS instrument over the

two sites in summer 2009. Most commonly used metrics and their combinations are used to ascertain the lidar metric that best predicts forest biomass. Chapter 4 analyzes the relationship between forest biomass and radar backscatter using airborne SAR data from the NASA/JPL UAVSAR instrument, that flew over the Harvard and Howland Forests in 2009 as well. The accuracy of biomass estimates from these backscatter measurements are ascertained with a backscatter error model. Chapter 5 looks closely at temporal decorrelation, the most unpredictable contributor to errors in estimates of height and biomass from repeat-pass interferometric measurements, complicating any parametric analysis of the InSAR-biomass relationship. Finally, Chapter 6 looks at PolInSAR results from multiple UAVSAR baselines near the Harvard Forest test site from the 2009 UAVSAR campaign, and the effect of temporal variations on PolInSAR height estimates.

CHAPTER 2

FOREST MENSURATION, ERRORS AND FIELD CAMPAIGNS

In a statistical sense, the only way to measure forest biomass without error would require cutting down all trees, drying and weighing them one by one. For obvious reasons this is not common practice. Felling a representative sample and weighing them is somewhat more common, however as the geographic scale increases it becomes prohibitively expensive and time consuming. It is common practice then to try and relate some aspect of a tree's structure to its biomass using dimensional analysis or allometry so that forest biomass can be estimated from much more practical measurements of structure.

2.1 Allometry

Dimensional analysis, or allometry refers to the relationships between certain elements of a natural object's size and shape. In forestry, for example, the diameter and volume of a tree are related. This allows for prediction of a tree's volume, and by association its mass, through a much simpler and practical measurement of its diameter. In simple Euclidean terms the volume of an object is proportional to a product of its diameter, D , and height, H

$$V \propto D^2 H. \quad (2.1)$$

However, most natural objects such as trees are not well described by simple Euclidean shapes especially because of complex structures such as tree crowns. The use of fractal

geometry (suggested in [84]) provides a more realistic alternative. Various studies have demonstrated the usefulness of this approach in relating tree diameter to crown dimensions in particular ([94, 139]) and to the overall structure of trees in general, so that the relationship between volume and a diameter-height product is given by

$$V \propto D^\alpha H^\beta \quad (2.2)$$

with both α, β positive and generally regarded to be bounded by $2 < \alpha + \beta < 3$. To take this a step further, analyses such as in [92] use biomechanics to report that height scales as a function of diameter, such that

$$H \propto D^\gamma \quad (2.3)$$

with $0 < \gamma \leq 1$. So equation 2.2 becomes

$$V \propto D^\alpha D^{\beta\gamma} = \epsilon D^{\alpha+\beta\gamma}. \quad (2.4)$$

with ϵ as the proportionality constant. Since mass, (or biomass when talking of trees) is a product of density (ρ) and volume, the total above ground biomass of a tree, M , can be written as

$$M = \rho\epsilon D^{\alpha+\beta\gamma} = aD^b. \quad (2.5)$$

In general a theoretical value of around $8/3$ has been suggested for the coefficient b [131]. In practice, both a and b have been shown to vary with tree species and ecological conditions among others. Whenever possible these coefficients are empirically determined for the various species encountered in a particular forest and documented in the form of weight tables.

2.1.1 Methods for developing allometric relationships

Interest in determining the above ground biomass in forests has led to a fairly large body of studies where weight tables have been developed and documented for different species in various biomes (i.e. ecological regions). However, there is no one standard method for generating allometric equations, in fact it is fairly common to use height as an independent variable along with diameter. In general the independent variables, such as diameter (at a height of 1.37m above the ground [7]), are measured for a representative sample of trees. These trees are then felled and separated into different components such as the stem, bark, branches and foliage. The fresh weight of each component is measured. Since the intent is to determine dry biomass the components are dried in ovens. However, it is impractical to dry an entire large tree. Instead different sampling methodologies are used. The stem (or bole) is cut into smaller pieces (1 to 2m in length) and fresh weight of each of the smaller sections is recorded. Discs (of a few centimeters in length) are cut from each section, labeled, weighed and dried in ovens. The dry weight of the discs is measured and the dry weight of the stem from which they were cut is estimated using weight-ratio type methods. Bark weight is obtained in a similar manner. Some studies account for stump weights by cutting the tree very close to the ground; those studies that do not, use similar weight ratio methods to estimate stump weight as well. To estimate the weight of a tree crown, most studies adopt some type of a stratified sampling approach. Approaches such as this involve cutting branches into sections of a certain size and separating them into classes or strata based on branch diameter, measured at some distance from the base. Randomly chosen branches from each diameter class are chosen for drying. Branches that belong to larger diameter classes are weighed much like the stem sections whereas smaller branches, including foliage bearing ones are dried intact. The entire crown weight is estimated using some form of class-dependent mean weight ratio estimator. Typically tree diameter, or dbh (diameter at

breast height - 1.37m [7]), is regressed to the dry weight of each component and to the projected total dry weight. In most cases, the projected total weight of the tree is different from the sum of its components. However, some recent studies (such as [97]) have outlined more statistically sound methods for using the sum of components for estimating total biomass. In either case, component or total weights are related to dbh using regression techniques. Since the variation of tree biomass is heteroscedastic, that is to say the variation increases with increasing diameter, the use of simple linear regression becomes complicated. Traditionally, this problem is circumvented by taking the logarithm of (2.5), such that

$$\log M = \log a + b \log D \quad (2.6)$$

and using linear regression to estimate $\log a$ and b . This solves the problems of heteroscedasticity, however the conversion from logarithmic back to arithmetic units causes a bias in the mean estimated weight. To correct for this artifact, Baskerville [8] suggested the following correction based on the effect of logarithms on probability distributions. If $\mu = \log M$ then

$$M_c = \exp(\mu + \sigma_{se}^2/2) \quad (2.7)$$

where M_c is the corrected mean weight and σ_{se} is the standard error in logarithmic units. The factor $e^{\sigma_{se}^2/2}$ is usually referred to as the bias correction factor and is published by most studies, however there is contention that this correction itself is biased for small sample sizes [65] so it isn't always published or used. In more recent studies, such as [97] and [71] the problem of heteroscedasticity is accounted for by modeling variance and using more sophisticated regression techniques.

2.1.2 Using existing allometric equations

Since it is rarely feasible to develop diameter-biomass allometries for a particular area of interest, it is common to utilize already developed allometric equations. Because of the large amount of studies that document such weight tables, it becomes important to be able to correctly identify the most representative equations. Typically, studies either focus on one, or a select few, species across multiple biomes or regions (e.g. [24] [70] [69] [112]), for multiple species that belong to a particular region or ecosystem (e.g. [136] [132] [10] [44] [133]), or literature that focuses on summarizing multiple studies (e.g. [118] [123] [65] [64]). These examples are not meant to be exhaustive, in fact hardly so, since these studies easily number in the hundreds. It is beyond the scope of this work to summarize all existing equations, however it would be remiss not to look at more than one of the studies that summarizes multiple field campaigns. Three studies are chosen for comparison here that are the key comprehensive compilations for Eastern USA temperate forests. Even though these studies summarize coefficients for the same power-law type diameter-biomass relationships of (2.5), they approach the analysis in distinct ways. The first, in Ter-Mikaelian [118] lists species-specific coefficients developed by other studies, the second in Jenkins et. al [65] develops new coefficients for species grouped into general categories using coefficients from previous field campaigns, and finally Lambert et. al [71] list species-specific coefficients calculated using raw data in a more rigorous statistical framework allowing a more accurate assessment of error. A short discussion on the three studies follows.

2.1.2.1 Single site allometry: Ter-Mikaelian [118]

The work by Ter-Mikaelian and Korzukhin [118] summarizes equations for sixty five North-American tree species from multiple studies conducted in the Northeastern United States. All coefficients are reproduced or recalculated for allometric equations

of the form given in (2.5). The original objective of this study¹ was to identify the reasons behind the observed variation between different allometric coefficients for species common to this region. The compilation of different allometric equations was a byproduct of this effort. A consequence of the intent to conduct a quantitative comparison was the documentation of standard error for most of the equations. This was a major reason for selecting this study over previously established works such as Tritton and Hornbeck [123] that provide a similar analysis.

In the development of allometric equations, there is no established rationale for which set of coefficients to choose for a particular species. It is common to use coefficients based on proximity of where the equation was developed and where it is to be applied. However that is not always the only consideration since large biases can be introduced if equations are used beyond the range of diameter values that were used for developing the original regressions. Furthermore, in some cases, species specific equations may not even be available, and applying coefficients developed for another species may cause large errors as well. It is hard to perfectly ascertain the amount of error introduced by using equations for different species or even non-site specific equations.

2.1.2.2 Ensemble allometry: Jenkins et. al. [65]

The study by Jenkins et. al from 2003 [65] attempts to rectify the spatial variability among allometric equations and the incompleteness of studies, such as [118], by compiling equations from all over the continental United States. This study aims to develop *generalizable* equations that would be applicable for a large set of species across varied biomes. It adopts a *meta-analysis* approach, as described in [98], for combining results from studies that may have used varying methodologies and standards. In short, it involves generating *pseudo-data* from published equations and

¹personal communication with Michael Ter-Mikaelian

combining all the *pseudo-data* to generate new regression coefficients. Here, instead of having species specific coefficients like in Ter-Mikaelian [118], species were categorized into eleven groups based on similarities in structure and allometric coefficients, among others. The allometric coefficients generated in this study are all for the same power-law relationship given in (2.5). Although the authors are meticulous in categorizing a plethora of species and careful to include a wide range of diameters, the drawbacks include the potential for introducing biases from non-species and non-site specific equations. Furthermore, estimates of standard error, derived from the *pseudo-data*, are sub-optimal and may be dominated by spatial variability in the diameter-biomass relationship.

2.1.2.3 BLUE allometry: Lambert et. al. [71]

The study by Lambert et. al. [71] attempts a more statistically rigorous approach to generate BLUE (best linear unbiased estimators) of biomass. The approach, based on methods proposed in [97], provides species specific equations generated by fitting raw diameter and biomass data collected over many sites across Canada. The use of raw data, instead of the pseudo-data approach used in Jenkins et. al. [65], allows for a more rigorous characterization of error. The approach outlined in [97] is a departure from the standard approaches in two distinct ways. First, it does not use the logarithms of diameter and biomass to circumvent the problem of heteroscedasticity. Instead it models the variance in a power-law sense, much like the diameter-biomass relationship itself. Secondly, it recognizes the possibility of correlation between component biomasses themselves. That is to say that the process of separating a tree into its components such as stem, bark, canopy, etc. to estimate its dry weight, may have errors that are not independent, and thus cannot simply be added. The method outlines an approach known as *Seemingly Unrelated Regression* or SUR to account for both of these factors so that the sum of component biomass estimates can be used

to generate estimates of total biomass. Furthermore, it allows the use of a variance-covariance matrix to account for correlations between components when estimating the total error. The coefficients for the biomass equations are different from (2.5), instead the estimation equation takes the form

$$M_{stem} = a_{stem} D^{b_{stem}} \quad (2.8)$$

$$M_{bark} = a_{bark} D^{b_{bark}} \quad (2.9)$$

$$M_{foliage} = a_{foliage} D^{b_{foliage}} \quad (2.10)$$

$$M_{branch} = a_{branch} D^{b_{branch}} \quad (2.11)$$

$$M_{total} = M_{stem} + M_{bark} + M_{foliage} + M_{branch} \quad (2.12)$$

where M_i is the dry biomass for the i^{th} component defined here as either stem, bark, foliage, branches or total, D is the diameter at breast height (dbh) and a_i , b_i are regression coefficients for a particular component.

2.2 Error Propagation Analysis

A slew of factors determine the error in estimating the biomass of a tree from a measurement of its diameter. These factors include (but are not limited to) [65] error in measurement of diameter, error in estimating coefficients of the allometric equations, the use of those coefficients outside of the species and ecosystem from which they were derived and inconsistencies in methodologies between different studies. It is not possible to perfectly account for every error source, however, an analysis of the error sources that can be characterized in some mathematical framework is necessary. Since the sources of error may depend on the choice of allometry, a treatment for the three allometries chosen for this work is presented.

2.2.1 Errors in single site allometry

Three potential error sources in forest biomass estimates are considered for this type of allometry: biomass error due to an error in measurement of tree diameter (σ_m), error in determining allometric coefficients (σ_a) and errors in using these allometric coefficients across different sites (σ_s). Assuming that the equations are chosen properly, these three error sources should account for most of the error in biomass estimates [13]. It is assumed that the three sources of error are independent, so that the total error in the estimates of tree weights can be written as

$$\sigma_t = \sqrt{\sigma_m^2 + \sigma_a^2 + \sigma_s^2}. \quad (2.13)$$

The error in *dbh* measurement can be propagated to an error in tree weight by using a Taylor series expansion of the allometric model, as suggested in [13]. Since the chosen allometric model consists of only diameter as the independent variable, the error in the above ground biomass can be written as a function of the measurement error as

$$\frac{\sigma_m^2}{M} = \frac{\sigma_D^2}{D^2} \left(\frac{\partial \ln(f)}{\partial \ln(D)} \right)^2 \quad (2.14)$$

where σ_D is the uncertainty in the measurement of diameter (D) and $\partial \ln(f) / \partial \ln(D)$ is the partial derivate of the natural log of the allometric model function, f , with respect to the natural log of the diameter. Since the allometric model function is of the form $f = aD^b$, the measurement error, σ_m , is given by

$$\sigma_m = M b \frac{\sigma_D}{D}. \quad (2.15)$$

The error in estimating model coefficients, σ_a depends on the correlation between diameter and weight data used in a particular study. Most studies report the standard error (or the root sum of squares of the fit residuals), σ_{se} , as a means of estimating

σ_a . Since standard errors are dependent on the methodology chosen for fitting diameter and weight, it is not simply equal to σ_a . Most studies summarized in [118] use log-transformed variables for regression and the standard error is also reported in logarithmic units. In such cases, σ_{se} must be transformed into arithmetic units. This is not as simple as using the inverse-logarithm since the statistics of random variables are skewed during this transformation. Baskerville [8] suggested the following conversion to estimate the allometric error from standard error in logarithmic units

$$\sigma_a^2 = \exp(2\sigma_{se}^2 + 2\mu) - \exp(\sigma_{se}^2 + 2\mu) \quad (2.16)$$

where μ is the logarithm of estimated biomass, i.e. $\mu = \log M$. Some studies use base-10 logarithms, in such cases a similar correction is used but with the corresponding anti-log function. A few studies in [118] provide standard error in arithmetic units, in those cases it is assumed that $\sigma_a = \sigma_{se}$.

The third component of error, σ_s (site error), captures the error in biomass estimates introduced by using coefficients developed at a site different to which the equation is later applied. No mathematical treatment of this error exists in forestry literature even though it is widely recognized as a potential uncertainty primarily driven by soil conditions and climate [25, 64]. However, an imperfect estimate of σ_s can be obtained by employing a *bootstrap* type approach [11]. The single-stage bootstrap is a technique from the non-parametric class of methods used in statistics for arriving at estimates of variation (or error) in data. It relies on resampling the data and using the spread of mean values from the various combinations of the resampled data to estimate the variance of errors. The approach used here takes advantage of the fact that the list of studies in [118] at times includes more than one set of allometric coefficients developed for a particular species by different researchers in different locations. In particular, this study lists nine sets of coefficients for red maples (*Acer rubrum*) and seven for paper birch (*Betula papyrifera*). Sites for these studies extend

from West Virginia in the south to Nova Scotia in the north, in essence sampling a range of ecoregions that may exist in the north-eastern United States. Multiple biomass estimates can be generated for a particular diameter by using various combinations of these equations. In fact, if all combinations are used, 126 biomass values can be obtained using combinations of five equations from a total of nine for Red Maple, and 35 for Paper Birch using combinations of four equations from a total of seven, providing enough samples for a crude estimate of the variance in allometric coefficients due to varying soil conditions etc.

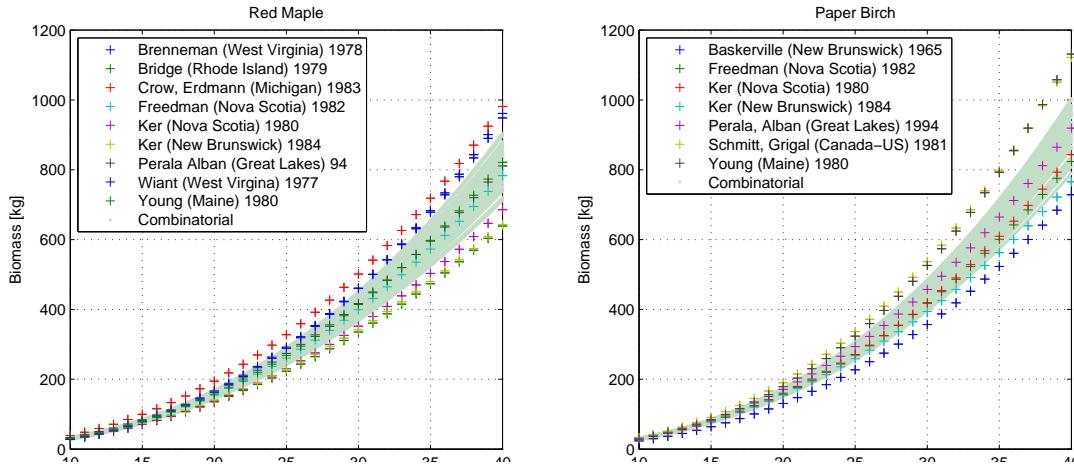


Figure 2.1: Variation between mean predicted values from the various allometric equations summarized in [118] for Red Maple (*Acer rubrum*) and Paper Birch (*Betula papyrifera*).

Figure 2.1 shows biomass estimates from the nine Red Maple equations and the seven Paper Birch equations and mean values from combinations of these equations. A sample standard deviation from the full set of combinations is computed and quadratic fits to diameter values are used as an estimate of σ_s .

To obtain uncertainties in estimates of forest biomass at some spatial scale, σ_{sp} , the per-tree errors (σ_t) are aggregated over the area of interest (assuming that the individual tree errors are uncorrelated), such that

$$\sigma_{sp}^2 = \sum_{i=1}^N \sigma_{t_i}^2 \quad (2.17)$$

where N is the total number of trees for the particular area in question, and $\sigma_{t_i}^2$ is the total error in tree weight for the i^{th} tree.

2.2.2 Error in ensemble equations

The treatment of error in ensemble equations is similar to that described in Section 2.2.1 with the exception of the site-error σ_s . The equations developed in Jenkins et. al [65] are derived from studies that encompass all of the continental United States. The standard error for each group of equations includes variability due to site conditions and captures the intrinsic variability between the species grouped together. The standard error is therefore expected to be much larger. However, due to the *meta-data* methodology used in this study, the standard errors are biased lower and not entirely reliable [71, 65]. In the absence of a better method of estimating variance, the standard error reported in [65] is used as an estimate of biomass uncertainty using the corrections described in (2.16), so that the total error is given by

$$\sigma_t = \sqrt{\sigma_m^2 + \sigma_a^2} \quad (2.18)$$

with σ_m estimated using equation 2.15.

2.2.3 Error in BLUE equations

The procedure outlined in Parresol [97] is more mathematically involved than those traditionally employed. Instead of relying on standard error estimates from regressions between diameter data and an estimate of tree weight, this method attempts to include the error in estimated tree weight as well as the regression error. Because the method estimates total tree weight as a function of the estimated component weights, with the addition of components used in Lambert et. al. [71] as one

particular functional form, the correlations of errors among the components cannot simply be ignored. This is accounted for by using variance-covariance matrices for the coefficients of the component equations using a statistical framework known as *NSUR* (non-linear seemingly unrelated regression). Furthermore, the variance of component weights is not constant over all observations either (heteroscedasticity). Instead of relying on the traditional approach of using logarithms to avoid the problem of heteroscedasticity, here the regression error for each component equation, e_i , is modeled such that if

$$M_i = a_i D^{b_i} + e_i \quad (2.19)$$

then e_i is functionally related to the diameter using

$$\ln e_i^2 = b + c_i \ln D. \quad (2.20)$$

In the above formulation, coefficient c_i is estimated by fitting diameter to the residuals e_i for each component and also the total biomass. The allometric error in total biomass estimates, σ_a is given by

$$\sigma_a^2 = S_{M_t}^2 + \sigma_{SUR}^2 \sigma_{ii} \psi_t(D) \quad (2.21)$$

where σ_{SUR}^2 is the SUR system variance or the residual sum of squares from the multiple non-linear regression analysis, $S_{M_t}^2$ is the estimated variance in total biomass due to errors in estimating the coefficients of component biomass equations and σ_{ii} is the residual root sum of squares from the particular equation of interest (in this case that of total biomass). The term $\psi_t(D)$ refers to the function that models heteroscedasticity, which takes the form of (2.20), i.e. $\psi_i(D) = D^{c_i}$. Lambert et. al.

publish coefficients c_i in [71] for each of the component equations including the total biomass. The variance of total biomass, $S_{M_t}^2$, is estimated using

$$S_{M_t}^2 = \mathbf{F}'_{ab} \hat{\Sigma}_{\mathbf{ab}} \mathbf{F}_{ab} \quad (2.22)$$

where $\hat{\Sigma}_{\mathbf{ab}}$ is the estimated variance-covariance matrix of the set of coefficients a_i, b_i . This matrix is estimated using the raw tree weight and diameter data. Lambert et. al [71] calculate and distribute² these matrices for each species. These are 8×8 matrices, for two parameters of each of the four component equations, that is, four a_i and four b_i rows/columns. The vector \mathbf{F}_{ab} is a row vector of the derivatives of the *model function* with respect to the fit parameters. Since the *model function*, given in (2.12), is a summation of components, the vector \mathbf{F}_{ab} can be calculated using

$$\mathbf{F}_{a_i b_i} = \frac{\partial}{\partial(a_i, b_i)} \left(\sum_{i=1}^4 a_i D^{b_i} \right) \quad (2.23)$$

where \mathbf{F}_{ab} is an 8×1 row vector. All parameters needed to estimate the total variance from this allometric analysis used in (2.21) through (2.23) are published in Lambert et. al [71].

The site-specific error, alluded to in Section 2.2.1, is not included in this analysis because the variability modeled by (2.21) is shown [71] to encompass the variability between the different equations summarized in [118]. The error due to measurement in diameter however needs to be propagated through the model function. This takes a slightly different form than (2.15) and can be estimated using the Taylor series approximation of the model function given by

$$\sigma_m^2 = \left(\frac{\partial M_{total}}{\partial D} \right)^2 \sigma_D^2 \quad (2.24)$$

²these matrices are not published in the article [71], they were obtained through personal communication

where M_{total} is given by (2.12). This can be simplified to show easily that

$$\sigma_m = \frac{\sigma_D}{D} \left(\sum_{i=1}^4 a_i b_i D^{b_i} \right) \quad (2.25)$$

where σ_D is the error in measuring tree diameter. The total error in estimating the tree biomass is given by

$$\sigma_t = \sqrt{\sigma_m^2 + \sigma_a^2} \quad (2.26)$$

where σ_m and σ_a are given by (2.25) and (2.21) respectively.

2.3 The Harvard Forest

The Harvard Forest near Petersham, MA is an ecological research facility that has been managed by the Harvard University since 1907. It is spread over 3000 hectares and is split mainly into three tracts: Prospect Hill, Tom Swamp and Slab City (see Figure 2.2). The forest is representative of the *Transition Hardwoods* of central New England [116]. The dominant species are red oak (*Quercus rubra*), red maple (*Acer rubrum*), white birch (*Betula papyrifera*), white pine (*Pinus strobus*) and eastern hemlock (*Tsuga canadensis*). Most of the forest is artificially planted over reclaimed agricultural land which occurred in the first half of the twentieth century [35]. Stands of a various species are maintained throughout the facility. Permanent study sites are spread over the forest where research is conducted in many topics of study such as biodiversity, conservation, forest-atmosphere carbon exchange and soil warming to name a few.

2.3.1 Harvard Forest field campaign

During summer of 2009 ground validation data was collected from fifteen hectares in the Harvard Forest. The survey area was divided into fifteen one hectare plots, with a plot measuring 200m by 50m. Each plot was further divided into 16 subplots,

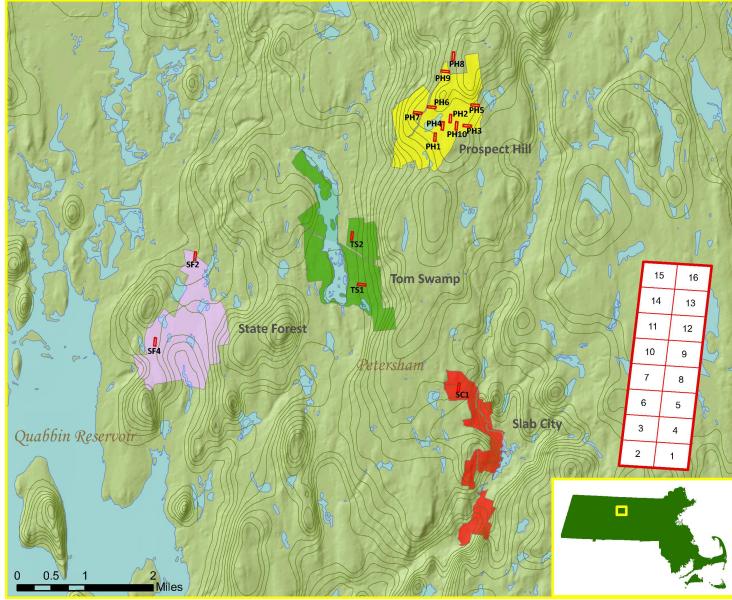


Figure 2.2: Harvard Forest tracts and plots. The inset shows a 1-hectare plot with its sixteen 25m by 25m subplots numbered one through sixteen.

each 25m by 25m. The orientation of each plot was chosen to be either 5 degrees for vertical plots, or 95 degrees for horizontal plots. The majority of these plots were set in the three Harvard Forest tracts of Prospect Hill, Tom Swamp and Slab City. Of the 15 plots, ten were in Prospect Hill (titled PH01 to PH10), two in Tom Swamp (TS01, TS02) and one in Slab City (SC01). The remaining two plots were set in the nearby State Forest (SF02 and SF04). Figure 2.2 shows the fifteen plot locations, their boundaries and the three Harvard Forest tracts. The inset in Figure 2.2 shows the sixteen subplots and the number scheme that was used to identify each subplot within the larger one-hectare plot. The choice of location of these plots was guided by forest species composition, age/structure, topography, accessibility and lidar/radar coverage. For instance plots PH1 and PH7 are set inside stands of predominantly tall and dense red pine trees, PH4 and PH2 are set inside a Hemlock stand, PH6 and PH10 are predominantly young deciduous trees etc. A GPS location of one corner of each plot was recorded. Corners for all other subplots were surveyed using a measuring

tape and compass. For each of the 240 subplots, diameter, species and condition (live or dead) for each tree larger than 10cm was cataloged. The GPS locations of the plot corners were re-surveyed in the summer of 2010 in an effort to assess the accuracy of the location of the ground plots. Because of the thick canopy cover at the Harvard Forest, an accuracy of better than 4m was hard to achieve.

2.3.2 Allometric equations for the Harvard Forest

Table 2.1 summarizes field diameter data for the 23 species and site-specific allometric equations used for each to estimate tree weights. All parameters and statistics listed here are for allometric equations of the form given in (2.5) with diameters expressed in units of centimeters and estimated weight in units of kilograms. In Table 2.1, the entries labeled Min D, Max D and NT, summarize the diameter data collected over the Harvard Forest. Min D and Max D are the minimum and maximum diameter recorded for a total of NT trees of the corresponding species. The table entry listed as ‘Type’ refers to the component biomass, in this case either AB: total above ground biomass or ST: stem wood + bark weight. Ideally equations that are regressed to total above ground biomass should be used, however such equations are not always available, in which case the ST equations are used instead. Since ST equations do not consider branch weights, biomass estimates obtained from those coefficients would be considerably lower and a potential source of bias, therefore such coefficients are used only sparingly to avoid large errors. Over the Harvard Forest data, ST equations are used only once for striped maples (*Acer pennsylvanica*) that are less than 0.3% of the total tree count and even less as a percentage of the total biomass, thus minimizing the bias in subplot or plot scale biomass estimates.

Table 2.1: Summary of diameter data for the 23 species catalogued at the Harvard Forest and the single-site allometric equations chosen from Ter-Mikaelian and Korzukhin[118]

Species	Min D	Max D	NT	Type	a	b	Range/Over	R ²	SEE	MTD	CF	N	Region	Author
Striped maple <i>Acer pennsylvanicum</i>	9.14	17.78	23	ST	0.0839	2.2300	(1-8)/23	0.987	0.279	ln	1.040	8	New Hampshire	Hocker and Early, 1983 [57]
Red maple <i>Acer rubrum</i>	8.12	69.21	2850	AB	0.1789	2.2334	(10-52)/23	0.98	0.116	ln	1.007	150	Michigan, Wisconsin	Crow and Erdmann, 1983 [24]
Sugar maple <i>Acer saccharum</i>	10.66	45.21	31	AB	0.1599	2.3376	(1-41)/1	0.993	0.141	ln	1.010	45	New Brunswick	Ker, 1980 [70]
Yellow birch <i>Betula alleghaneensis</i>	10.166	53.85	234	AB	0.1684	2.4150	(1-55)/0	0.994	0.099	log	1.011	14	New Hampshire	Whittaker, 1974 [132]
Black birch <i>Betula lenta</i>	9.652	48.514	305	AB	0.0629	2.6606	(5-50)/0	0.990	0.0133	log	1.002	8	West Virginia	Brenneman, 1978 [10]
Paper birch <i>Betula papyrifera</i>	8.89	81.65	499	AB	0.1074	2.4313	(3-33)/16	0.99	0.141	ln	1.010	45	Nova Scotia	Ker, 1980 [69]
Grey birch <i>Betula populifolia</i>	9.5	9.5	1	AB	0.1218	2.3123	(1-23)/0	0.99	0.141	ln	1.010	44	Nova Scotia	Ker, 1980 [69]
American chestnut <i>Castanea dentata</i>	10.414	18.542	7	AB	0.1599	2.3376	(1-41)/0	0.993	0.141	ln	1.010	45	New Brunswick	Ker, 1980 [70]
American beech <i>Fagus grandifolia</i>	8.89	71.95	130	AB	0.1957	2.3916	(1-60)/2	0.994	0.089	log	1.009	14	New Hampshire	Whittaker, 1974 [132]
White ash <i>Fraxinus americana</i>	3.175	59.56	78	AB	0.1535	2.3213	(1-28)/25	0.992	0.141	ln	1.010	46	New Brunswick	Ker, 1980 [70]
Tamarack <i>Larix laricina</i>	41.14	44.19	2	AB	0.1599	2.3376	(1-41)/0	0.993	0.141	ln	1.010	45	New Brunswick	Ker, 1980 [70]
Hophornbeam <i>Ostrya virginiana</i>	13.46	13.46	1	AB	0.1074	2.4313	(3-33)/16	0.99	0.141	ln	1.010	45	Nova Scotia	Ker, 1980 [69]
Red pine <i>Pinus resinosa</i>	4.318	53.361	1230	AB	0.0778	2.4171	(3-46)/11	0.993	0.111	ln	1.006	69	Upper Great Lakes	Perala, Alban, 1994 [100]
Red spruce <i>Picea rubens</i>	9.906	2.451	103	AB	0.2066	2.1830	(1-35)/9	0.982	0.107	log	1.013	15	New Hampshire	Whittaker, 1974 [132]

Species	Min D	Max D	NT	Type	a	b	Range/Over	R ²	SEE	MTD	CF	N	Region	Author
White pine <i>Pinus strobus</i>	8.89	111.76	947	AB	0.0696	2.4490	(3-66)/28	0.99	0.2092	ln	1.022	35	Maine	Young, 1980 [136]
Bigtooth aspen <i>Populus grandidentata</i>	10.033	40.132	9	AB	0.0983	2.3373	(1-34)/2	0.99	0.156	ln	1.012	30	Nova Scotia	Freedman, 1982 [39]
Pin cherry <i>Prunus pensylvanica</i>	8.89	10.795	6	AB	0.1556	2.1948	(3-24)/0	0.99	0.372	ln	1.071	30	Maine	Young, 1980 [136]
Black cherry <i>Prunus serotina</i>	8.89	48.260	197	AB	0.1225	2.4253	(5-40)/11	0.99	20.41	abs	n/a	19	West Virginia	Wiant, 1977 [133]
White oak <i>Quercus alba</i>	8.89	57.15	187	AB	0.0472	2.7010	(5-40)/18	0.986	32.66	abs	n/a	19	West Virginia	Wiant, 1977 [133]
Red oak <i>Quercus rubra</i>	7.112	136.906	2102	AB	0.0643	2.6598	(5-40)/302	0.988	35.87	abs	n/a	19	West Virginia	Wiant, 1977 [133]
Black oak <i>Quercus velutina</i>	11.684	51.81	16	AB	0.0945	2.5030	(5-40)/1	0.99	22.68	abs	n/a	19	West Virginia	Wiant, 1977 [133]
American basswood <i>Tilia americana</i>	13.71	13.71	1	AB	0.0617	2.5328	(5-50)/0	0.96	0	abs	n/a	13	West Virginia	Brenneman, 1978 [10]
Eastern hemlock <i>Tsuga canadensis</i>	4.572	89.916	1473	AB	0.0991	2.3617	(3-51)/129	n/a	0.130	ln	1.0085	36	Maine	Young, 1980 [136]

The fields a , b are the coefficients of the allometric equation. The term ‘Range’ in field ‘Range/Over’ refers to the range of diameter values that either the chosen study reported or was estimated by [118] and ‘Over’ refers to the number of trees in the diameter data that exceed this limit. Fields ‘MTD’, ‘ R^2 ’, ‘SEE’, ‘N’ are representative statistics highlighting the performance of the regression between tree diameter and biomass. The abbreviation, ‘MTD’ refers to the method used in fitting the two variables. Studies chosen in this analysis either use log-transformed (\ln : \log_e , \log : \log_{10}) data for linear regression or use weighted-non-linear (abs) regression. R^2 , the coefficient of determination, is reported by most studies and summarized [118]. The SEE (standard error in estimate) is either reported by the studies or calculated from the data summaries. It is listed in units based on the fit methodology, i.e. either log, \ln or abs. The parameter ‘N’ refers to the number of samples (trees) used in the regression of tree diameter and weight data, while ‘CF’ is the correction factor suggested in [8] to correct for biases caused by the conversion between arithmetic and logarithmic units. Finally the last two fields in Table 2.1 refer to the location of the study site and the authors that documented the field campaign in the corresponding cited article. Ideally there would be coefficients and statistics for every species, however, that is not the case. The SEE and CF for black birch (*Betula velutina*) were estimated from the summary in Tritton and Hornbeck [123], coefficients for American chestnut (*Castanea dentata*) and tamrack (*Larix laricina*) were not summarized, so the coefficients from sugar maple (*Acer saccharum*) were used instead, as suggested in [44] and [64] respectively. Similarly, coefficients for paper birch (*Betula papyrifera*) were used for hophornbeam (*Ostrya virginiana*). The SEE statistic for total weight was not available for red pine (*Pinus resinosa*) so the standard error of the ‘stem-bark’ equation was used. The study does not summarize the error from any of the equations from Young et. al. [136]. In those cases, SEE was estimated from the data summary in [136].

Table 2.2: The ensemble equation coefficients from Jenkins et. al [65] with the twenty three species at the Harvard Forest grouped into eight categories.

Tree/Shrub species	<i>a</i>	<i>b</i>	Max Dbh	R ²	SEE
Aspen/Alder/Willow	0.1098	2.3867	70	0.95	0.507441
	Bigtooth aspen (<i>Populus grandidentata</i>)				
Soft Maple/Birch	0.1477	2.3651	66	0.96	0.491685
	Striped maple (<i>Acer pennsylvanicum</i>), red maple (<i>Acer rubrum</i>), yellow birch (<i>Betula alleghaniensis</i>), black birch (<i>Betula lenta</i>), paper birch (<i>Betula papyrifera</i>), gray birch (<i>Betula populifolia</i>)				
Mixed Hardwood	0.0837	2.4835	56	0.98	0.360458
	American chestnut (<i>Castanea dentata</i>), white ash (<i>Fraxinus americana</i>), hophornbeam (<i>Ostrya virginiana</i>), pin cherry (<i>Prunus pensylvanica</i>), black cherry (<i>Prunus serotina</i>), American basswood (<i>Tilia americana</i>)				
Hard maple/Oak/Beech	0.1336	2.4342	73	0.99	0.236483
	Sugar maple (<i>Acer saccharum</i>), American beech (<i>Fagus grandifolia</i>), white oak (<i>Quercus alba</i>), red oak (<i>Quercus rubra</i>), black oak (<i>Quercus velutina</i>)				
Cedar/ Larch	0.1309	2.2592	250	0.98	0.294574
	Tamarack (<i>Larix laricina</i>)				
True fir/ Hemlock	0.0790	2.4814	230	0.99	0.182329
	Eastern hemlock (<i>Tsuga canadensis</i>)				
Pine	0.0792	2.4349	180	0.98	0.253781
	Red pine (<i>Pinus resinosa</i>), white pine (<i>Pinus strobus</i>)				
Spruce	0.1253	2.3323	250	0.98	0.250424
	Red spruce (<i>Picea rubens</i>)				

Table 2.2 summarizes the coefficients from the ensemble equations from Jenkins et. al. [65]. Both coefficients a, b are for allometric equations of the form given in (2.5). This study divides most species seen in the United States into eleven broad categories. All twenty three species seen at the Harvard Forest fall into eight of those eleven groups. Since this study pays attention to the range of diameter values that the coefficients are estimated over, especially towards the higher end, not many trees measured at the Harvard Forest exceed the limits of these equations.

Table 2.3 summarizes the coefficients of the diameter-biomass equations from Lambert et. al. [71] of the form given in (2.8) through (2.12). The coefficients for the four component equations are listed as a_i, b_i for the i^{th} component (stem, bark, branches or foliage, respectively). The error parameters σ_{SUR} and σ_{tt} are the SUR system variance and error in the total biomass equation as described in Section 2.2.3. The field $\psi_t(D)$ refers to the coefficient of the function that models heteroscedasticity of error in the total biomass equation, i.e. coefficient ‘ c ’ in $\psi_t(D) = D^c$. The fields ‘Range’ refers to the range of diameter values the equations were regressed over and ‘N’ is the number of samples or trees of each species used in the regressions. For species whose equations could not be found, coefficients for a general softwood (gymnosperm) or hardwood (angiosperm) equations were used.

Table 2.3: Summary of coefficients from the BLUE allometric equations of Lambert et. al. [71] for the twenty three species catalogued at the Harvard Forest

Species	a_{stem}	b_{stem}	a_{bark}	b_{bark}	$a_{branches}$	$b_{branches}$	$a_{foliage}$	$b_{foliage}$	σ_{SUR}	σ_{tt}	$\psi_t (D)$	Range	N
Stripe maple (<i>Acer pennsylvanicum</i>)	0.0871	2.3702	0.0241	2.1969	0.0167	2.4807	0.0390	1.6629	1.0000	0.0373	3.63	-	-
Red maple (<i>Acer rubrum</i>)	0.1014	2.3448	0.0291	2.0893	0.0175	2.4846	0.0515	1.5198	1.0006	0.0050	4.25	2.4-56.0	177
Sugar maple (<i>Acer saccharum</i>)	0.1315	2.3129	0.0631	1.9241	0.0330	2.3741	0.0393	1.6930	1.0002	0.0828	3.28	2.0-57.8	235
Yellow birch (<i>Betula alleghaneensis</i>)	0.1932	2.1569	0.0192	2.2475	0.0305	2.4404	0.1119	1.3973	1.0003	2.8971	2.34	0.8-70.3	280
Black birch (<i>Betula lenta</i>)	0.1754	2.1616	0.0381	2.0991	0.0085	2.7790	0.0373	1.6740	1.0075	0.0344	3.64	1.1-55.3	117
Paper birch (<i>Betula papyrifera</i>)	0.0593	2.5026	0.0135	2.4053	0.0135	2.5532	0.0546	1.6351	1.0001	0.0067	3.81	1.5-53.6	606
Grey birch (<i>Betula populifolia</i>)	0.0720	2.3885	0.0168	2.2569	0.0088	2.5689	0.0099	1.8985	2.2562	0.0016	3.92	2.2-22.7	43
American chestnut (<i>Castanea dentata</i>)	0.0871	2.3702	0.0241	2.1969	0.0167	2.4807	0.0390	1.6629	1.0000	0.0373	3.63	-	-
American beech (<i>Fagus grandifolia</i>)	0.1478	2.2986	0.0120	2.2388	0.0370	2.3680	0.0376	1.6164	1.0102	0.4720	2.92	1.8-46.3	177
White ash (<i>Fraxinus americana</i>)	0.1861	2.1665	0.0406	1.9946	0.0461	2.2291	0.1106	1.2277	1.0053	0.0363	3.62	2.4-53.7	109
Tamrack (<i>Larix laricina</i>)	0.0625	2.4475	0.0174	2.1109	0.0196	2.2652	0.0801	1.4875	1.0003	0.0052	3.88	1.8-44.5	575
Hophornbeam (<i>Ostrya virginiana</i>)	0.1929	1.9672	0.0671	1.5911	0.0278	2.1336	0.0293	1.9502	0.7228	0.0679	2.73	5.2-18.5	14
Red pine (<i>Pinus resinosa</i>)	0.0564	2.4465	0.0188	2.0572	0.0033	2.7515	0.0212	2.0690	1.0005	0.0052	3.95	1.3-55.1	371
Red spruce (<i>Picea rubens</i>)	0.0989	2.2814	0.0220	2.0908	0.0005	3.2750	0.0066	2.4213	1.0956	0.0059	3.64	6.5-45.3	55
White pine (<i>Pinus strobus</i>)	0.0997	2.2709	0.0192	2.2038	0.0056	2.6011	0.0284	1.7166	1.0016	0.3142	2.76	1.5-68.7	199
Bigtooth aspen (<i>Populus grandidentata</i>)	0.0959	2.3430	0.0308	2.2240	0.0047	2.6530	0.0080	2.0149	1.0026	0.3645	2.35	1.9-39.2	100
Pin cherry (<i>Prunus pensylvanica</i>)	0.0871	2.3702	0.0241	2.1969	0.0167	2.4807	0.0390	1.6629	1.0000	0.0373	3.63	-	-
Black cherry (<i>Prunus serotina</i>)	0.3743	1.9406	0.0679	1.8377	0.0796	2.0103	0.0840	1.2319	1.0599	60.766	1.52	0.9-49.6	78
White oak (<i>Quercus alba</i>)	0.0762	2.3335	0.0338	1.9845	0.0113	2.6211	0.0188	1.7881	1.0043	0.0011	4.70	2.2-74.3	61
Red oak (<i>Quercus rubra</i>)	0.1754	2.1616	0.0381	2.0991	0.0085	2.7790	0.0373	1.6740	1.0075	0.0344	3.64	1.1-55.3	117
Black oak (<i>Quercus velutina</i>)	0.0871	2.3702	0.0241	2.1969	0.0167	2.4807	0.0390	1.6629	1.0000	0.0373	3.63	-	-
American basswood (<i>Tilia americana</i>)	0.0562	2.4102	0.0302	2.0976	0.0230	2.2382	0.0288	1.6378	1.0101	0.1755	2.91	3.7-54.8	80
Eastern hemlock (<i>Tsuga canadensis</i>)	0.0619	2.2381	0.0139	2.2382	0.0217	2.2653	0.0776	1.6995	1.0005	0.0156	3.60	1.3-51.4	235

2.3.3 Comparison of biomass estimates and errors

Individual tree weights were estimated over the Harvard Forest using coefficients listed in Tables 2.1, 2.2 and 2.3 with the appropriate biomass and error equations. Figure 2.3 shows the mean tree weights (in units of kilograms) for four of the major tree species encountered at the Harvard Forest as a function of the diameters measured during the Harvard Forest field campaign. Error bars, using 95% confidence intervals are also shown. Since a majority (90 percent) of the measured diameters are less than 40cm, the diameter range in Figure 2.3 is truncated at 40cm. In order to calculate the error in estimated tree weight a diameter measurement accuracy, a σ_D of 2% was chosen [47]. The species-specific variance-covariance matrices needed in (2.22) are given in Appendix A. As expected the mean tree weights of softwoods (red pines

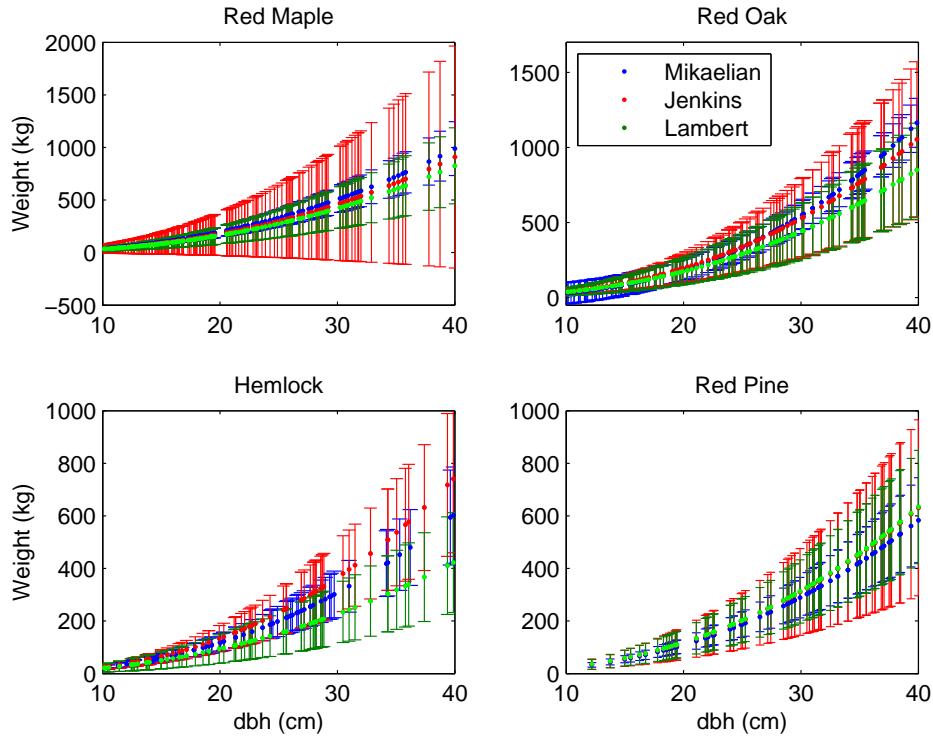


Figure 2.3: Estimated tree weights with 95% confidence intervals for the four major tree species at the Harvard Forest.

and hemlocks) are lower than the mean tree weights of hardwoods (red maples and red oaks). Red maples have lower specific wood density so they tend to have lower biomass values than hardwoods such as oaks. This is noticeable in Figure 2.3 with the weights of red maples only slightly larger than those of hemlocks and pines of the same size. Even though red maples are more numerous in the Harvard Forest dataset, red oaks have the most biomass of all species catalogued. The more noticeable trend in Figure 2.3 is the large uncertainty in tree weight estimates. In all four cases shown here, the error exceeds 100% of the mean tree weight for larger diameters. The mean biomass and error estimates from the BLUE allometric equations tend to be less than or equal to estimates from the other two allometric equations for most species.

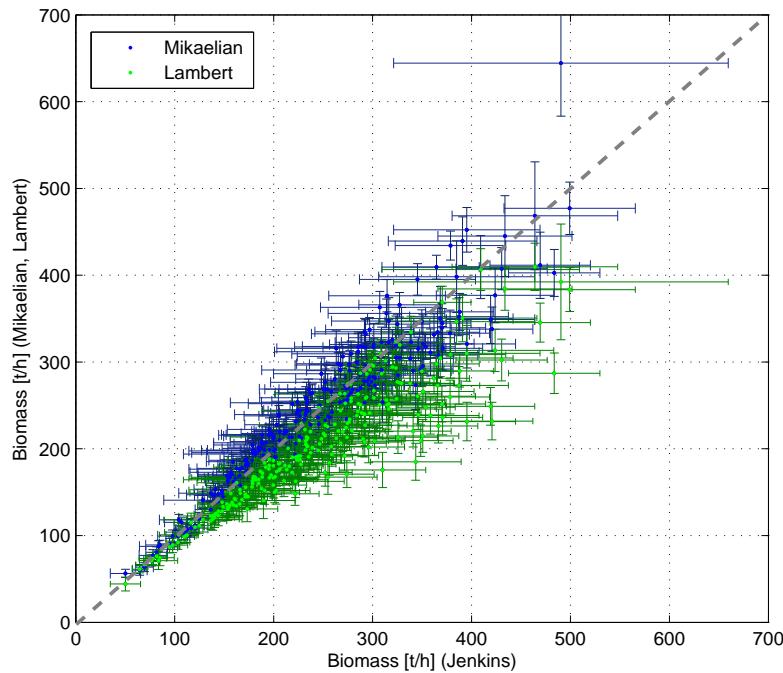


Figure 2.4: Comparison of subplot-level biomass estimated using the Site-specific (Ter-Mikaelian [118]), Ensemble (Jenkins [65]) and BLUE (Lambert [71]) allometries.

These per-tree biomass estimates were added for trees that belong to a particular subplot to generate spatial estimates of biomass. To convert the units of these subplot biomass estimates to tons/hectare the mean values were multiplied by a scaling factor of 16/1000. The biomass errors for the subplots, obtained by using the root sum of squares of the individual tree errors, were also multiplied by the same scaling factor to convert the errors into units of tons/hectare. Figure 2.4 compares the mean biomass estimates for the 240 subplots obtained using the three different allometries. Estimates from the Ensemble allometry (Jenkins [65]) are plotted against subplot-level biomass estimates from the Single-site (Ter-Mikaelian [118]) and BLUE (Lambert [71]) allometries. The errors in biomass estimates from each of the allometries are shown as error-bars of widths corresponding to 95% confidence intervals. The mean estimates from the Single-site and Ensemble equations are consistent, however estimates from the BLUE equations are consistently lower. The subplot-level biomass estimates over the Harvard Forest range from 50tons/ha to 500tons/ha with mean values of roughly 200tons/ha from BLUE and 250tons/ha from the other two allometries.

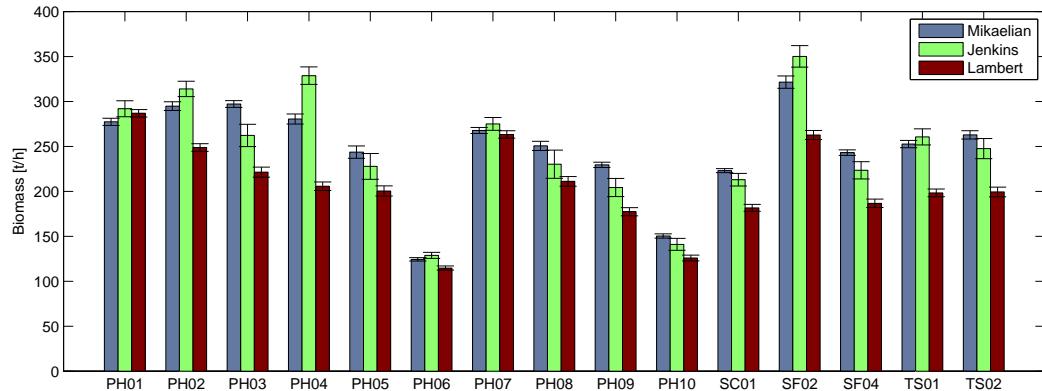


Figure 2.5: Hectare level biomass estimates using Ensemble (Jenkins [65]), Single-site (Ter-Mikaelian [118]) and BLUE (Lambert [71]) allometries.

Figure 2.5 shows the per-tree biomass estimates aggregated for the corresponding plots to generate hectare-level biomass estimates from the three allometric equations. Estimates of errors are shown as error-bars using 95% confidence intervals. The biomass values for the fifteen hectares over the Harvard Forest range from 115 to 350 tons/hectare. The confidence intervals for these hectare-level biomass estimates are fairly narrow, with mean errors of 2tons/ha for BLUE and Single-site allometries and slightly higher errors of 4tons/ha for estimates using the Ensemble allometry.

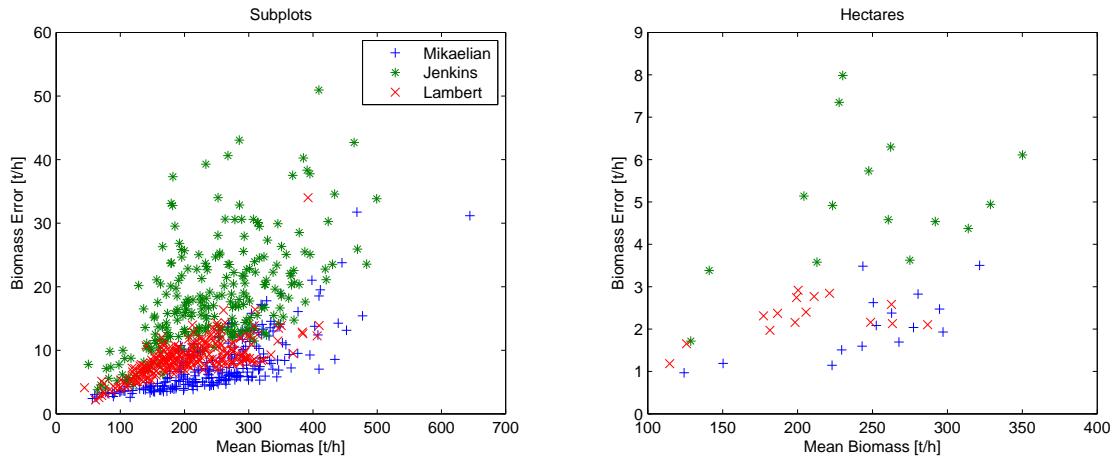


Figure 2.6: Subplot and plot level biomass error as a function of mean biomass from the Ensemble (Jenkins [65]), Single-site (Ter-Mikaelian [118]) and BLUE (Lambert [71]) allometries.

Figure 2.6 shows a comparison of the estimated errors as a function of mean biomass at hectares and subplots for the three sets of allometric equations. The error from the Ensemble equations is higher with a mean value of 18.8tons/ha at subplots and 4.95tons/ha at hectares. The mean error for both the Single-site and BLUE allometries is roughly the same at 2tons/ha and 7tons/ha for hectare and subplot level estimates respectively. The biomass error at subplots has a strong dependence on the mean biomass value for either allometry. Errors from the Ensemble equations also display an increase in their variance as a function of the mean biomass values at both spatial scales. This effect is less noticeable for the other two allometries.

2.4 Howland Forest, Maine

The Howland Forest research facility, managed by the University of Maine since 1989, is spread over roughly 500 acres in central Maine near the town of Howland, 35 miles north of Bangor. Figure 2.7 shows the two tracts, one near the town of Howland and the other near Penobscott, where plots, of similar dimensions as those established at the Harvard Forest, were laid out and surveyed. The forest, a boreal-northern hardwood transitional forest, consists mainly of spruce, fir, hemlock, pines and maples. The topography of the region is generally flat and the field sites are also laid out in low relief areas.

2.4.1 Howland Forest field campaign

At the Howland Forest eleven one-hectare plots were surveyed during the summer of 2009 as part of the field campaigns over the north east region of the US (which the Harvard Forest was also part of). Data from twelve such sites was also collected at Penobscott. A total of twenty eight species were encountered during the two field campaigns. Of those, eight account for roughly 90% of the total tree count and biomass. The field data consisted of species information, diameters and condition over one hectare plots with dimensions of $200\text{m} \times 50\text{m}$, comprising of sixteen $25\text{m} \times 25\text{m}$ subplots. The field campaigns over Howland and Harvard were part of the same DESDynI data collection effort so the methodology over the two sites were deliberately similar.

2.4.2 Allometric equations for the Howland Forest

Of the twenty eight species catalogued at the Howland Forest, seventeen were encountered at the Harvard Forest as well. For species common to both datasets, allometric equations summarized in Section 2.3.2 were used to generate biomass and error estimates over the Howland Forest.

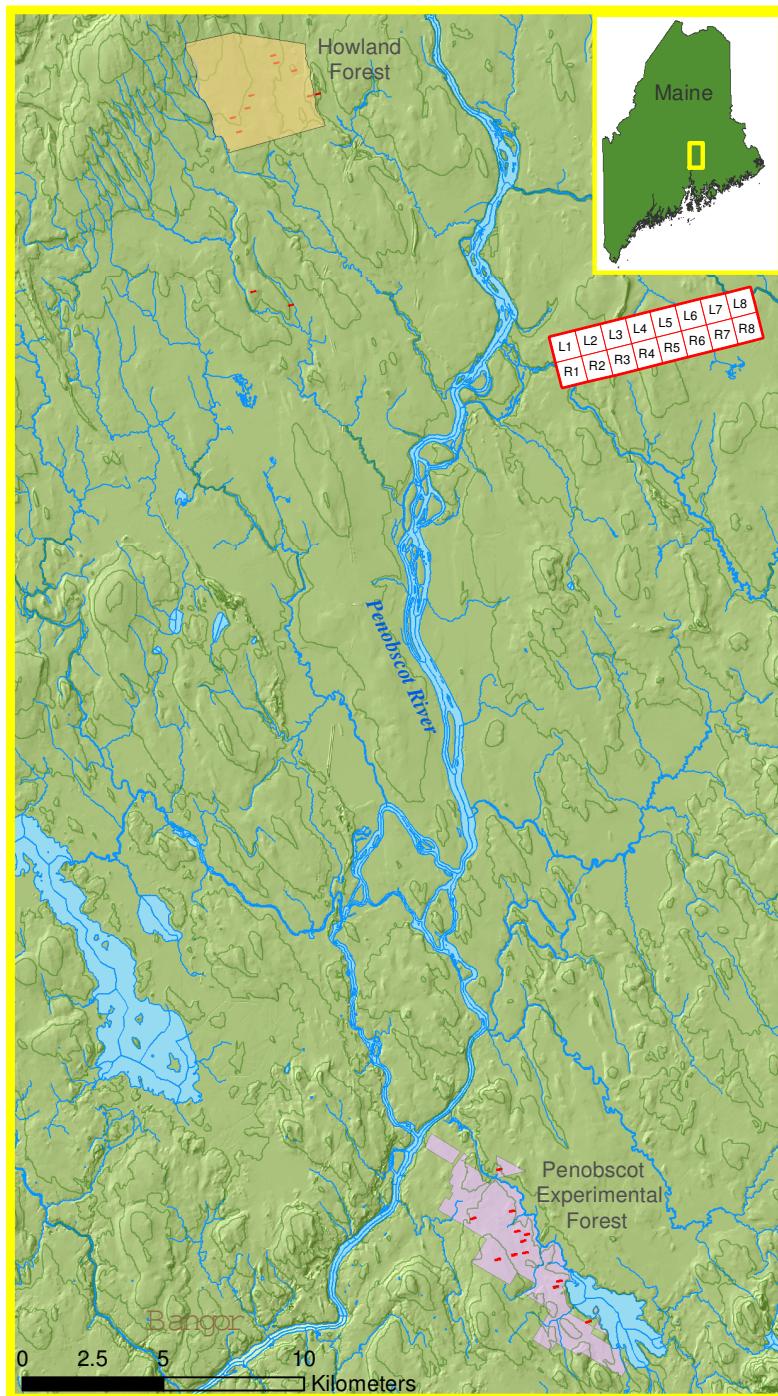


Figure 2.7: Howland Forest Research facility includes two sites, one near the town of Howland and the other near Penobscott in central Maine.

For the species that were only seen at the Howland Forest and not the Harvard Forest, coefficients from the single-site (Ter-Mikaelian [118]) are summarized in Table 2.4. Coefficients for green ash (*Fraxinus pennsylvanica*) were not summarized in [118] therefore those of black ash (*Fraxinus nigra*) were used. Similarly, coefficients from bigtooth aspen (*Populus grandidentata*) were used for balsam poplar (*Populus balsamifera*). The species for a number of trees were left undetermined. In such cases the coefficients for balsam fir (*Abies balsamea*) were used.

The eleven species unique to the Howland Forest dataset belong to the same eight categories for the Ensemble equations as summarized in Table 2.2. The balsam fir (*Abies balsamea*) belongs to the *True Fir/Hemlock* group, mountain maple (*Acer spicatum*) to the *Soft Maple/Birch* category, while black ash (*Fraxinus nigra*), green ash (*Fraxinus pennsylvanica*) and American elm (*Ulmus americana*) are classified as *Mixed Hardwoods* in [65]. For the two spruce species, Norway and black (*Picea abies*, *Picea mariana*) coefficients from the *Spruce* group were used, while coefficients from the *Cedar/Larch* group were used for white cedar (*Thuja occidentalis*). Both balsam poplar (*Populus balsamifera*) and trembling aspen (*Populus tremuloides*) are grouped into the *Aspen/Adler/Willow* category. Biomass for the unidentified trees was estimated using the *Mixed Hardwood* equation.

Allometric coefficients form Lambert et. al [71] for the species unique to the Howland forest are summarized in Table 2.5. Here, coefficients for most of the species were available, except for mountain maple (*Acer spicatum*), green ash (*Fraxinus pennsylvanica*) and Norway spruce (*Picea abies*). For the first two coefficients from the general hardwood equations were used, while the coefficients from the softwood equation was used for Norway spruce. The biomass of all the unidentified trees was estimated using a general hardwood equation as well.

Table 2.4: Summary of diameter data for the eleven species unique to the Howland Forest dataset and the corresponding single-site allometric equations chosen from Ter-Mikaelian [118]

Species	Min D	Max D	NT	Type	a	b	Range/Over	R ²	SEE	MTD	CF	N	Region	Author
Balsam fir <i>Abies balsamea</i>	8.6	40	3102	AB	0.0690	2.4975	(3-40)/0	0.970	0.123	ln	1.008	40	Ontario	Honer, 1971 [59]
Mountain maple <i>Acer spicatum</i>	11.0	23.4	6	AB	0.2040	2.2524	(1-20)/3	0.990	0.074	log	1.006	15	New Hampshire	Whittaker, 1974 [132]
Black ash <i>Fraxinus nigra</i>	17.5	17.5	1	AB	0.1634	2.3480	(4-32)/0	0.953	0.348	ln	1.062	18	Upper Great Lakes	Perala, Alban, 1994 [100]
Green ash <i>Fraxinus pennsylvanica</i>	17.5	17.5	1	AB	0.1634	2.3480	(4-32)/0	0.953	0.348	ln	1.062	18	Upper Great Lakes	Perala, Alban, 1994 [100]
Norway spruce <i>Picea abies</i>	10.6	40.0	47	AB	0.2722	2.1040	(12-44)/0	0.960	0.0152	ln	1.012	30	New York	Jokela, 1986 [66]
Black spruce <i>Picea mariana</i>	10.0	27.00	94	AB	0.1683	2.1777	(2-34)/0	0.99	0.199	ln	1.020	49	Nova Scotia	Ker, 1980 [69]
Balsam poplar <i>Populus balsamifera</i>	41.8	45.2	2	AB	0.0687	2.5153	(5-40)/2	0.99	1.780	abs	1.00	19	West Virginia	Wiant, 1977 [133]
Trembling aspen <i>Populus tremuloides</i>	7	58	191	AB	0.2065	2.2490	(15-40)/25	0.988	0.037	log	1.002	191	Wisconsin	Pastor, Bockheim, 1981 [99]
White cedar <i>Thuja occidentalis</i>	10.0	30	745	AB	0.1148	2.1439	(2-30)/82	0.991	0.141	ln	1.010	46	New Brunswick	Ker, 1980 [70]
American elm <i>Ulmus americana</i>	13.0	22.7	6	AB	0.0825	2.4680	(4-29)/0	0.991	0.1418	ln	1.011	14	Upper Great Lakes	Perala, Alban, 1994 [100]

Table 2.5: Summary of coefficients from the BLUE allometric equations of Lambert et. al. [71] for the eleven species catalogued only at the Howland Forest

Species	a_{stem}	b_{stem}	a_{bark}	b_{bark}	$a_{branches}$	$b_{branches}$	$a_{foliage}$	$b_{foliage}$	σ_{SUR}	σ_{tt}	$\psi_t (D)$	Range	N
Balsam fir (<i>Abies balsamea</i>)	0.0534	2.4030	0.0115	2.3484	0.0070	2.5406	0.0840	1.6695	1.0053	0.0047	3.63	1.5-42.4	639
Mountain maple (<i>Acer spicatum</i>)	0.0871	2.3702	0.0241	2.1969	0.0167	2.4807	0.0390	1.6629	1.0000	0.0373	3.63	-	-
Black ash (<i>Fraxinus nigra</i>)	0.0941	2.3491	0.0323	2.0761	0.0448	1.9771	0.0538	1.3584	1.0029	0.0085	3.82	2.0-43.1	73
Green ash (<i>Fraxinus pennsylvanica</i>)	0.0871	2.3702	0.0241	2.1969	0.0167	2.4807	0.0390	1.6629	1.0000	0.0373	3.63	-	-
Norway spruce (<i>Picea abies</i>)	0.0648	2.3927	0.0162	2.1959	0.0156	2.2916	0.0861	1.6261	1.0000	0.0065	3.82	-	-
Black spruce (<i>Picea mariana</i>)	0.0477	2.5147	0.0153	2.2429	0.0278	2.0839	0.1648	1.4143	1.0005	0.0091	3.55	1.6-37.2	1534
Balsam poplar (<i>Populus balsamea</i>)	0.0510	2.4529	0.0297	2.1131	0.0120	2.4165	0.0276	1.6215	1.0032	0.0049	4.02	2.0-53.2	20
Trembling aspen (<i>Populus tremuloides</i>)	0.0605	2.4750	0.0168	2.3949	0.0080	2.5214	0.0261	1.6304	1.0000	0.0199	3.54	0.7-47.2	773
White cedar (<i>Thuja occidentalis</i>)	0.0654	2.2121	0.0114	2.1432	0.0335	1.9367	0.0499	1.7278	1.0009	0.0315	2.96	2.1-66.2	184
American elm (<i>Ulmus americana</i>)	0.0402	2.5804	0.0073	2.4895	0.0401	2.1826	0.0750	1.3436	1.3426	0.0337	3.28	0.7-55.2	81

2.4.3 Comparison of biomass estimates and errors

Biomass estimates for the 23 plots and corresponding 368 subplots at the Howland Forest were estimated by aggregating the tree weights obtained using the three different allometric equations. Unlike the Harvard Forest dataset, a larger number of low-biomass sites were sampled during the field campaigns over the Howland Forest. Figure 2.8 shows the comparison of subplot-level biomass estimates from the three allometric equations; Ensemble (Jenkins [65]), Single-Site (Ter-Mikaelian [118]) and BLUE (Lambert [71]). Errors in these biomass estimates are shown as error-bars using 95% confidence intervals.

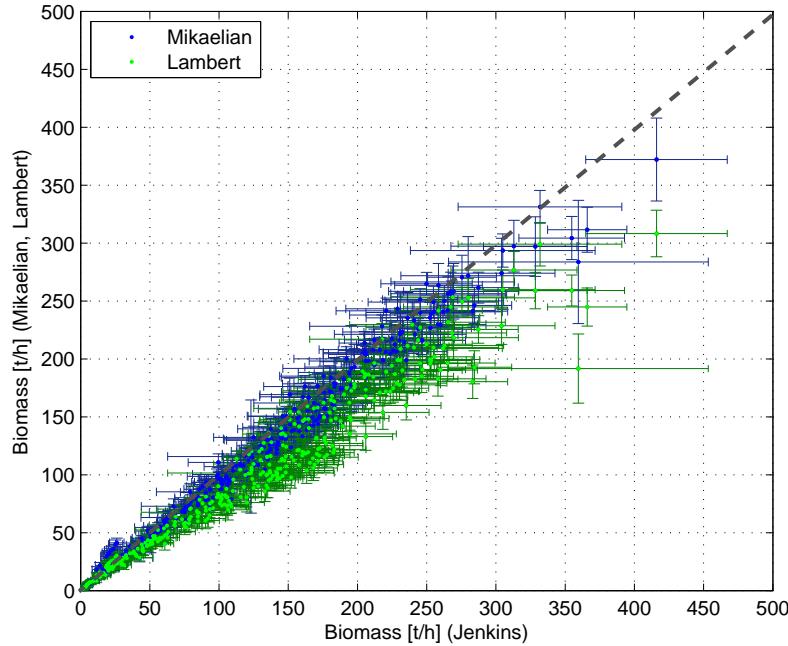


Figure 2.8: Comparison of subplot-level biomass estimated using the Site-specific (Ter-Mikaelian [118]), Ensemble (Jenkins [65]) and BLUE (Lambert [71]) allometries over the Howland Forest.

As was the case with the Harvard Forest estimates, the BLUE allometry (shown here in green) consistently underestimates the biomass values compared to the Ensemble equations. Mean biomass estimates from single-site and ensemble equations,

however, are similar. The range of biomass values at the Howland Forest at subplot-scales is similar to the Harvard Forest (of roughly 450tons/ha) however there are more subplots with fairly low biomass values (less than 50tons/ha) that are not present in the Harvard Forest dataset.

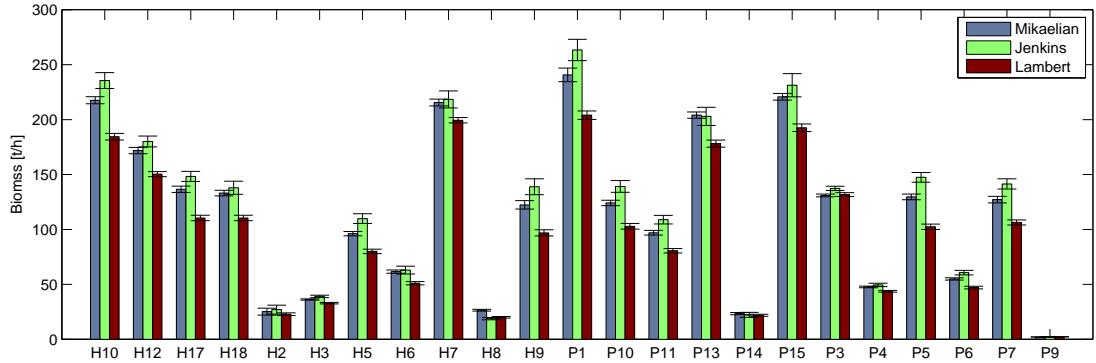


Figure 2.9: Comparison of hectare-level biomass estimated using the Site-specific (Ter-Mikaelian [118]), Ensemble (Jenkins [65]) and BLUE (Lambert [71]) allometries over the Howland Forest.

Figure 2.9 shows the biomass estimates for the twenty three hectares at the Howland Forest obtained from the three different allometric equations. The biomass estimates from these sites, with the descriptor H for Howland, and P for Penobscott, range from close to zero (plot P9) to about 270tons/ha (plot P1). The high biomass values at Howland are not as high as those at the Harvard Forest (of up to 350tons/ha) owing primarily due to the larger number of hardwoods at the Harvard Forest site. However, the low biomass sites, such as P9 at Howland, are much lower than any found at the Harvard Forest, primarily because those are near clear-cut regions with very few small trees (about 77 trees in total with diameters ranging from 6 to 13 cm). Such sites were not surveyed at the Harvard Forest.

Figure 2.10 shows the error estimates from the three allometries as a function of mean estimated biomass at the two spatial scales (subplots and hectares). These errors, much like those for the Harvard Forest increase as a function of the mean

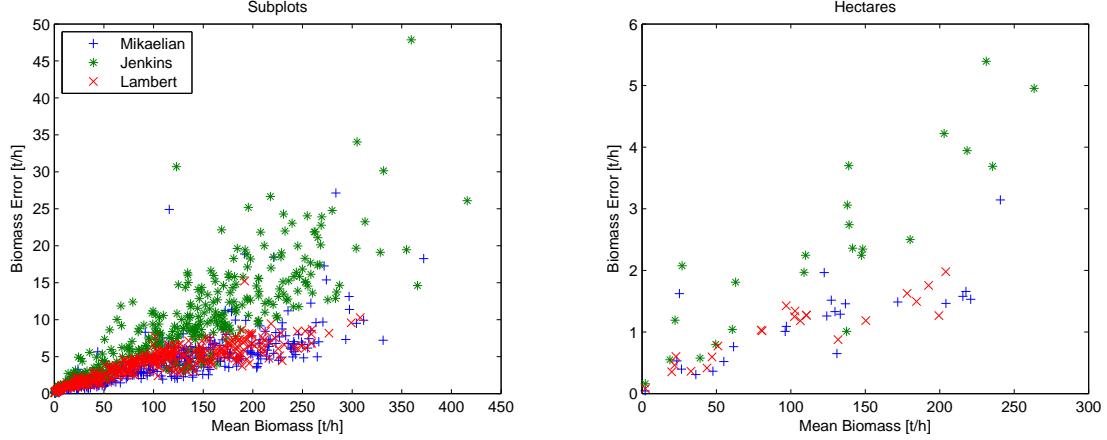


Figure 2.10: Subplot and plot level biomass error as a function of mean biomass from the Ensemble (Jenkins [65]), Single-site (Ter-Mikaelian and Korzhukhin [118]) and BLUE (Lambert [71]) allometries over the Howland Forest.

biomass value for either the subplots or the hectares. The variance of the errors from the Ensemble equations increase a function of the mean biomass as well. The mean value of errors from the Ensemble equations are highest at 8.7tons/ha and 2.3tons/ha for subplots and hectares respectively. Much like Harvard Forest, both the Single-site and BLUE allometries yield lower and consistent mean error values of roughly 1ton/ha and 4 tons/ha for hectares and subplots.

Table 2.6 summarizes statistics for errors in biomass estimates at subplot- and hectare-scales at the Howland and Harvard Forests. The statistics in terms of the mean, the maximum and the minimum for each of the three allometries are expressed in units of tons/ha. The errors in biomass estimates at the Howland Forest are lower in general than the errors at the Harvard Forest. This difference can primarily be attributed to the lower biomass estimates at the Howland Forest. Errors from the Ensemble allometry seem to be largest among the three allometries, with errors in estimates from the BLUE allometry consistently the lowest.

Table 2.6: Comparison of biomass errors from the three allometries at the two study sites. All errors are listed in units of tons/ha.

		Subplots			Hectares		
		Single-site	Ensemble	BLUE	Single-site	Ensemble	BLUE
Harvard	Mean	7.740	18.578	8.828	2.096	4.950	2.286
	Min	2.415	3.799	2.171	0.970	1.716	1.183
	Max	31.735	86.245	33.992	3.503	7.984	2.917
Howland	Mean	4.179	8.786	3.947	1.173	2.373	1.027
	Min	0.099	0.213	0.102	0.043	0.160	0.079
	Max	27.138	47.858	15.220	3.143	5.395	1.979

CHAPTER 3

LIDAR MEASUREMENT AND ERRORS

Airborne or spaceborne remote sensing instruments do not measure individual tree diameters, instead rely on measurements of tree heights, canopy cover, density among others as means for estimating forest biomass. Profiling lidars, optical instruments capable of illuminating forests with light pulses and detecting the reflected light are commonly used to determine forest characteristics such as tree heights. Height estimates from lidars with larger footprints (of tens of meters) have been shown to relate well to stand biomass in various forests and ecoregions [88, 86]. Structure metrics from full-waveform lidars, instruments that sample the entire reflected waveforms, have been shown to be even better predictors of forest biomass [76, 77, 29, 60]. Even though the accuracy of these estimates are generally much poorer than the biomass estimates obtained using individual tree diameters, the ability of lidars to cover large regions make them extremely useful instruments for generating large-scale biomass estimates. In this chapter, we explore the accuracy of various lidar biomass relationships over the Harvard and Howland Forests using lidar and ground validation data.

3.1 Lidar Data Over the Harvard Forest

The laser vegetation imaging sensor (LVIS) is an airborne scanning laser altimeter developed by the NASA Goddard Spaceflight Center (GSFC) [9]. The instrument is a nadir looking profiler that is capable of sampling full waveform returns. It can cover large swaths by scanning the laser up to 7° off-nadir with a footprint that varies

between 20 to 80m. For each location the laser echoes are sampled and processed to generate geolocated full waveform data or moments that include ground height, canopy top and quartiles of the lidar energy returns.

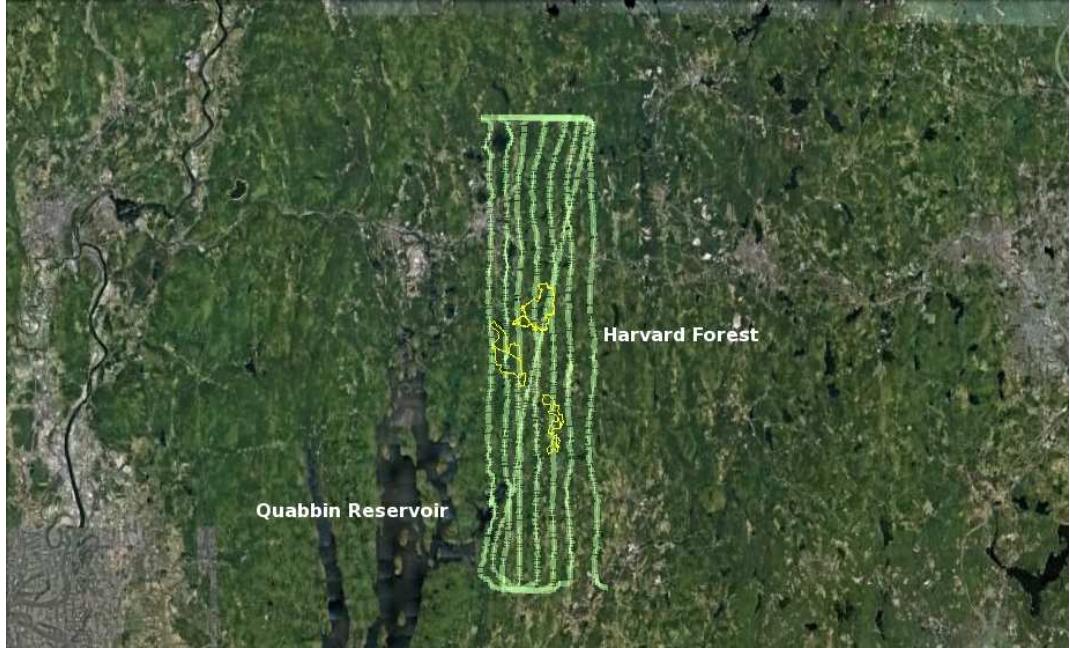


Figure 3.1: The LVIS nadir track is highlighted by green lines while boundaries of the Harvard forest are shown by the yellow lines. LVIS first imaged this area in 2003 and then in 2009.

Figure 3.1 shows the LVIS nadir track over the Harvard forest and surrounding region for a deployment in 2003. A similar profiling campaign took place in August, 2009 in conjunction with the field data collection. Full waveform data from both LVIS deployments have been processed by GSFC.

3.1.1 LVIS shot selection methodology

The LVIS data is distributed in two forms, as the full waveform data and estimated RH metrics (quartiles of the return waveform) for every laser shot. Each shot (identified by its corresponding shot number) is associated a latitude, longitude and height triplet with the first and last sample of the waveform. The choice of which LVIS shot should belong to a particular subplot or plot was based on percentage area

of overlap between the two, defined as the area of intersection as a percentage of the total area of the LVIS footprint. The center of the 25m diameter LVIS footprint was chosen as the latitude, longitude pair associated with the ground return. Most subplots have more than one shot from either the 2009 or the 2003 datasets with at least 5 percent overlap, with a mean overlap of 74 percent over all subplots.

3.1.2 LIVS RH metrics and field biomass

Most studies have reported strong relationships between lidar RH metrics and field biomass. It is expected that some ecological process controls the structure of forests as they grow and add to their biomass, in the hope that a measure of structure will somehow reflect biomass. Such a process is not perfectly understood and as a result the relationships between lidar metrics and biomass are inconsistent; varying over different biomes and study sites with factors such as species, age and climate (to name a few) affecting the outcome of such relationships, possibly differently.

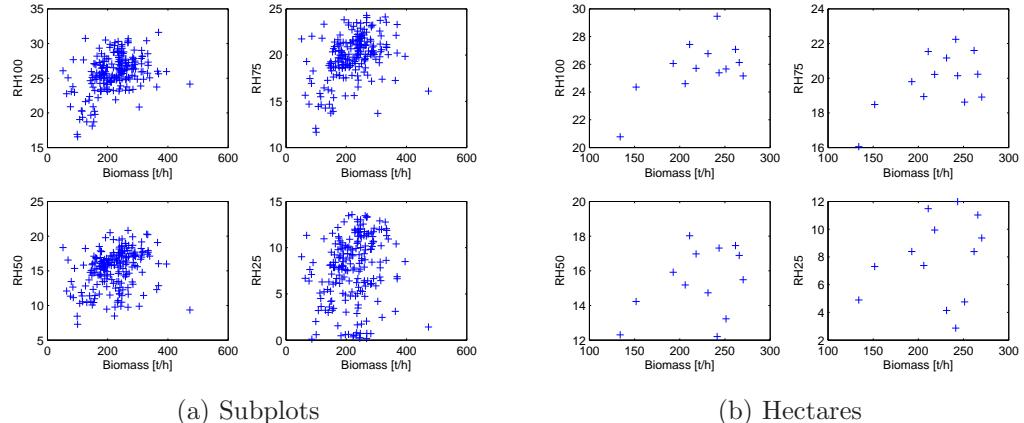


Figure 3.2: LVIS RH metrics plotted against the biomass values from the 208 subplots and the 13 one-hectare plots for an overlap of at least 5 percent or more.

The relationship between the four RH metrics and biomass at the Harvard Forest is shown in Figure 3.2 at both the one-hectare and subplot levels. For these scatter

plots only the shots with at least 5 percent overlap were selected.

Metric	R^2	RMSE [t/h]
RH100	0.170	61.231
RH75	0.170	61.233
RH50	0.086	64.250
RH25	0.031	66.173

Metric	R^2	RMSE [t/h]
RH100	0.337	34.462
RH75	0.343	35.847
RH50	0.137	41.086
RH25	0.052	43.054

(a) Subplots

(b) Hectares

Table 3.1: Regression statistics for LVIS RH metrics and field biomass for 5 Percent overlap.

Regression statistics for a linear fit between lidar RH metrics and field biomass are summarized for subplot and hectare scales in Table 3.1a 3.1b respectively. The RMSE (the root mean square error), in units of tons per hectare, is comparable to the standard deviation of the ground biomass, which is 66.85 t/h for subplot level estimates and 40.95 t/h for hectare level estimates.

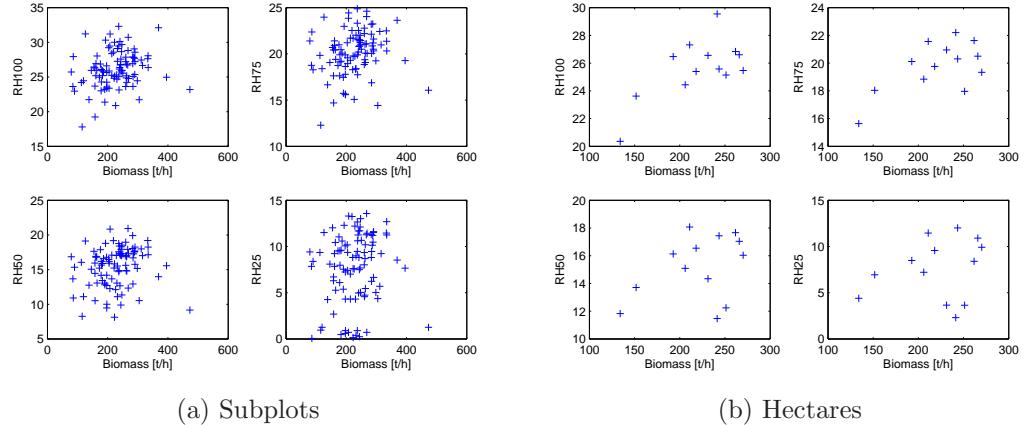


Figure 3.3: LVIS RH metrics plotted against the biomass at subplot and hectare scales for an overlap of 75 percent or more.

Figure 3.3 shows the relationship between hectare and subplot level biomass estimates and LVIS RH metrics for an overlap of at least 75 percent or more. The number of subplots that satisfy this criterion reduce to 97 instead of 208.

Metric	R ²	RMSE [t/h]	Metric	R ²	RMSE [t/h]
RH100	0.049	65.830	RH100	0.398	32.855
RH75	0.034	66.355	RH75	0.386	34.652
RH50	0.013	66.582	RH50	0.142	40.954
RH25	0.016	66.948	RH25	0.055	42.980

(a) Subplots

(b) Hectares

Table 3.2: Regression statistics for LVIS RH metrics and field biomass for at least 75 percent overlap.

Table 3.2 summarizes regression statistics for linear fits between biomass and LVIS shots with overlap of 75% or more. The R² statistic for subplot level estimates reduces significantly whereas that of the hectare level estimates increases slightly. There could be two reasons for the higher correlations seen at hectare scales, one that the LVIS shots and subplots are not perfectly geolocated causing low correlations at subplots, or that the relationship between structure and biomass only exists over larger spatial scales. It is hard to identify which may be the case, however, it is likely that there is a contribution from both. In either case, the relationship between any of the LVIS RH metrics and field biomass at the Harvard Forest is not very strong, in fact it is much weaker than reported in most studies relating biomass and lidar. This analysis focuses on understanding the reasons why this relationship is not similar to previous studies.

3.1.3 Discussion on LVIS RH metrics

LVIS RH metrics distributed by GSFC, however carefully processed aren't guaranteed to be error free. It is worth the effort to look at the metrics themselves more carefully.

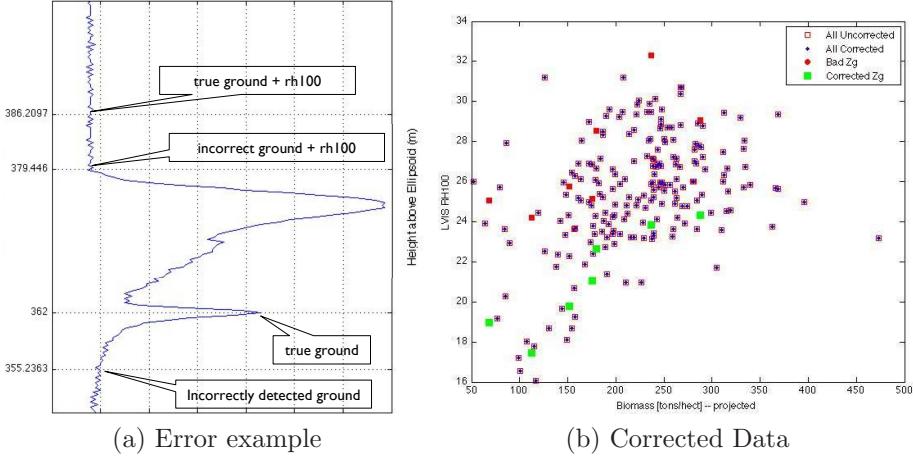


Figure 3.4: An example of errors in ground detection and LVIS RH100 corrected for badly detected ground values.

Most remote sensing data have contributions from a fair number of error sources. Of the ones that may be present in the LVIS data distributed by GSFC, the two that can be readily tested are ground detection and estimation of the energy quartiles. Figure 3.4a shows an example of an erroneously detected ground return in the LVIS data and how that affects estimation of the RH100 metric. In this particular case, for some reason the ground was detected about 5 meters below the true ground. The RH100 metric is essentially the height above the detected ground at which 100 percent of the waveform energy has been returned. An offset of 5m in the detection of ground directly translates to an overestimate of RH100 by the same amount. In Figure 3.4b LVIS RH100 metric from shots with maximal overlap are plotted against the corresponding subplot biomass. From visual inspection it was determined that seven of the 208 waveforms had errors in ground detection and therefore the data had

overestimated RH100, in some cases by up to 8m. Once the overestimated RH100 values were corrected fit statistics got better (the results of the corrected RH100 data points are shown by green squares in Figure 3.4b), but only marginally. It is hard to identify and completely eliminate all ground detection errors, therefore for the purposes of this analysis these errors are assumed to be present in the data with only negligible impact.

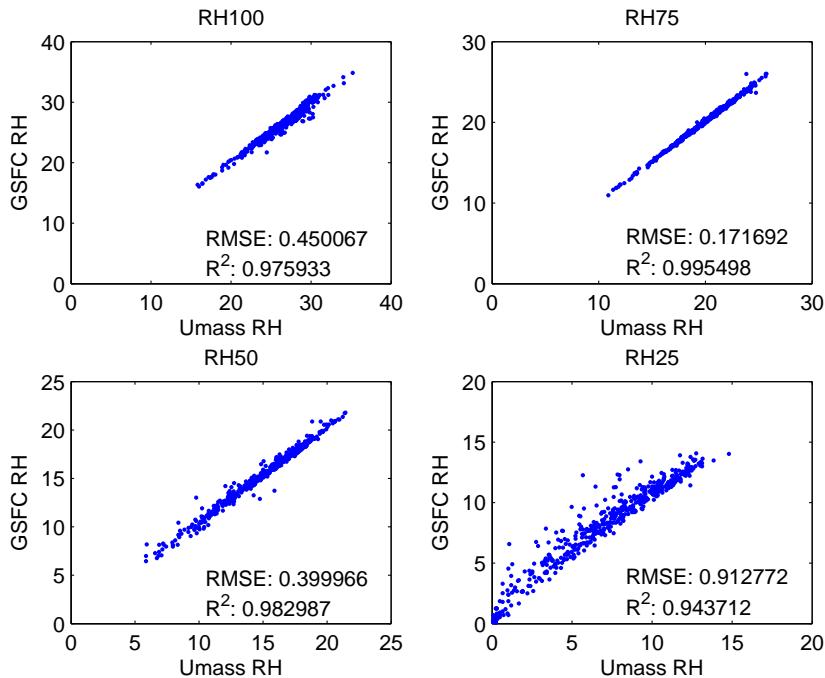


Figure 3.5: Comparison of LVIS RH metrics distributed by GSFC and RH metrics calculated at UMass.

To test whether there were any large inconsistencies in the estimation of RH metrics themselves, especially for shots over the field sites, the RH metrics were recalculated. To do so, the raw waveform was filtered and integrated up to each height. The height at which with integrated waveform energy reached 100 percent of the total energy in the waveform was said to be RH100, similarly the height of 75 percent energy return was estimated at RH75 and so on. The results of the recalculated RH metrics are plotted against the RH data distributed by GSFC in

Figure 3.5. There is relatively good agreement with the independently estimated metrics over the field sites, suggesting the absence of large inconsistencies. Testing these two error sources allows us to eliminate their effects as the driving factors behind the low correlation between biomass and the RH metrics seen over Harvard.

3.1.4 Performance of other lidar metrics

Various studies have used metrics other than the energy quartiles and found better correlation of those metrics with field biomass. It would be remiss not to revisit relevant metrics and see whether any perform better over the Harvard Forest field sites than the energy quartiles did. There is a large body of studies that analyze lidar relationships with biomass, however for the purpose of this analysis, studies that focused on large footprint full waveform lidar were reviewed. The following inexhaustive list summarizes lidar metrics presented in the reviewed literature

1. MEANRH100: Mean of RH100 over all relevant shots [86, 30, 60]
2. MAXRH100: Maximum of RH100 over all relevant shots [75]
3. MINRH100: Minimum of RH100 over all relevant shots [75]
4. MODRH100: Mode of RH100 over all relevant shots [77, 78]
5. STDRH100: Standard deviation of RH100 over all relevant shots [77, 78]
6. RH90: Mean of RH90 (90th energy percentile) over all relevant shots [29]
7. RH75: Mean of RH75 over all relevant shots [29]
8. RH50: Mean of RH50 (also called HOME) over all relevant shots [76, 29, 30, 31, 60]
9. RH25: Mean of RH25 over all relevant shots [29]
10. RH10: Mean of RH10 (10th energy percentile) over all relevant shots [29]
11. RH05: Mean of RH05 (5th energy percentile) over all relevant shots [29]
12. HTRT: Ration of RH100 and RH50 [30]
13. HG: Ground elevation [86]

14. MCH: Mean canopy height, calculated from the canopy height profile (CHP) as the first moment of the CHP averaged over all relevant shots [86, 76, 75, 77, 78]
15. STDMCH: Standard deviation of the mean canopy height over all relevant shots [77, 78]
16. QMCH: Quadratic mean canopy height, calculated as the second moment of the CHP averaged over all relevant shots [86, 76, 75, 77, 78]
17. STDQMCH: Standard deviation of the QMCH over all relevant shots [77, 78]
18. RVT: Energy of the canopy portion of a waveform [86]
19. COVER: Ratio of RVT to the total waveform energy corrected for ground albedo [75]
20. MCHCOV: Product of MCH and COVER [77, 78]
21. QMCHCOV: Product of QMCH and COVER [77, 78]
22. MCT: Mean canopy transmittance, calculated from the canopy transmittance profile and its first moment averaged over all relevant shots [77, 78]
23. STDMCT: Standard deviation of the MCT over all relevant shots [77, 78]
24. TH50: 50th percentile of the canopy transmittance profile averaged over all shots [77, 78]

The RH metrics in this list are all derived form the full lidar waveform as described earlier, however, some metrics require modification of the lidar waveform as described in [76] to produce the canopy height profiles.

The steps involved in generating the canopy height profile, shown in Figure 3.6, include filtering of the raw waveform, delineation of the ground and canopy portions, calculation of the transmittance height profile, correcting for attenuation effects and differencing the attenuation corrected transmittance height profile to obtain the canopy height profile. First, the raw waveform (or the waveform data distributed by GSFC) is filtered to remove noise. Estimates of the LVIS noise floor, on a shot by shot

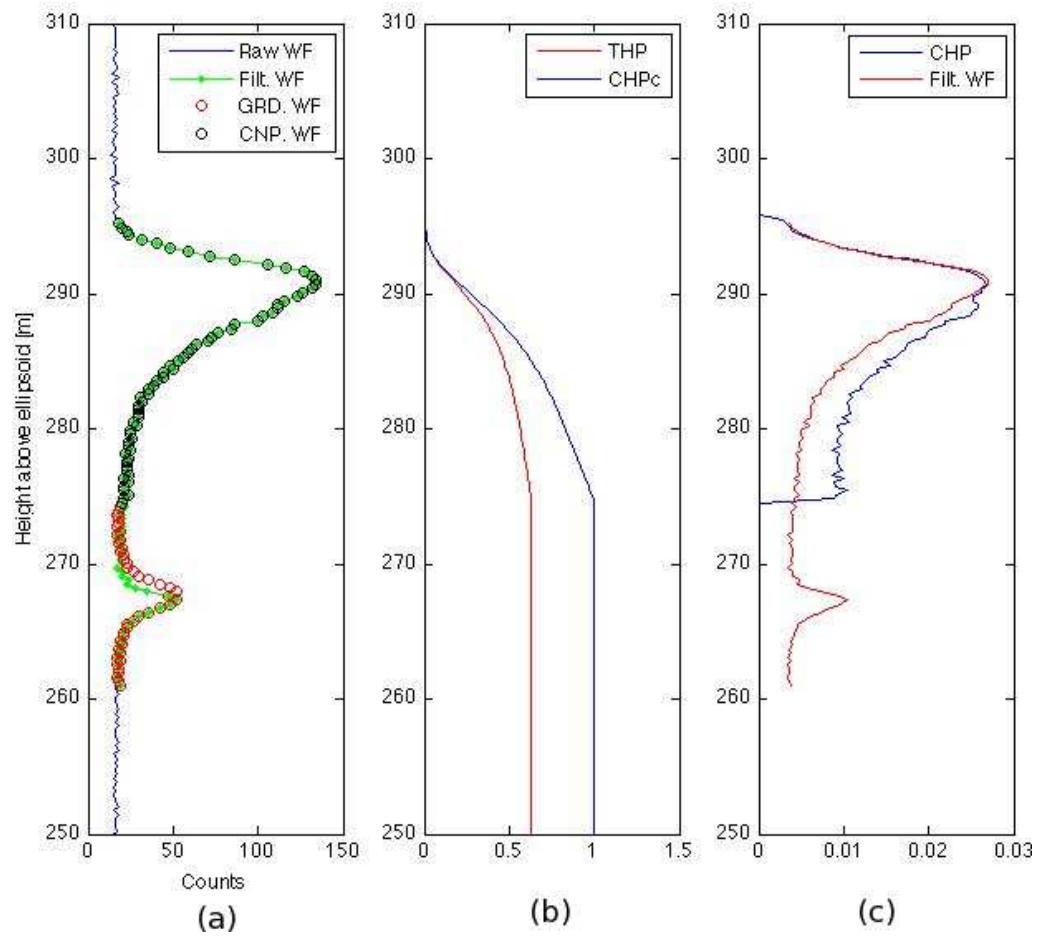


Figure 3.6: Illustration of the steps involved in calculating the canopy height profile.

basis, are distributed with the waveform data but can also be verified by averaging the waveform over a region that does not contain canopy returns. An example of a filtered waveform is shown in Figure 3.6a by the green line. To separate ground and canopy returns, the portion of the waveform below the detected ground is mirrored above it and called the ground return, $R_g(h)$. This is shown in Figure 3.6a with red circles. The canopy part of the waveform, $R_v(h)$ is estimated by subtracting $R_g(h)$ from the filtered waveform. The transmittance height profile (THP) is estimated from $R_v(h)$ and $R_g(h)$ by taking a ratio of the energy from canopy returns to the total energy, given by the relationship

$$THP(h) = \frac{\sum_{z=h_{max}}^h R_v(z)}{\sum R_v(z) + 2 \sum R_g(z)} \quad (3.1)$$

where the factor of 2 multiplying the ground energy contribution is to account for the lower ground albedo. Figure 3.6b shows the THP calculated using the ground and canopy parts of the waveform shown in Figure 3.6a. The cumulative canopy height profile, an intermediate product, is calculated from the THP using the MacArthur Horn equation [83]

$$CHPc(h) = -\ln(1 - THP(h)) \quad (3.2)$$

where \ln denotes the natural logarithm. The $CHPc$, shown in Figure 3.6b by the blue line, is differenced to obtain the canopy height profile, CHP . Figure 3.6c shows both the filtered raw waveform and the CHP . The mean canopy height, MCH , is estimated from the canopy height profile as

$$MCH = \sum_{h=0}^{h_{max}} CHP(h) \times h. \quad (3.3)$$

$QMCH$, the quadratic mean canopy height, is calculated from the CHP as

$$QMCH = \sqrt{\sum_{h=0}^{h_{max}} CHP(h) \times h^2}. \quad (3.4)$$

The variable COVER is computed as the ratio of total energy in R_v to the total waveform energy, i.e. the sum of R_v and R_g , with R_g corrected for albedo. The variable, MCT is calculated from the THP using

$$MCT = \sum_{h=0}^{h_{max}} THP(h) \times h. \quad (3.5)$$

Regression statistics for individual linear fits of all these lidar variables to field biomass at the subplot and hectare level biomass estimates are summarized in Tables 3.3a and 3.3b respectively for an LVIS overlap of at least 5 percent. Almost all of these lidar variables work better at hectare levels, however the best performing variables are still the RH metrics, with variables associated with RH100 performing best. Table 3.4 summarizes the fit statistics for shots with at least 75 percent overlap. Variables related to the RH metrics still perform best. Since the number of shots per subplot reduce significantly because of the stringent 75% overlap criterion (most subplots have at most one shot) the metrics that rely on standard deviations are inaccurately estimated which could be a reason for the lack of correlation. Even though, the number of subplots that satisfy the 75% overlap criterion have reduced to 97 from 240 (in the case of 5% overlap), that is still a significant number of samples for the correlations statistics to be of use. At hectare scales, the number of shots are not significantly affected by the 75% overlap criterion. The statistics for hectare level fits are summarized in Table 3.4b. The best R^2 values are obtained for RH100 related metrics for hectare level averages and the 75% overlap. However, at a maximum of 0.39, these R^2 values are very low compared to those reported in literature, some as high as 0.96. The RMSE numbers by themselves seem to suggest that the hectare scale correlations perform much better as the RMSE is reduced by one half compared

Lidar Metric	R ²	RMSE	Lidar Metric	R ²	RMSE
1. MEANRH100	0.170	61.231	1. MEANRH100	0.337	34.462
2. MAXRH100	0.178	60.951	2. MODRH100	0.163	38.726
3. MINRH100	0.178	60.931	3. MAXRH100	0.084	44.083
4. MODRH100	0.026	66.327	4. MINRH100	0.262	37.989
5. STDRH100	0.098	63.837	5. STDRH100	0.429	33.407
6. RH90	0.162	61.531	6. RH90	0.350	35.644
7. RH75	0.170	61.233	7. RH75	0.343	35.847
8. RH50	0.086	64.250	8. RH50	0.137	41.086
9. RH25	0.031	66.173	9. RH25	0.052	43.054
10. RH10	0.035	66.034	10. RH10	0.152	40.711
11. RH05	0.039	65.894	11. RH05	0.194	39.690
12. HTRT	0.001	67.194	12. HTRT	0.002	44.175
13. HG	0.000	67.199	13. HG	0.006	44.088
14. MCH	0.068	64.870	14. MCH	0.138	41.054
15. STDMCH	0.000	67.212	15. STDMCH	0.018	43.817
16. QMCH	0.086	64.252	16. QMCH	0.163	40.448
17. STDQMCH	0.002	67.159	17. STDQMCH	0.128	41.296
18. RVT	0.030	66.201	18. RVT	0.071	42.621
19. OVER	0.026	66.333	19. COVER	0.072	42.590
20. MCHCOV	0.069	64.863	20. MCHCOV	0.153	40.692
21. QMCHCOV	0.065	64.981	21. QMCHCOV	0.142	40.957
22. MCT	0.003	67.115	22. MCT	0.001	44.198
23. STDMCT	0.006	67.018	23. STDMCT	0.228	38.860
24. TH50	0.009	66.903	24. TH50	0.039	43.348

(a) Subplots

(b) Hectares

Table 3.3: Regression statistics for various Lidar Metrics and field biomass for an LVIS overlap of at least 5 percent. The RMSE values are in units of tons/hectare.

Lidar Metric	R ²	RMSE	Lidar Metric	R ²	RMSE
1. MEANRH100	0.049	65.830	1. MEANRH100	0.398	32.855
2. MAXRH100	0.041	66.095	2. MODRH100	0.350	34.136
3. MINRH100	0.041	66.095	3. MAXRH100	0.038	43.128
4. MODRH100	0.038	66.191	4. MINRH100	0.490	31.570
5. STDRH100	0.055	65.608	5. STDRH100	0.431	33.361
6. RH90	0.046	65.922	6. RH90	0.402	34.180
7. RH75	0.034	66.355	7. RH75	0.386	34.652
8. RH50	0.027	66.583	8. RH50	0.142	40.954
9. RH25	0.016	66.948	9. RH25	0.055	42.980
10. RH10	0.013	67.065	10. RH10	0.153	40.694
11. RH05	0.021	66.801	11. RH05	0.184	39.947
12. HTRT	0.000	67.496	12. HTRT	0.001	44.195
13. HG	0.007	67.261	13. HG	0.002	44.185
14. MCH	0.003	67.388	14. MCH	0.162	40.484
15. STDMCH	0.011	67.127	15. STDMCH	0.044	43.228
16. QMCH	0.006	67.317	16. QMCH	0.187	39.870
17. STDQMCH	0.007	67.262	17. STDQMCH	0.201	39.515
18. RVT	0.024	66.699	18. RVT	0.080	42.425
19. COVER	0.007	67.276	19. COVER	0.075	42.537
20. MCHCOV	0.021	66.805	20. MCHCOV	0.156	40.631
21. QMCHCOV	0.019	66.869	21. QMCHCOV	0.147	40.850
22. MCT	0.034	66.341	22. MCT	0.006	44.085
23. STDMCT	0.036	66.287	23. STDMCT	0.187	39.878
24. TH50	0.027	66.601	24. TH50	0.022	43.737

(a) Subplots

(b) Hectares

Table 3.4: Regression statistics for various Lidar Metrics and field biomass for an LVIS overlap of at least 75 percent. The RMSE values are in units of tons/hectare.

to the subplot scales. However the range of biomass values at hectare scales is also smaller which causes the RMSE values to appear lower.

To allow for a comparison of the performance of lidar metrics between the two spatial scales, one possible solution would be to look at RMSE values as a fraction of the standard deviation of field biomass. The standard deviation of biomass data is representative of the natural biomass variation of the forest while the RMSE is an estimate of the accuracy of the biomass estimates obtained using an independent source (in this case lidar). To be able to trust lidar derived biomass, the RMSE should ideally be a small fraction of the biomass standard deviation. The RMSE as a fraction of the standard deviation not only allows ready comparison between the two spatial scales, but an indication of how good the estimators perform as well. At subplot levels, the best RMSE of 60.9t/h (for MINRH100 at 5% overlap), is approximately 90% of the biomass standard deviation (66.5t/h), whereas the best performing RMSE at hectare scales (MINRH100 at 31.57t/h for 75% overlap), is about 76% of the standard deviation of field biomass at hectare scales (42.33t/h). The improvement in the normalized RMSE no longer appears to be as significant, from 90% to 76%, instead of the factor of two reduction for the non-normalized RMSE values.

Combination of lidar metrics

A few of the studies reviewed also look at combinations of the various lidar metrics as predictors of biomass. The method for choosing which lidar metrics to include in the linear combinations is not always specified, some seem to employ principal component analysis-type techniques. In any case, the basic idea is to choose variables that are somewhat uncorrelated with each other while still related to the biomass data. Such combinations therefore tend to be site specific. In the case of Harvard Forest, to see which of the lidar variables should be included in a linear combination, the regression statistics for individual metrics summarized in Tables 3.3 and 3.4 were

used along with a correlation matrix of the twenty four variables computed over the field data.

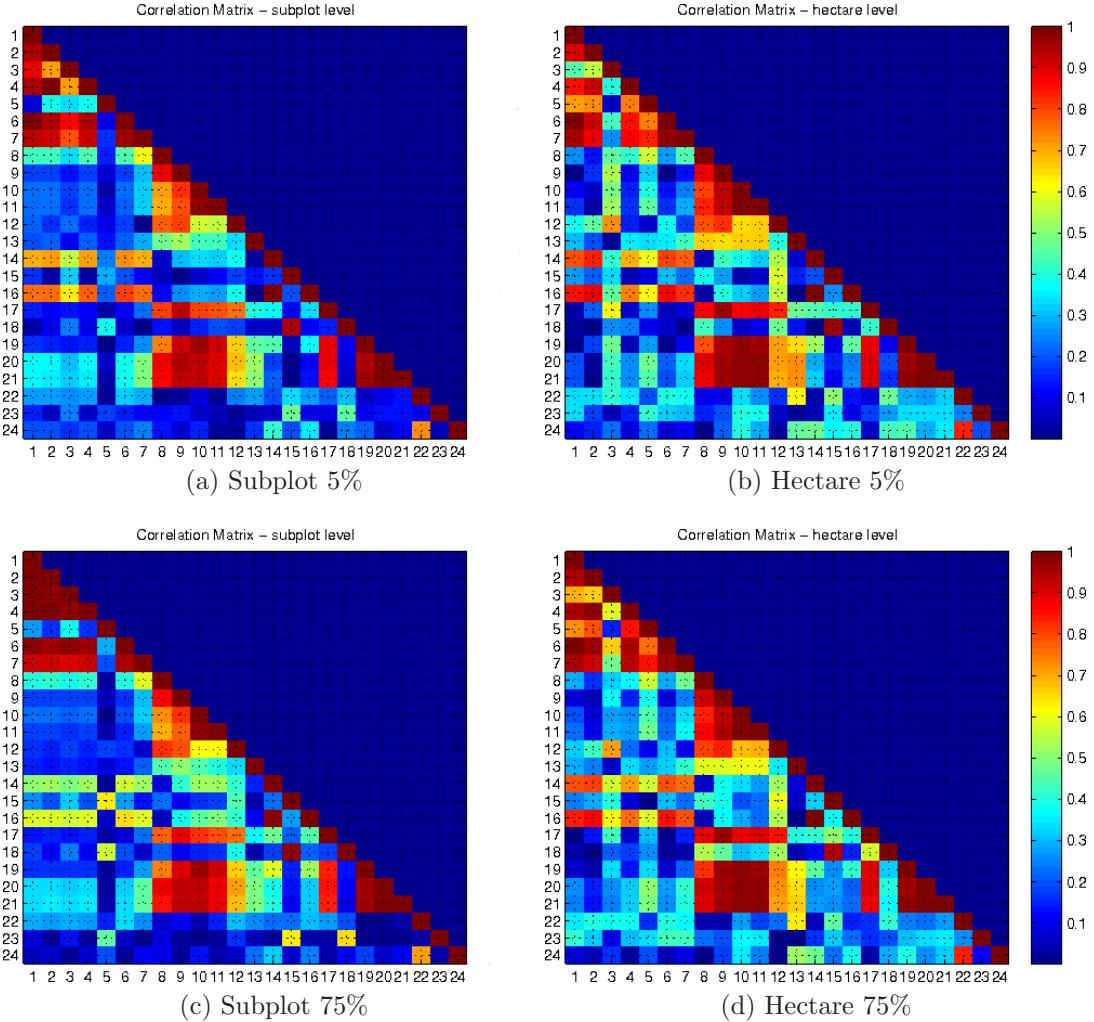


Figure 3.7: Correlation matrices for the 24 lidar variables for hectares and subplots with 5 and 75% shot overlap percentage.

Figure 3.7 shows the four correlation matrices for subplot and hectare scales and overlap percentages of 5 and 75%. Based on these matrices and the regression statistics of each individual variable to biomass data, the set of variables chosen for linear regression at subplot levels included MEANRH100, MAXRH100, MINRH100 and RH75. For hectare scales this set of variables consisted of MEANRH100, MINRH100, STDRH100, RH50, RH75 and STDMCT.

Overlap	Subplots		Hectares	
	R ²	RMSE	R ²	RMSE
5%	0.171	61.05	0.366	33.70
75%	0.028	66.20	0.384	33.21

Table 3.5: Regression statistics for linear combinations of lidar variables to biomass data. The RMSEs are in units of tons/hectare.

The regression statistics for linear combination of these variables are given in Table 3.5. It is important to point out that the R² values shown in Table 3.5 are calculated using the adjusted R² metric that accounts for the artificially high R² values if the number of variables are on the order of the number of observations by essentially scaling the regular R² metric by a ratio of the number of observations to the number of variables used in the regression. Comparing the numbers in Table 3.5 and Tables 3.3, 3.4 it appears that linear combinations of the LVIS metrics do not perform significantly better than the individual metrics themselves. One reason for this could be the choice of variables. Other linear combinations of the twenty four metrics for subplot level biomass relationships do no perform any better. Similarly, no other combination of the LVIS metrics has resulted in a significant increase in performance at hectare scales either.

3.1.5 Extending the range of field biomass values

There seems to be more of a relationship between biomass and lidar metrics averaged over hectares than over subplots. However, the correlation is not as strong as has been reported in literature (with R² values generally 0.8 and higher). One possible cause for the lack of correlation between lidar data and field biomass could be the narrow range of biomass values observed at the Harvard Forest at hectare scales, ranging from 134 to 270 t/h (a difference of 136 t/h). At subplot levels, however, this range is much larger, at about 420 t/h. So far, the subplots that belong to a

physical hectare had been aggregated to form hectare level estimates. This reduced the variability in field data but was likely more reflective of the true state of the forest.

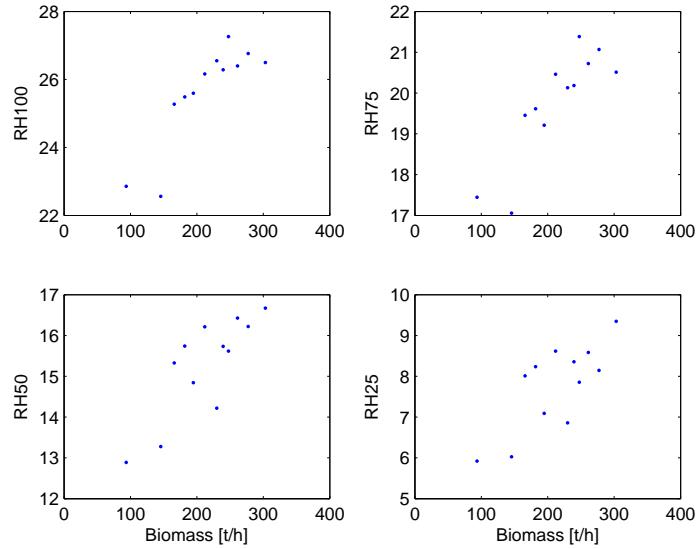


Figure 3.8: LVIS RH metrics plotted against biomass data over hectares created by aggregating subplots from a sorted list.

To test whether better correlations would exist if the range of biomass values was larger at hectare scales, subplots were sorted according to their biomass values, blocks of sixteen subplots were then aggregated from this sorted list to create biomass and lidar metrics on hectare levels. In essence this created hectares with a larger variation in biomass (from 90 to 300 t/h). The LVIS RH metrics form these sorted hectares are plotted against biomass for an overlap of 35% in Figure 3.8. The 35% overlap criterion was chosen to maximize shot overlap without excluding any subplots from the analysis.

Regression statistics for the individual lidar variables are summarized in Table 3.6 for the sorted hectares. Regression from a linear combination of a set of these variables, which consisted of MEANRH100, MAXRH100, RH50 and MCHCOV, results

Lidar Metric	R ²	RMSE
1. RH100	0.717	31.744
2. MODRH100	0.702	32.557
3. MAXRH100	0.607	37.366
4. MINRH100	0.706	32.347
5. STDRH100	0.305	49.712
6. RH90	0.690	33.199
7. RH75	0.720	31.533
8. RH50	0.634	36.087
9. RH25	0.545	40.234
10. RH10	0.737	30.591
11. RH05	0.720	31.543
12. HTRT	0.056	61.264
13. HG	0.021	60.255
14. MCH	0.499	42.199
15. STDMCH	0.069	61.631
16. QMCH	0.511	41.683
17. STDQMCH	0.066	57.626
18. RVT	0.532	40.774
19. COVER	0.571	39.061
20. MCHCOV	0.733	30.832
21. QMCHCOV	0.709	32.163
22. MCT	0.096	62.412
23. STDMCT	0.047	60.993
24. TH50	0.202	53.252

Table 3.6: Regression statistics for various Lidar Metrics and field biomass for an LVIS overlap of at least 35 percent over hectares formed by aggregating subplots from a sorted list. The RMSE values are in units of tons/hectare.

in an R^2 value of 0.82 and an RMSE of 25.25 t/h. This RMSE is roughly 42% of the standard deviation of the sorted biomass data (59.6 t/h). All of these results are a considerable improvement upon the case where the biomass data was not sorted, in fact the results from the linear combination are comparable to other studies. This would seem to suggest that the small variation in field data was the reason why there was such a weak correlation between lidar metrics and biomass over the Harvard Forest. However, if the combination that produces an R^2 value of 0.82, reflects a real relationship between lidar and field biomass and not some statistical artifact, it ought to be able to perform almost as well on the unsorted data. Ideally this relationship would be applied to an independent data set from a similarly composed forest to test its predictive capabilities. Given the scarcity of field data, however, the unsorted data could also serve as the independent data set, however imperfect.

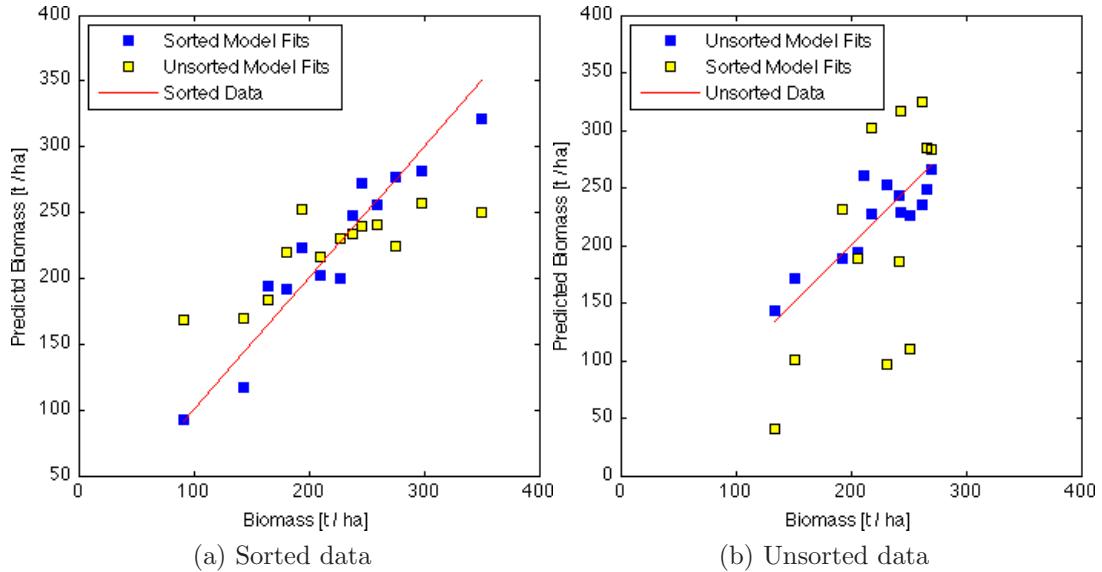


Figure 3.9: Performance of models developed by regressing lidar variables to biomass data averaged over hectares created from sorted and unsorted data.

Two models developed using a linear combination of a set of lidar variables that consisted of MEANRH100, MAXRH100, RH50 and MCHCOV over biomass data from hectares that were created from sorted and unsorted subplots. The model de-

rived from sorted data was applied to unsorted hectares and vice versa. Figure 3.9a shows the performance of the sorted and unsorted models at predicting sorted data, whereas Figure 3.9b plots the results of applying the two models to biomass and lidar data from unsorted hectares. The regression statistics, R^2 and RMSE for applying the unsorted model to sorted data are 0.64 and 40.89t/h (68% of the biomass standard deviation), whereas applying the sorted model to unsorted data results in an R^2 and RMSE of 0.19 and 37.92t/h (92% of the biomass standard deviation). This seems to suggest that the model developed from unsorted data is more reflective of a height to biomass relationship than the sorted data model. The sorted data model seems to only predict the sorted data, which would suggest that the high R^2 values are more of a statistical artifact of over-fitting the data than a reflection of a real relationship.

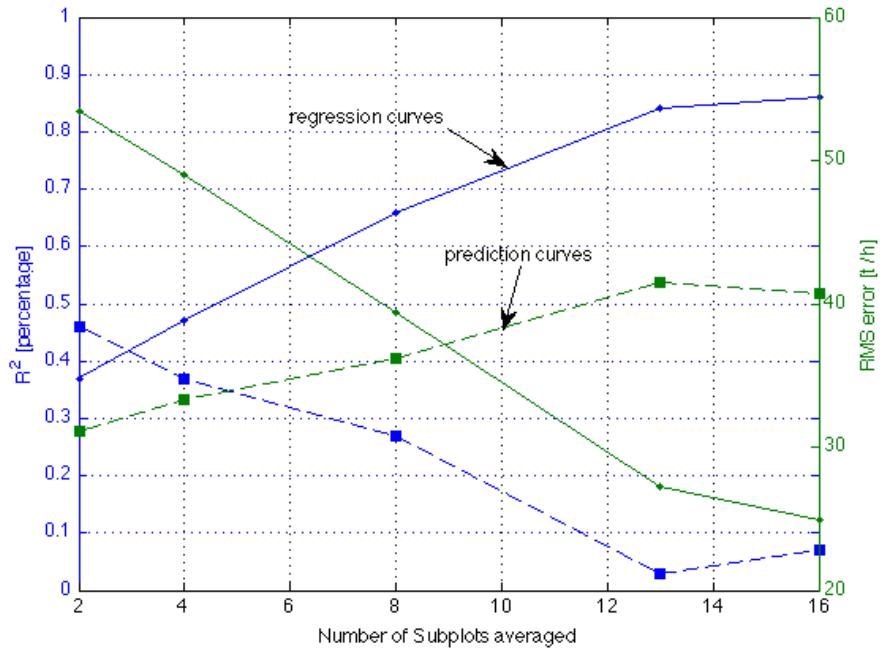


Figure 3.10: Sorted model is applied to both the sorted data (regression curve, shown by the solid lines) and the unsorted data (prediction curve, shown by the dashed lines) for varying number of subplots aggregated.

Figure 3.10 shows the R^2 and RMSE statistics for the sorted model as it is applied to the sorted data (labeled regression curve) and unsorted data (called prediction curve) for varying number of subplots aggregated. The regression statistics predictably get better with more averaging, however the exact opposite happens with the prediction curve which may be another indication of an over-fitted model.

The extension of the dynamic range of biomass data using the sorting method did lead to better fits, however those statistics are misleading. The prediction approach, demonstrative as it may be, does not however conclusively prove that field data with naturally larger dynamic ranges wouldn't lead to better lidar-biomass models, neither does it shed light on why the relationship between lidar variables and biomass is not as strong over the Harvard Forest as seen in other sites. It would be useful to conduct a similar analysis over a field site where lidar is shown to predict biomass well. The Howland Forest research site, located in central Maine, is expected to be ecologically similar to the Harvard Forest. Furthermore, a good correlation between lidar and biomass has been reported for this forest.

3.2 Lidar Data Over the Howland Forest

LVIS was flown over the Howland Forest in 2009 covering all the field sites. The RH metrics derived from full waveform data from that field campaign are available from the GSFC LVIS repository. Figure 3.11 shows the four energy quartiles (RH100, RH75, RH50 and RH25) over subplots (3.11a) and hectares (3.11b) for a minimum of 5% overlap between the LVIS shots and the subplot (overlap is defined the same way as before).

The performance of LVIS 2009 energy quartiles is summarized in Table 3.7 as R^2 and RMSE for linear fits between each metric and dbh derived biomass. From the scatter plots in Figure 3.11 and fit statistics in Table 3.7 it is apparent that LVIS RH

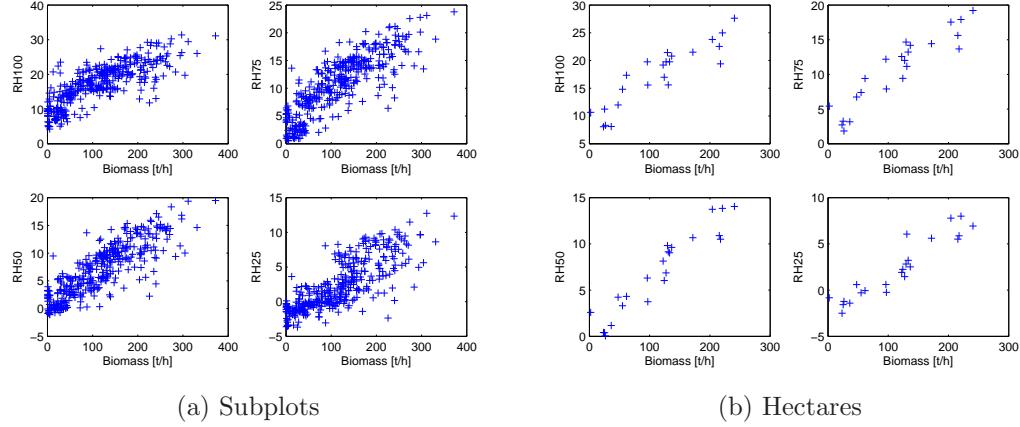


Figure 3.11: LVIS RH metrics plotted against the biomass values from the 364 subplots and the 23 hectare plots for an overlap of 5 percent or more.

metrics have a significantly stronger relationship to field biomass at Howland than at Harvard. A maximum R^2 of 0.89 and a minimum RMSE of 23.5t/h (as 32% of the standard deviation of field biomass), both for RH50 at hectare scales, are comparable to results reported in other studies.

Metric	R^2	RMSE [t/h]
RH100	0.632	47.885
RH75	0.700	43.208
RH50	0.718	41.904
RH25	0.646	46.938

Metric	R^2	RMSE [t/h]
RH100	0.811	31.508
RH75	0.854	27.625
RH50	0.895	23.511
RH25	0.870	26.148

(a) Subplots

(b) Hectares

Table 3.7: Regression statistics for LVIS RH metrics and field biomass for 5 percent overlap.

Figure 3.12 plots the four RH metrics against biomass data for an overlap of at least 75%. A total of 207 of 364 subplots satisfy this overlap criterion. Table 3.8 summarizes regression statistics for linear fits between the energy quartiles from the

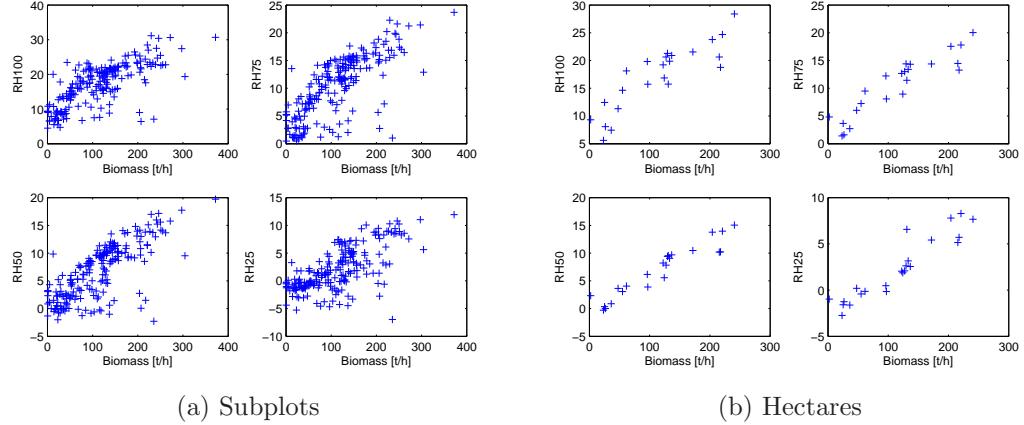


Figure 3.12: LVIS RH metrics plotted against the biomass values from 193 of 364 subplot and the 23 hectare plots for an overlap of 75 percent or more.

shots with 75% overlap and field data. Performance of the linear fits between LVIS and biomass at the Howland Forest does not get significantly worse with higher overlap thresholds as was the case at the Harvard Forest. In fact the worst performing RH metric over Howland performs significantly better than any single metric over the Harvard Forest at either the subplot or hectare scales regardless of overlap percentage.

Metric	R ²	RMSE [t/h]
RH100	0.518	49.961
RH75	0.563	47.555
RH50	0.568	47.315
RH25	0.467	52.524

Metric	R ²	RMSE [t/h]
RH100	0.752	36.025
RH75	0.833	29.555
RH50	0.889	24.119
RH25	0.866	26.516

(a) Subplots

(b) Hectares

Table 3.8: Regression statistics for LVIS RH metrics and field biomass for 75 percent overlap.

More LVIS data was collected over Howland Forest in a similar campaign in 2003. The relationship between RH metrics from that dataset and field biomass are similar

to results from the 2009 dataset. Even though the R^2 values are slightly lower and the root mean squared errors are slightly larger, they are still much better than the results over the Harvard Forest.

With such similar data collection and analysis methodologies over the Harvard and Howland forests, and the similarity of results over Howland to those reported in literature, it is tempting to conclude that the Howland results are more representative of height to biomass relationships. That the problem at Harvard may have been with the field data, site selection or the nature of the Harvard Forest itself is not an entirely unreasonable conclusion. However, it may be instructive to take a closer look at the differences between the two analyses. Assuming that there were insignificant differences in measurement of tree diameters, offsetting errors in determining GPS coordinates of plot locations, the major differences between the two sites are

- An increased presence of low biomass sites at the Howland Forest.
- Species composition of the two forests.

In the following few sections these differences will be explored in more detail with the intent of making the Harvard and Howland Forest results more comparable.

3.2.1 Low biomass sites

The field data at Howland forest includes subplots and hectares with low biomass values and tree counts. Such plots are absent from the Harvard Forest dataset. To assess the impact of these low biomass sites on the performance of RH metrics, we could either look for low biomass sites around Harvard forest, or exclude the low biomass sites from the Howland dataset. For the purposes of this analysis, the latter is chosen. The choice of low biomass sites was established using a threshold of 20 trees or less per subplot (or 360 trees or less per hectare) and a biomass of less than 80t/h (both lower than the lower numbers at Harvard).

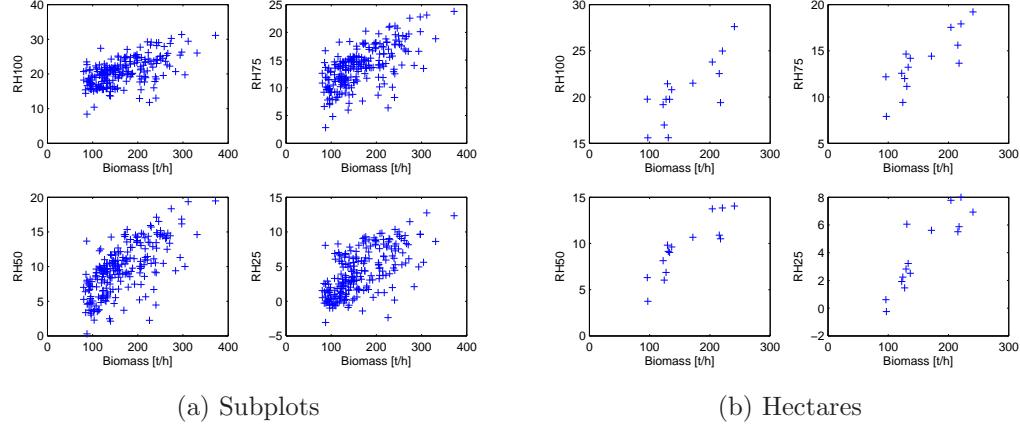


Figure 3.13: LVIS RH metrics plotted against the biomass estimates from subplots and hectares excluding low-biomass datapoints for an overlap of 5 percent or more.

Metric	R ²	RMSE [t/h]
RH100	0.333	47.120
RH75	0.413	44.209
RH50	0.449	42.838
RH25	0.410	44.299

(a) Subplots 5%

Metric	R ²	RMSE [t/h]
RH100	0.560	32.425
RH75	0.656	28.684
RH50	0.753	24.323
RH25	0.752	24.366

(b) Hectares 5%

Metric	R ²	RMSE [t/h]
RH100	0.216	47.911
RH75	0.259	46.555
RH50	0.307	45.040
RH25	0.286	45.715

(c) Subplots 75%

Metric	R ²	RMSE [t/h]
RH100	0.447	36.372
RH75	0.583	31.589
RH50	0.720	25.894
RH25	0.732	25.321

(d) Hectares 75%

Table 3.9: Regression statistics for LVIS RH metrics and field biomass for 5 and 75 percent overlaps excluding low biomass sites.

Figure 3.13 shows the scatter plots and Table 3.9 summarizes the regression statistics for linear fits between the four LVIS RH metrics and biomass data over Howland forest with low biomass sites excluded from the analysis. The number of subplots with high biomass and 5% LVIS overlap reduce to 225, while the number for 75% overlap reduce to 132 subplots. The number of one-hectare plots reduce from 23 to 15 for both cases. There are still enough subplots or plots for the correlation statistics to be meaningful. For 5% overlaps the R^2 values of linear fits between RH50 and biomass reduce from 0.72 seen in Table 3.7a to 0.45, while the RMSE values remain essentially the same, around 42t/h. A similar trend is seen in subplot level regressions at 75% overlap. The drop in R^2 values for hectare scale estimates is also significant, from 0.89 to 0.72, without having a significant impact on the RMSE values. The lack of impact on RMSE can be explained by the fact that the eliminated data points were low biomass and also low height, so any RMS error would be small anyway. The impact on R^2 , however, is harder to explain. Since R^2 is a commonly used metric for judging the strength of lidar-biomass relationships, the large impact (with 50% reduction in the case of subplots) is all the more important. A possible explanation is that these low biomass sites were masking the weak relationship between height and biomass by artificially increasing the range of biomass values (as was done in the case of sorted hectares at Harvard Forest). Since we are interested in determining the biomass in forests where low biomass areas are hard to find anyway, using accuracy measures of lidar-radar relationships over these artificially extended ranges might be misleading.

Comparing the statistics for high-biomass only field data and LVIS RH metrics at the Howland and Harvard Forests, it is apparent that LVIS still performs better at the Howland Forest with the best R^2 value at the Harvard Forest of 0.40 and 0.72 at the Howland Forest, and RMSE of 76% and 61% of the biomass standard deviation

respectively. Although the difference in statistics is smaller than before, the difference is statistically significant.

3.2.2 Species composition

On a species by species basis, the difference in biomass values arrived at by using different allometric equations is easily explained. However when biomass is estimated over hectares, where species are often mixed, species composition would also impact biomass estimates obtained by using overlapping sets of allometric equations. The extent of this impact would be different from one field site to another. Howland Forest for example is 72% coniferous by biomass whereas at the Harvard Forest coniferous trees only contribute to 33% of the total biomass. The impact of varying species composition on Lidar-biomass relationships, can be minimized by simplifying the allometries to two equations only, one for deciduous and the other for coniferous trees. This would cause marked deviations from some true biomass, however since the objective is to compare two forests the bias can be set aside for now.

Two equations were chosen to represent coniferous and deciduous trees, and applied to both the Howland and Harvard Forest datasets. Table 3.10 summarizes regression statistics for linear fits between four LVIS RH metrics and biomass at the Harvard and Howland Forests for subplots and hectares for an overlap of 5 percent or more, while Table 3.11 lists the same statistics for LVIS shots with overlaps of 75% or more. Although the metric that displays best correlations with this ‘coniferous-deciduous’ biomass is different at Howland and Harvard (RH100 at Harvard and RH50 at Howland), with an R^2 value of 0.68 and 0.61 for hectares the strength of these correlations is very similar, furthermore, the best RMSEs at both locations, as a percentage of biomass is approximately the same at 58%. Studies such as [91] have analyzed data over the Howland and Harvard Forests and have concluded that structure metrics from coniferous trees have a better correlation with biomass than

Metric	R ²	RMSE [t/h]
RH100	0.289	80.438
RH75	0.247	82.807
RH50	0.043	93.351
RH25	0.006	95.107

(a) Harvard Subplots

Metric	R ²	RMSE [t/h]
RH100	0.513	46.316
RH75	0.412	50.920
RH50	0.035	67.578
RH25	0.083	69.137

(b) Harvard Hectares

Metric	R ²	RMSE [t/h]
RH100	0.381	56.149
RH75	0.446	53.126
RH50	0.471	51.895
RH25	0.429	53.935

(c) Howland Subplots

Metric	R ²	RMSE [t/h]
RH100	0.587	36.805
RH75	0.648	33.989
RH50	0.709	30.898
RH25	0.686	32.095

(d) Howland Hectares

Table 3.10: Regression statistics for LVIS RH metrics and field biomass over Harvard and Howland Forests estimated using the ‘coniferous-deciduous’ allometry for 5 percent overlap.

Metric	R ²	RMSE [t/h]
RH100	0.114	89.453
RH75	0.073	91.469
RH50	0.003	94.889
RH25	0.001	95.108

(a) Harvard Subplots

Metric	R ²	RMSE [t/h]
RH100	0.618	41.030
RH75	0.494	47.205
RH50	0.021	67.121
RH25	0.078	68.982

(b) Harvard Hectares

Metric	R ²	RMSE [t/h]
RH100	0.248	57.610
RH75	0.286	56.150
RH50	0.323	54.693
RH25	0.287	56.114

(c) Howland Subplots

Metric	R ²	RMSE [t/h]
RH100	0.474	41.543
RH75	0.579	37.191
RH50	0.684	32.210
RH25	0.675	32.650

(d) Howland Hectares

Table 3.11: Regression statistics for LVIS RH metrics and field biomass over Harvard and Howland Forests estimated using coniferous and deciduous equations for 75 percent overlap.

deciduous trees. Those conclusions, we find may be driven by the high correlations seen over the Howland Forest that is mostly coniferous and the low correlations found over the Harvard Forest which is mostly deciduous.

3.3 Combining Howland and Harvard Forest Datasets

With the ‘coniferous-deciduous’ allometry leading to comparable correlations between lidar and biomass at Howland and Harvard it is possible to see how lidar performs when data from the two sites is combined. Figure 3.14 shows the subplot and hectare level LVIS RH metrics against biomass data derived using the ‘coniferous-deciduous’ allometry at both the Harvard and Howland Forest, while Table 3.12 summarizes the fit statistics for the combined data.

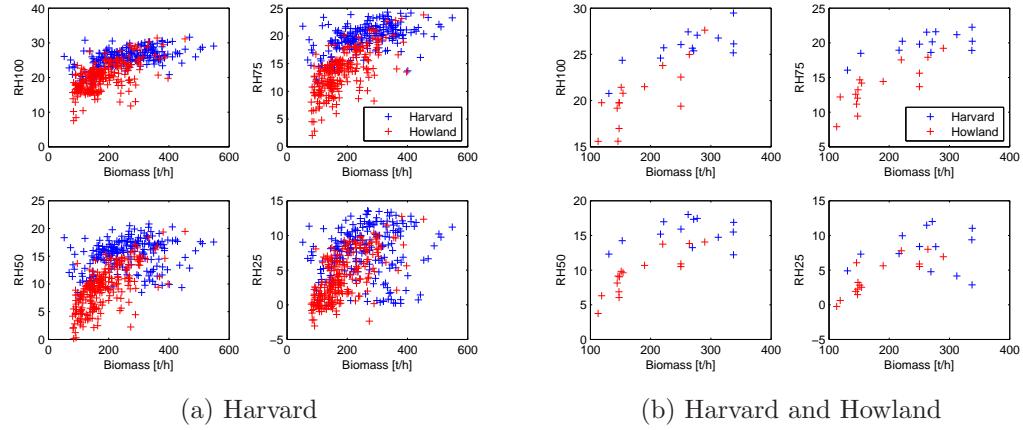


Figure 3.14: LVIS RH metrics plotted against data from both Harvard and Howland Forests using ‘coniferous-deciduous’ allometry and 5% overlap.

Figure 3.15 and Table 3.13 summarize the performance of the four LVIS RH quartiles and biomass data from both forests using a 75% overlap for LVIS shots. The performance of the combined data is more similar to results at the Howland Forest, with all RH metrics somewhat correlated with biomass, except that the best

performing metric is RH100 which explained the biomass best at Harvard. The best R^2 statistic (for RH100 at 75% overlap) is 0.66 and an RMSE of 39t/h (50% of the biomass variation). These seem to be consistent with fit statistics at the individual sites.

Metric	R^2	RMSE [t/h]
RH100	0.423	69.553
RH75	0.427	69.344
RH50	0.341	74.372
RH25	0.233	80.230

Metric	R^2	RMSE [t/h]
RH100	0.661	41.989
RH75	0.632	43.766
RH50	0.518	50.087
RH25	0.358	57.829

(a) Subplots

(b) Hectares

Table 3.12: Regression statistics for LVIS RH metrics over both the Howland and Harvard Forest with biomass values estimated using ‘coniferous-deciduous’ allometry and 5 percent overlap.

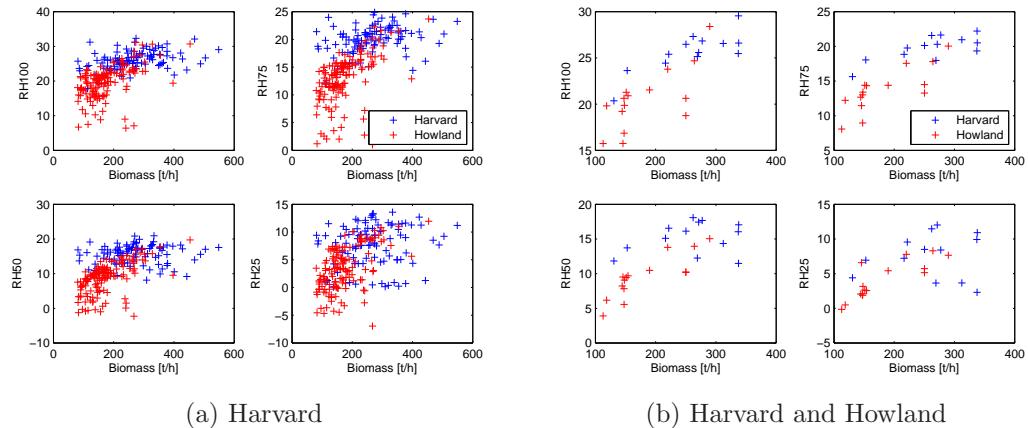


Figure 3.15: LVIS RH metrics plotted against data from both Harvard and Howland Forests using ‘coniferous-deciduous’ allometry and 75% overlap.

3.3.1 Effects of various allometric equations

So far, the lidar metrics over the Harvard and Howland sites have been compared to biomass estimates obtained from the Single-Site allometric equations from

Metric	R ²	RMSE [t/h]	Metric	R ²	RMSE [t/h]
RH100	0.346	74.335	RH100	0.649	42.744
RH75	0.363	73.356	RH75	0.636	43.511
RH50	0.322	75.668	RH50	0.512	50.394
RH25	0.232	80.532	RH25	0.336	58.815

(a) Subplots

(b) Hectares

Table 3.13: Regression statistics for LVIS RH metrics over both the Howland and Harvard Forest with biomass values estimated using ‘coniferous-deciduous’ allometry and 75 percent overlap.

Ter-Mikaelian [118]. As noted in Chapter 2 there are differences between biomass estimates obtained from different allometric equations. The impact of using different allometric equations can be readily analyzed. Table 3.14 lists the R² and RMSE statistics for fits between the four LVIS RH metrics, from shots with at least 50% overlap and biomass estimates obtained using the three allometric equations discussed in Chapter 2, namely the Single-Site (Ter-Mikaelian [118]), the Ensemble (Jenkins [65]) and the BLUE (Lambert [71]) allometries. As described before, allometric equations are based on diameter-biomass relationships developed at a particular site and the Single-site equations are a compilation of species-specific equations from studies conducted at sites most similar to the Harvard and the Howland Forests. The Ensemble equations provide coefficients for estimating biomass calculated from diameter-biomass data averaged over the continental United States, thus including variations due to differing biomes. While the BLUE allometry also captures cross-biome variations using data from sites spread across Canada, it estimates coefficients for the diameter-biomass relationships using more sound statistical methods and better data than the Ensemble equations of Jenkins et. al. [65]. The resultant biomass estimates over the field sites using the three different allometries are summarized in Chapter 2, however Table 3.14 summarizes the impact of using different allometric equations on the lidar-biomass

relationship itself. The difference in the R^2 values for different allometries is small for any of the RH metrics, however the RMSE is higher for the Ensemble allometry by a small margin, which can be attributed to the higher mean biomass estimates obtained by this allometry compared to the others. Overall, using different allometric equations, does not significantly alter the relationship between lidar metrics and field biomass.

Metric	R^2	RMSE [t/h]	Metric	R^2	RMSE [t/h]
Single-Site (Ter-Mikaelian [118])					
RH100	0.397	64.558	RH100	0.722	34.540
RH75	0.442	62.096	RH75	0.760	32.126
RH50	0.395	64.639	RH50	0.697	36.058
RH25	0.266	71.214	RH25	0.479	47.299
Ensemble (Jenkins [65])					
RH100	0.378	65.088	RH100	0.638	38.629
RH75	0.405	63.629	RH75	0.638	38.596
RH50	0.356	66.212	RH50	0.562	42.490
RH25	0.261	70.922	RH25	0.410	49.304
BLUE (Lambert [71])					
RH100	0.384	52.527	RH100	0.671	32.717
RH75	0.410	51.414	RH75	0.695	31.518
RH50	0.270	57.183	RH50	0.519	39.571
RH25	0.135	62.252	RH25	0.295	47.927

(a) Subplots

(b) Hectares

Table 3.14: Regression statistics for LVIS RH metrics using 50% overlap over both the Howland and Harvard Forest with biomass values estimated using the three allometric equations; Single-Site (Ter-Mikaelian [118], Ensemble (Jenkins [65]), BLUE (Lambert [71])).

3.3.2 Saturation in RH metrics

For the combined RH metrics there seems to be a saturation trend at subplot scales which is absent from data aggregated to hectares. The lack of trend could either be real or hidden at hectares because of the fewer data samples. Since the Harvard data is in general higher in biomass, saturation could explain why all RH metrics didn't do so well at Harvard. Fit coefficients suggest that there is indeed a smaller slope to the Harvard data for RH50 and RH25. Subplot data is fairly noisy which makes it hard see the saturation trend clearly. However, if the combined subplots were sorted by biomass and then averaged to create the sorted hectares as done previously, the trend may become clearer. Figure 3.16 plots the LVIS RH metrics averaged over sixteen

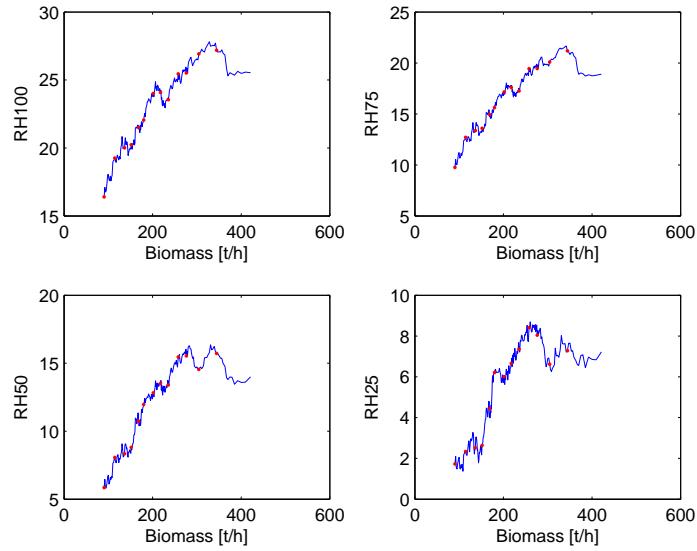


Figure 3.16: LVIS RH metrics averaged over 16-subplots chosen from a list of subplots sorted by biomass.

subplots chosen from the sorted list. The blue line in Figure 3.16 is for data samples from an overlapping window of sixteen samples used to select subplots from the sorted list, whereas the red-dots are the samples from the non-overlapping sixteen-sample window. The saturation trends in the RH50 and RH25 metrics are quite apparent,

this is reflected in the R^2 values over Harvard. The saturation is less noticeable in RH100 and RH75 which was probably the reason why these quartiles were better correlated with biomass over Harvard. Since the Howland Forest data had been on the lower biomass scale saturation at RH50 and RH25 was probably not an issue, therefore all metrics performed consistently.

In this chapter LVIS data from two sites, the Harvard and Howland Forests was analyzed to assess the relationship between lidar metrics and field biomass. Preliminary analysis suggested good correlations over the Howland Forest and poor performance over the Harvard Forest. Metrics other than the simple energy quartiles were also tested over Harvard to see if there was some other metric better related to biomass without much success. RH100 was shown to still be the best predictor of biomass. Combinations of LVIS metrics didn't attain the high correlations reported in literature either. It was observed that using different types of allometric equations to predict biomass from diameter data has a significant impact on lidar performance. Use of simplified equations and excluding low biomass data from the Howland Forest dataset seemed to make the lidar-biomass relationships over the two sites comparable, with RMSE values of 50% of the biomass standard deviation the best result and corresponding R^2 values between 0.6 and 0.7 at hectare scales. At subplots the relationships between height and biomass seemed weaker, with the best R^2 values between 0.3 and 0.4 and RMSEs between 60 and 70% of the biomass variation. Combining the two datasets using simplified allometries suggest a possible saturation trend in the height-biomass relationship, explaining the poor performance of lidar over the Harvard Forest where generally higher biomass values were recorded.

CHAPTER 4

RADAR BACKSCATTER, ERROR MODEL AND OBSERVATIONS

A backscattered radar signal consists of contributions from three major scattering mechanisms: *i) single bounce*: from rough surfaces such as fields and water, *ii) double bounce*: from edges, such as those of buildings or tree trunks and the ground (named as such to indicate that the radar wave bounces twice before returning to the radar), and *iii) volumetric scattering*: from a diffuse object, such as a canopy (with an arbitrary number of bounces). The backscattered signal is effected by each scattering mechanism depending on the nature of the scattering object, radar wavelength, polarization and the radar viewing geometry among others. Polarimetric radars are capable of isolating the different scattering mechanisms for a particular target, for instance the difference between like polarizations (HH-VV) tends to be dominated by double bounce, their sum (HH+VV) by single bounce scatterers and cross-polarization (HV) is mostly volumetric scattering. In closed canopy forests such as the Harvard and Howland Forests, volumetric scattering tends to be the dominant scattering mechanism for an L-band radar viewing at incidence angles larger than 30°. Since the intensity of backscatter from a volumetric target increases as the size of the volume increases, it is expected that cross-polarization data in such a configuration should be related to the size of the trees and consequently to their biomass. However, this is complicated by several factors, for instance, the loss of a signal's energy as it propagates through a canopy (known as signal extinction) reduces the intensity of volumetric backscatter and while the extinction may be affected by factors such as moisture content in the canopy, those factors aren't always related to the size of the

volume. Furthermore, different tree species could have the same volumetric extent but different biomass levels because of their wood densities, complicating the simple backscatter to biomass inversion scenario.

4.1 Backscatter to Biomass

Regardless of the complications in relating forest biomass to radar backscatter, a large body of study exists where various backscatter and biomass relationships have been developed and analyzed. The backscatter-biomass models invariably take the form of a power-law relationship. A study of the relationship between SIR-C (L-band, C-band) and X-SAR (X-band) polarimetric data over sites in Northern Michigan presented in Dobson et. al. [27] uses power-law relationships between polarimetric backscatter variables and structural elements of a forest such as basal area, tree heights, crown and trunk biomass. Similarly, studies in the tropical forests [82, 81], or woodlands [80] using L- or C-band spaceborne SAR imagery also rely on a power-law curves to establish a relationship between radar backscatter and field biomass. In some instances, such as Le Toan et. al. [72] and Saatchi et. al. [109], multiple-linear or multiple-quadratic regressions have been used to relate various polarimetric backscatter variables to field biomass. Here, we rely on a power-law relationship between backscatter and biomass, using a simplified form of the model used in [56]

$$\sigma = \beta_0 (1 - e^{-\beta_1 M}) + \beta_2 M^\alpha e^{-\beta_1 M} \quad (4.1)$$

where σ is the radar backscatter, M is the biomass and α , an empirically determined factor, is usually fixed to 0.2. This model will be used throughout the analysis presented in this chapter to ascertain the uncertainty in biomass estimates from radar backscatter.

4.1.1 Error in radar estimates of biomass

The set of coefficients, $\boldsymbol{\beta} = [\beta_0, \beta_1, \beta_2]$ in (4.1) are generally obtained for a particular field site using non-linear regressions between radar data and corresponding field estimates of biomass. In regression analysis the accuracy of estimates is often characterized by confidence and prediction intervals. To see how these bounds can be developed for biomass estimates from radar backscatter, we consider first a general non-linear regression scenario, where a variable \mathbf{Y} is related to a set of variables given by \mathbf{X} , such that

$$\mathbf{Y} = f(\boldsymbol{\beta}, \mathbf{X}) + \boldsymbol{\epsilon} \quad (4.2)$$

where \mathbf{Y} is an $n \times 1$ vector, referred to as the *response*, while \mathbf{X} is an $n \times k$ matrix of n measurements from k *predictors*. The set of coefficients, $\boldsymbol{\beta}$, is a $p \times 1$ dimensional vector while $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_n)$ is referred to as the *error in equation*, or *model error*. The confidence interval on estimates of \mathbf{Y} for the estimated set of coefficients $\hat{\boldsymbol{\beta}}$, is given by

$$\hat{y}_i \pm t_{\alpha/2} \sqrt{S_{y_i}^2} \quad (4.3)$$

where S_{y_i} , the standard error for the i^{th} response, \hat{y}_i , can be estimated using the *delta* method, given by [11]

$$S_{y_i}^2 = \mathbf{F}'_{\boldsymbol{\beta}} \hat{\Sigma}_{\boldsymbol{\beta}} \mathbf{F}_{\boldsymbol{\beta}} \quad (4.4)$$

where $\hat{\Sigma}_{\boldsymbol{\beta}}$ is the covariance matrix of the regression coefficients and the vector $\mathbf{F}_{\boldsymbol{\beta}}$ is a set of partial derivatives of the model function with respect to the coefficients, $\hat{\boldsymbol{\beta}}$, such that

$$\mathbf{F}_{\boldsymbol{\beta}} = \left[\frac{\partial f}{\partial \beta_1}, \dots, \frac{\partial f}{\partial \beta_p} \right]^T. \quad (4.5)$$

The covariance matrix, $\hat{\Sigma}_{\boldsymbol{\beta}}$ is estimated by [32]

$$\hat{\Sigma}_{\boldsymbol{\beta}} = (\mathbf{J}^T \mathbf{J})^{-1} \quad (4.6)$$

where \mathbf{J} , sometimes referred to as the Jacobian, is an $n \times p$ matrix constructed such that each row is the vector \mathbf{F}_β^T evaluated at \mathbf{X}_i for $i = 1, \dots, n$. Similarly, a prediction interval is estimated by

$$\hat{y}_i \pm t_{\alpha/2} \sqrt{S_{y_i}^2 + \hat{\sigma}^2} \quad (4.7)$$

where $\hat{\sigma}^2$ is the variance of the error in equation, given in vector form as $\boldsymbol{\epsilon}$.

In the backscatter-biomass regression scenario, the non-linear relationship between the observed backscatter and estimated biomass is given in (4.1). Since this equation describes a *forward model*, the backscatter coefficient, σ , is the response (\mathbf{Y}) while biomass, M , is the sole predictor (\mathbf{X}), for n measurements with $k = 1$. Confidence and prediction intervals, from (4.3) and (4.7) would then characterize uncertainty of the estimated backscattering coefficient. These intervals would have to be numerically inverted to obtain the uncertainties in biomass estimates. A numerical inversion is unavoidable since a closed form solution to (4.1) does not exist.

4.1.2 Including parametric measurement error

The calculation of confidence intervals described in (4.3) and (4.7) does not include a treatment of measurement errors. If there are errors in measuring the response and the predictor (in this case backscatter and biomass respectively) the uncertainty in biomass estimates should be effected when those error-prone measurements are used to estimate the coefficients in the first place. In a general, non-linear regression scenario with *additive* measurement error, the terminology changes to include errors in both the predictor, \mathbf{X} and the response, \mathbf{Y} . Borrowing the terminology from [11, 45], the error prone measurement of the response \mathbf{Y} is denoted by \mathbf{D} , such that

$$D_i = y_i + q_i \quad (4.8)$$

where q_i is the error in measurement of response (i.e. the backscatter), such that $E(q_i | y_i, \mathbf{x}_i) = 0$ and $Var(q_i | y_i, \mathbf{x}_i) = \sigma_{qi}^2$. Similarly, an error prone measurement of

the predictor (the biomass), \mathbf{X} is given by \mathbf{W} , such that

$$\mathbf{W}_i = \mathbf{x}_i + \mathbf{u}_i, \quad i = 1, \dots, n, \quad (4.9)$$

where \mathbf{u}_i is a vector of the error in measurement of the predictors with $E(\mathbf{u}_i | y_i, \mathbf{x}_i) = 0$, $Cov(\mathbf{u}_i | y_i, \mathbf{x}_i) = \Sigma_{ui}$ and $Cov(\mathbf{u}_i, q_i | y_i, \mathbf{x}_i) = \Sigma_{uqi}$. The covariance matrix, Σ_{ui} (the covariance of errors in the measurement of \mathbf{X}_i , the predictors) is a $k \times k$ matrix while Σ_{uqi} is the covariance between the errors in predictor and response. Since there is no practical reason for errors in measurement of backscatter and biomass to be related, Σ_{uqi} is ignored here by setting it to zero. Finally the relationship used in the non-linear regression algorithms is given by

$$\mathbf{D} = f(\boldsymbol{\beta}, \mathbf{W}) + \boldsymbol{\epsilon} \quad (4.10)$$

where the error in equation, $\boldsymbol{\epsilon}$, remains the same $n \times 1$ vector with variance $\hat{\sigma}^2$. The calculation of confidence and prediction intervals is still the same as described in (4.3) and (4.7), however estimation of the covariance matrix, $\hat{\Sigma}_{\boldsymbol{\beta}}$ is no longer as simple as given in (4.6). The covariance matrix must now account for three more error sources, namely q_i , the error in response, \mathbf{u}_i , the error in predictors and the covariance of the two errors Σ_{uqi} . This problem is simpler in a linear-regression scenario, where equations have been developed that allow for estimation of the covariance matrix with additive measurement error [11]. The case of non-linear regressions is however much more complicated. There is no one method for estimating the covariance matrix, $\Sigma_{\boldsymbol{\beta}}$ in a non-linear regression setting with measurement error in both the response and predictor. An analytical solution in such a case has to be tailored to the particular model function. While such an analysis could be conducted for the backscatter-biomass case, doing so would require much deeper insight into the nuances of statistics, and though that would be of some academic value it is avoided here for simplicity.

The most common workaround is to use simulation or Monte-Carlo type techniques. The problem is two fold, the first is to correct the bias in coefficients, β , introduced by measurement error, and second to estimate the covariance matrix itself. The solution to the first problem is handled by the method called SIMEX [23], while the second by the *bootstrap* [11].

4.1.2.1 Parametric bootstrapping

The more common form of the bootstrap, called the *one-stage bootstrap* estimates the statistics of the data, such as the variance, by using different or all possible combinations of the data (with or without replacing samples). Such a method is useful as it makes no assumptions about the statistics of the error. It does however lack the ability to assess the impact of noise that may not be present in that particular dataset, limiting its usefulness when an analysis of the measurement errors is required. Parametric bootstrap, or a *two-stage bootstrap*, on the other hand, simulates a large number of datasets using realizations of measurement errors with the intent of relying on sample statistics to estimate standard errors and thus the confidence intervals. A description of the method follows.

Given a set of predictors \mathbf{W}_i (in this case, simply biomass) and an estimate of their error variances $\hat{\sigma}_{ui}^2$, we construct a large number of imperfectly measured predictors such that [11]

$$\mathbf{W}_{bi} = \mathbf{W}_i + \mathbf{u}_{bi} \quad i = 1, \dots, n, \quad (4.11)$$

where $\mathbf{u}_{bi} \sim N(0, \hat{\sigma}_{ui}^2)$ for $b = 1, \dots, B$, where B is the number of bootstrap realizations and should be large. If the number of predictors were larger than one, then $\hat{\sigma}_{ui}^2$ is replaced with $\hat{\Sigma}_{ui}$ which requires knowledge of not just the variance but also the covariances between the errors in measuring the predictors. The response without any measurement error but with an error in equation (the model error) is then calculated such that [11]

$$Y_{bi} = f(\hat{\beta}_c, \mathbf{W}_{bi}) + \epsilon_b \quad i = 1, \dots, n, \quad (4.12)$$

where $\hat{\beta}_c$ are unbiased coefficients (obtained using SIMEX, to be described later) and $\epsilon_{bi} \sim N(0, \hat{\sigma}^2)$. The error-prone measurement of the response is then calculated using [11]

$$D_{bi} = Y_{bi} + q_{bi} \quad i = 1, \dots, n, \quad (4.13)$$

where $q_{bi} \sim N(0, \hat{\sigma}_{qi}^2)$, and $\hat{\sigma}_{qi}^2$ is the estimate of variance of the measurement error in the response (in this case backscatter). For each bootstrap sample, a set of coefficients $\hat{\beta}_b$ are estimated using non-linear fits between \mathbf{D}_b and \mathbf{W}_b . Finally, the covariance matrix, $\hat{\Sigma}_\beta$, is estimated by computing the sample covariance of the coefficients, $\hat{\beta}_b$ for $b = 1, \dots, B$.

The implementation of this method is complicated by problems of non-convergence of the non-linear fitting algorithms. Often, convergence relies on a good initial guess, which is not easy to automate. Since the number of bootstrap samples, B , needs to be large enough to have meaningful estimates of the covariance matrix, non-convergence could potentially make the problem of estimating the confidence intervals impossibly hard. Furthermore, obtaining information about the variance of the error in equation, $\hat{\sigma}^2$ is not simple. This quantity is generally not known *a priori*, certainly not in the backscatter-biomass case. While it is logical to use the mean squared residuals from the fit as an estimate, the residuals tend to be contaminated by measurement error in both the response and the predictor [11]. An estimate of the variance of ϵ would thus require removing, or deconvolving, the effects of measurement error from the residuals. This is often quite complicated and model dependent. An approximate solution can be borrowed from linear regression, given by [11]

$$\hat{\sigma}^2 = \frac{\sum_i (D_i - f(\hat{\beta}_c, \mathbf{W}_i))^2}{n - k} - \hat{\sigma}_q^2 - \mathbf{F}_X^T \hat{\Sigma}_u \mathbf{F}_X \quad (4.14)$$

where

$$\hat{\sigma}_q^2 = \sum_i \frac{\sigma_{qi}^2}{n}, \quad \hat{\Sigma}_u = \sum_i \frac{\hat{\Sigma}_{ui}}{n} \quad (4.15)$$

and

$$\mathbf{F}_X = \left[\frac{\partial f}{\partial X_1}, \dots, \frac{\partial f}{\partial X_k} \right]^T. \quad (4.16)$$

In the backscatter-biomass case of (4.1), since \mathbf{X} is simply M , so \mathbf{F}_X reduces to a scalar, $\mathbf{F}_X = \partial f / \partial M$. In essence, the use of (4.14) through (4.16) has the effect of subtracting estimates of measurement error from the fit residuals to approximate model error or error in equation.

4.1.2.2 SIMEX: Simulation Extrapolation

The introduction of measurement error in both the response and predictor biases the estimated regression coefficients [11]. The SIMEX method attempts to estimate the amount of bias caused by error in measurement of \mathbf{X} and \mathbf{Y} [23]. These unbiased coefficients are needed for the bootstrap procedure described above. Essentially, SIMEX assesses the trend in coefficients as a function of measurement error. Mathematically, it postulates that $\beta_j(\lambda)$ for $j = 1, \dots, p$ (for p coefficients) is unbiased for $\lambda = -1$ if the covariance matrix of the measurement errors of the predictors is given by $Cov(\mathbf{u}_i) = (1 + \lambda) \Sigma_{ui}$ instead of Σ_{ui} and the variance of the measurement error of the response by $(1 + \lambda) \hat{\sigma}_q^2$ rather than $\hat{\sigma}_q^2$. The procedure for computing $\beta_j(\lambda)$ at $\lambda = -1$ is similar to the bootstrap itself. Here, for each λ , SIMEX predictors with known covariances are calculated such that [11]

$$\mathbf{W}_{si}(\lambda) = \mathbf{W}_i + \lambda^{1/2} \mathbf{u}_{si} \quad i = 1, \dots, n, \quad (4.17)$$

for $s = 1, \dots, S$ (for large values of S) with $Cov(\mathbf{u}_i) = \Sigma_{ui}$. Similar, the error in response is added for each λ to calculate the SIMEX $D_{si}(\lambda)$, such that

$$D_{si}(\lambda) = D_i + \lambda^{1/2} q_{si} \quad i = 1, \dots, n. \quad (4.18)$$

Coefficients $\hat{\beta}_s(\lambda)$ are estimated for each SIMEX sample and averaged to calculate

$$\hat{\beta}(\lambda) = \sum_s \frac{\hat{\beta}_s(\lambda)}{S}. \quad (4.19)$$

Typically, second order quadratic polynomials are fitted to each of the coefficients, $\beta_j(\lambda_s)$ and the unbiased estimate of $\hat{\beta}$ is obtained by evaluating those polynomials at $\lambda = -1$ [23, 11].

The parametric bootstrap and SIMEX comprise a mathematical framework where errors in measurements of biomass and backscatter can be included in an assessment of the accuracy of biomass estimates from radar backscatter. The errors in measurements of biomass at the Howland and Harvard Forests were characterized in Chapter 2. A description of the errors in measuring backscatter follows.

4.2 Backscatter Error Model

Uncertainties in radar measurements are a consequence of the random nature of the scattering process and the complexity of a radar instrument itself. The contributors to what would amount as error in a radar backscatter measurement when biomass estimation is of interest include, among others, image speckle, calibration errors, topographic effects and variability due to atmospheric changes such as soil moisture. The contribution of each error source to the total error budget depends on a slew of factors ranging from systematic instrument parameters to weather conditions. To be able to inform the framework laid out in previous sections regarding error in biomass estimation from radar, a thorough accounting of the radar backscatter error budget is in order. The radar error model developed in [56], models the total error in backscatter, $\Delta\sigma$, as a sum of four major error sources

$$\Delta\sigma = \Delta\sigma_s + \Delta\sigma_t + \Delta\sigma_c + \Delta\sigma_a \quad (4.20)$$

where $\Delta\sigma_s$ captures the effect of image speckle, thermal and multiplicative noise, $\Delta\sigma_t$ describes the temporal variation in radar backscatter unrelated to biomass change, $\Delta\sigma_c$ is the error in backscatter calibration and $\Delta\sigma_a$ is the error in backscatter due to incorrect area projection normalization.

4.2.1 Image speckle, thermal and multiplicative noise

The first term in (4.20), $\Delta\sigma_s$, combines the effect of image speckle (due to the coherent nature of radar scattering), thermal noise and sources of multiplicative noise such as ambiguities, impulse response sidelobes and quantization noise. This term is modeled in [56] as a modified version the scatterometer error equation, given by

$$\Delta\sigma_s = \frac{1}{\sqrt{N}} \left(\frac{1}{\sqrt{N_{o_s}}} + \frac{1}{\sqrt{N_{o_t}}} \frac{1 + \text{SNR/MNR}}{\text{SNR}} \right) \sigma \quad (4.21)$$

where N is the number of spatial looks or averages, while N_{o_t} is the total number of repeat observations as the sum of speckle identical N_{o_i} and speckle diverse N_{o_s} observations, $N_{o_t} = N_{o_s} + N_{o_i}$, where this distinction is introduced to account for observations from slightly different observation geometries where speckle is correlated (speckle identical) such as in interferometry and observation geometries that are very dissimilar where speckle is uncorrelated (speckle diverse) such as from ascending and descending passes of a satellite. The variable SNR in (4.21) refers to signal-to-noise ratio from thermal noise only while MNR is the multiplicative-noise ratio often due to observation errors such as range and azimuth ambiguities. The number of spatial looks, N can be computed for an imaged area using

$$N = \frac{A_g \sin \theta_i}{\kappa_\rho \Delta\rho \kappa_s \Delta s} \quad (4.22)$$

where A_g is the area of region over which image pixels are spatially averaged (for example the area of the field sites where biomass is estimated), θ_i is the local incidence

angle, $\Delta\rho$ and Δs are the range and azimuth resolution (or spacing) respectively, while κ_ρ and κ_s are the broadening factors in range and azimuth, accounting for the loss of resolution due to use of windowing in the image formation process. The range resolution is given by

$$\Delta\rho = \frac{c}{2B} \quad (4.23)$$

where c is the speed of light and B is the signal bandwidth. The maximum azimuth resolution, Δs for a synthetic aperture system is given by

$$\Delta s = \frac{L_{az}}{2} \quad (4.24)$$

where L_{az} is the azimuth dimension of the antenna. The broadening factors, κ_ρ and κ_s can be approximated by

$$\kappa_{\rho,s} \approx 1.6363 - 0.6363\sqrt{\eta} \quad (4.25)$$

where the parameter η depends on the type of windowing employed during the compression of the range or azimuth spectra, for instance $\eta = 1$ corresponds to uniform weighting, $\eta = 0$ corresponds to the Hanning window, $\eta = 0.08$ corresponds to the Hamming window etc. The signal-to-noise ratio, SNR, is given by

$$\text{SNR} = \frac{\sigma}{\bar{\sigma}_0} \quad (4.26)$$

where σ is the backscatter value and $\bar{\sigma}_0$ is the system-specific noise equivalent backscatter coefficient. The multiplicative noise ratio, MNR, is given by

$$\frac{1}{\text{MNR}} = \text{ISLR}_\rho + \text{ISLR}_s + \frac{1}{\text{AMB}_t} + \frac{1}{\text{QNR}} \quad (4.27)$$

where ISLR_ρ and ISLR_s are the integrated sidelobe ratios in range and azimuth for sidelobes arising from the range and azimuth compression techniques associated with

a pulse compression SAR system. AMB_t is the total signal-to-ambiguity ratio (including the effect of both range and azimuth ambiguities), referring to the ambiguities that arise from the uncertainty in associating a backscattered radar signal with its corresponding transmitted signal. The quantization-noise-ratio, QNR, a ratio of the signal power to the quantization noise power, is dependent on the quantization scheme, but can be approximated by

$$\text{QNR} \approx 2.0 + 6.02b \quad (4.28)$$

for QNR expressed in dB and b referring to the number of bits used in quantization. The integrated sidelobe ratios, ISLR_ρ and ISLR_s are essentially ratios of the total energy in the sidelobes to the energy in the mainlobe of a compressed range or azimuth signal. These ratios are dependent on the type of window used in the image formation process, however, they are approximated by polynomial fits to numerical evaluations of the energy integrals as a function of the broadening factor η as [56]

$$\text{ISLR}_{\rho,s}(\eta) = -46.965 + 104.11\eta - 112.59\eta^2 + 43.124\eta^3 \quad (4.29)$$

with the ISLR expressed in dB and η the same window specific parameter described earlier.

4.2.2 Temporal variation

Temporal variability in radar backscatter, given by the variable $\Delta\sigma_t$ in (4.20) can in principle encompass effects from many physical processes such as wind, weather, flooding etc. Although change in forest biomass may also affect radar backscatter, that change would not be considered an error, rather an observation. Here, temporal variation refers to changes in backscatter from one observation to another from all processes that have no impact on forest biomass. It is difficult to have a generic model that explains temporal variation for all possible scenarios, as a result this error is invariably treated differently from one data set to another.

4.2.3 Error in backscatter calibration

The sampled radar signal, a unitless digital number ranging from 0 to $2^b - 1$ for a b -bit digitizer, is converted to physical units by normalizing that number by a calibration constant. The calibration error term, $\Delta\sigma_c$ refers to the errors introduced by unknowns in the computation of the calibration constant. Systematic calibration errors arise from uncertainties in pointing and geolocation, while random errors arise from uncalibrated behavior of radar electronics. The calibration error is modeled in [56] as

$$\Delta\sigma_c = \frac{1}{\sqrt{N_{ot}}}\Delta\sigma_{ran} + \frac{1}{\sqrt{N_{ot}}}\Delta\sigma_p + \frac{1}{\sqrt{N_{os}}}\frac{1}{\sqrt{N}}\sqrt{\frac{A_{dem}}{A_{pix}}}\Delta\sigma_{geo} \quad (4.30)$$

where $\Delta\sigma_{ran}$ are the random errors from uncalibrated behavior of the electronics, $\Delta\sigma_p$ the errors in calibration due to pointing uncertainty and $\Delta\sigma_{geo}$ the error due to inaccuracies of the digital elevation model (DEM) used for orthorectification, while A_{pix} and A_{dem} are the areas of the image and DEM pixels respectively. Since pointing and random calibration errors tend to be correlated for a particular scene, they are unaffected by spatial averaging and are thus modeled without any dependence on N . The random calibration error term $\Delta\sigma_{ran}$ captures the effect of uncertainty in estimating the instrument gain needed to compute the calibration constant. It is calculated from an instrument specific residual random calibration term, $\Delta\sigma_{dB}$ (specified in decibels), by

$$\Delta\sigma_{ran} = \sigma_t (10^{\Delta\sigma_{db}/10} - 1) \quad (4.31)$$

where σ_t is the true backscatter. The effects of antenna patterns on radar backscatter are usually removed through calibration by using measured antenna patterns and an estimate of the antenna pointing direction. Uncertainties in the antenna pointing results in an incorrect compensation of the backscatter, and therefore error in a backscatter measurement of biomass, denoted by $\Delta\sigma_p$. This error term would depend on the antenna patterns themselves. This error, however can be approximated by assuming a two-way *sinc* pattern for the antenna gain, such that

$$\Delta\sigma_p \approx -\sigma \left(\frac{\Delta G_{r,\rho}}{G_{r,\rho}} + \frac{\Delta G_{r,s}}{G_{r,s}} + \frac{\Delta G_{t,\rho}}{G_{t,\rho}} + \frac{\Delta G_{t,s}}{G_{t,s}} \right) \quad (4.32)$$

where $G_{x,y}$ refers the transmit and receive (t, r) gain compensation in range and azimuth (ρ, s). Using global averaged values of the gain compensation term, the relative gain error in the range direction is given as a function of the pointing error, $\Delta\theta$ as

$$\frac{\Delta G_{x,\rho}}{G_{x,\rho}} = \frac{4\Delta\theta}{\theta_{1/2}} \ln \left[\text{sinc} \left(\frac{\pi}{2k} \right) \right] \quad (4.33)$$

where $\theta_{1/2}$ is the elevation half-power beamwidth, and k is a beam shaping factor (assumed to be 1.136 for a uniformly weighted array; a common configuration). Similarly, the relative gain error in the azimuth direction is given by

$$\frac{\Delta G_{x,s}}{G_{x,s}} = \frac{2k\Delta\theta}{\pi} \left(\frac{\ln [\text{sinc}(\pi/2k)]}{\int_0^{\pi/2k} \text{sinc}^2(x) dx} \right). \quad (4.34)$$

Finally, the geolocation error, $\Delta\sigma_{geo}$, arising from errors in the DEM that cause errors in estimating the radar pointing direction and therefore an incorrect antenna gain calibration, is calculated from (4.33) with the pointing error given by

$$\Delta\theta = \frac{\Delta h_{dem}}{\rho} \quad (4.35)$$

where Δh_{dem} is the height accuracy of the DEM and ρ is the range.

4.2.4 Errors due to pixel area normalization

In a calibrated radar, the radar-cross section is normalized to the scattering area to calculate the radar backscattering coefficient, σ . Ground topography can have a significant impact on the scattering area and therefore the backscatter coefficient. To estimate biomass, topographic effects in backscatter must be accounted for in the processing chain. However, uncertainties in the ground topography introduce error

in the corrected backscatter. This effect is captured in the term $\Delta\sigma_a$ in (4.20), and is modeled in [56] as

$$\Delta\sigma_a = \frac{1}{\sqrt{N_{os}}} \sqrt{\frac{A_{dem}}{A_{pix}}} \frac{1}{\sqrt{N}} \left(\frac{\Delta A_c}{A_c} \right) \sigma \quad (4.36)$$

where the relative area projection error $\Delta A_c/A_c$, a function of the DEM height error, Δh_{dem} is given by

$$\frac{\Delta A_c}{A_c} = \frac{\sin(2\tau_\rho)\sin^2(\tau_s) + \sin(2\tau_s)\sin^2(\tau_\rho)}{2 - 2\sin^2(\tau_\rho)\sin^2(\tau_s)} \Delta\tau + \frac{\cos(\theta - \tau_\rho - \tau_s)}{\sin(\theta - \tau_\rho)\cos(\tau_s)} \Delta\tau \quad (4.37)$$

where τ_ρ and τ_s are the mean slopes in the range and azimuth direction for incidence angle θ , while $\Delta\tau$ is the error in determining the range or azimuth slopes (considered to be the same here), given by

$$\Delta\tau = \frac{\Delta h_{dem}}{L_{dem}(1 + \tan^2(\tau_x))} \quad (4.38)$$

where Δh_{dem} is the DEM height error, L_{dem} is the DEM posting and τ_x is slope in either range or azimuth direction.

4.3 Radar Measurements

During the summer of 2009, in conjunction with the field campaigns outlined in chapter 2, NASA JPL's UAVSAR (uninhabited aerial vehicle synthetic aperture radar) [55] was flown over the Harvard and Howland forest regions. The flightlines were designed to be in a race-track configuration so that the left looking antenna could image the same area from two different aspect angles (by flying at opposite headings, hence the term racetrack). At the Harvard Forest, UAVSAR collected a total of forty scenes on five days, the 6th, 8th, 13th, 16th and 17th of August, with twenty scenes from each heading of 5° and 185°. At the Howland Forest, UAVSAR

was flown at headings of 167° and 347° collecting ten scenes at each heading on four days, the 5^{th} , 7^{th} and the 14^{th} of August. Table 4.1 lists some key parameters of the UAVSAR instrument.

Table 4.1: UAVSAR instrument parameters.

Parameter	Value
Frequency	1.26GHz
Bandwidth	80MHz
Polarization	HH, HV, VH, VV
Look Angles	$25^\circ - 65^\circ$
Swath Width	16km
Resolution	$1.6m \times 0.66m$

UAVSAR data is distributed by NASA/JPL in either the slant-range geometry or as ground projected images using the 30m resolution SRTM DEM for orthorectification. In either format, the antenna pattern artifacts are removed in calibration by using precise antenna models and best estimates of pointing angle. The two artifacts that do remain in the distributed data are due to pixel area effects from topographic variations and the dependence of radar backscatter on incidence angle [43, 125]. Before UAVSAR backscatter data can be used to estimate biomass, both these artifacts must be removed [12].

Topographic variations cause the area of imaged pixels to be projected differently in the radar look direction. This variation causes the backscattering coefficient, which is normalized to the scattering area, to vary with topography. For the UAVSAR data used in this analysis, topographic artifacts are removed using the algorithm outlined in [126], where the corrected backscatter coefficient σ_{pix} is given by

$$\sigma_{pix} = \sigma \cos \psi \quad (4.39)$$

where σ is the calibrated, ground-projected backscatter coefficient (as distributed by NASA/JPL), and the projection angle ψ is given by

$$\cos \psi = \hat{n} \cdot (\hat{x} \times \hat{R}) \quad (4.40)$$

where \hat{n} is the unit surface normal, \hat{x} is the unit vector in the along-track direction, and \hat{R} is the unit look vector. The surface normal is computed from the resampled co-registered 30m SRTM DEM that is distributed with the UAVSAR data, while the look vector is computed from UAVSAR peg parameters [54]. Methods for treating regions of layover and shadowing do exist, however, there aren't any such regions around the Harvard and Howland Forests. For this analysis such pixels are simply masked to zero. The second artifact, that of the dependence of volume backscatter on the incidence angle, or the backscatter law, is compensated by [125, 43]

$$\sigma_{inc} = \sigma_{pix} \left(\frac{\cos \theta}{\cos \theta_i} \right) \quad (4.41)$$

where θ is the reference incidence angle (where no correction is applied) while θ_i is the local incidence angle, given by [126]

$$\cos \theta_i = \hat{n} \cdot \hat{R} \quad (4.42)$$

where \hat{n} and \hat{R} are the unit surface normal and look vectors described before. While the pixel area correction is valid everywhere, the backscatter law in (4.41) is only valid over forested regions since it relies on an assumption of the trend in volume backscatter as a cosine of the local incidence angle. This trend isn't valid for all scattering mechanisms. Figure 4.1 shows the two corrections applied to one of the scenes acquired at a heading of 5° over the Harvard Forest. The left most image is the calibrated, ground projected image distributed by JPL. The variation in backscatter

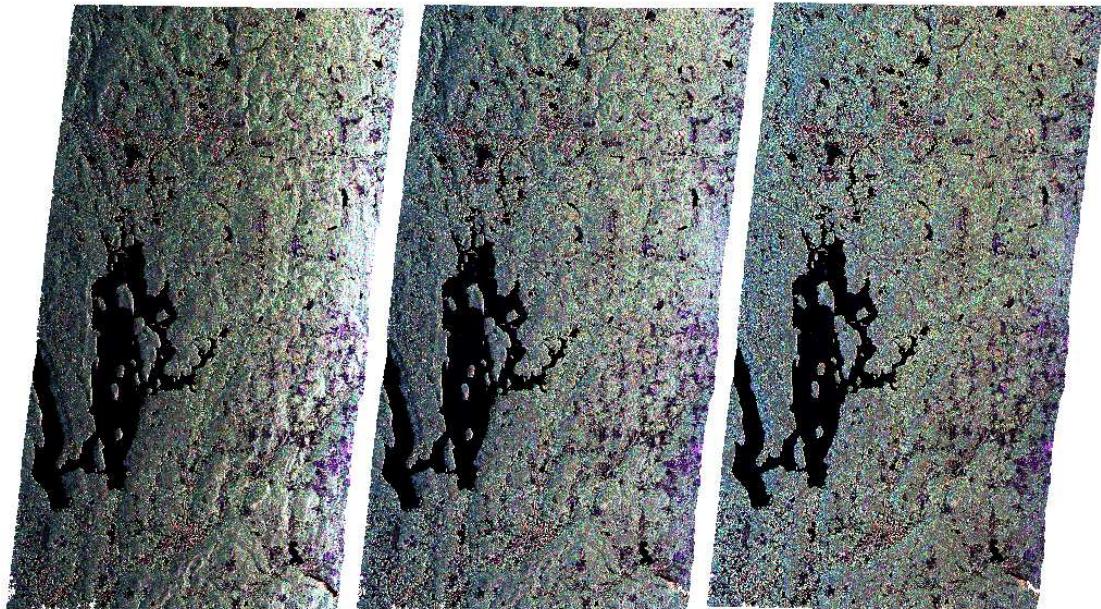


Figure 4.1: Polarimetric composite (HH-red, HV-green, VV-blue) of three stages of the UAVSAR backscatter correction, the left plane shows the calibrated image distributed by JPL, the center image is corrected for pixel area variations while the image on the right is corrected for both the pixel area variations and the backscatter law.

intensity as a result of pixel area and incidence angle variation across the swath is quite apparent. The center image has been corrected for pixel area variations due to topography, given by (4.39). The effect of this correction is visible as a decrease in backscatter intensity from hills facing the radar as well as a reduction of the cross-swath trend in backscatter. The right-most image in Figure 4.1 has been corrected for the backscatter law using (4.41) as well as pixel area variations. As expected (and desired) almost all topographic and cross-swath trends are visibly absent from this image. There are however some trends in the polarimetric composition, evident by the dominance of the blue component, representing the VV-channel, at father ranges. Since the backscatter-law assumed for this analysis is independent of polarization this trend is expected [125].

Table 4.2: List of UAVSAR scenes chosen over the Harvard and Howland Forests.

Harvard Forest			Howland Forest		
Line	Flight	Date	Line	Flight	Date
00502	09057-001	08/08/09	16701	09054-008	08/05/09
00502	09065-005	08/17/09	16701	09054-010	08/05/09
00503	09057-003	08/08/09	16701	09056-008	08/07/09
00505	09057-005	08/08/09	16701	09056-012	08/07/09
00506	09065-007	08/17/09	16701	09061-008	08/14/09
00508	09060-000	08/13/09	16701	09061-010	08/14/09
00509	09057-007	08/08/09	16702	09054-012	08/05/09
00509	09060-002	08/13/09	16702	09054-014	08/05/09
00512	09065-003	08/17/09	16702	09054-016	08/05/09
00513	09065-001	08/17/09	16702	09056-014	08/07/09

Table 4.2 lists the identifiers for ten UAVSAR scenes chosen for analyzing the backscatter-biomass relationship over the Harvard and Howland Forests. The acqui-

sition date for each scene is also listed. All of these scenes are corrected for pixel area and the backscatter law, as described earlier. The UAVSAR orthorectified imagery is distributed in the equiangular projection using the WGS84 ellipsoid with equal pixel spacing of 5.556×10^{-5} degrees (approximately 6.2m in the E/W direction and 4.5m in the N/S direction). Pixels corresponding to a particular field site are chosen from the imagery based on the amount of overlap between a pixel and its corresponding ground site. An analysis of the backscatter-biomass relationship over the Harvard and Howland forests using these scenes follows.

4.3.1 Backscatter-biomass relationships

With field estimates of biomass, discussed in chapter 2 and backscatter intensity data from UAVSAR collected at approximately the same time this dataset allows for a ready analysis of the relationship between radar backscatter and biomass in the temperate forests of the North-eastern United States. Although a framework has been established that allows for an analysis of errors in biomass estimates from backscatter in the presence of measurement error in both backscatter and field-biomass, it is instructive to look at biomass uncertainty without considering measurement error first.

4.3.1.1 Harvard Forest

Figures 4.2, 4.3, 4.4 plot the subplot and hectare scale field estimates of biomass obtained from the three different allometries (Single-site: Ter-Mikaelian [118], Ensemble: Jenkins [65] and BLUE: Lambert [71]) against average cross-polarized (HVHV) backscatter data from the ten UAVSAR scenes listed in Table 4.2 over the 240 subplots and 15 hectares at the Harvard Forest. The backscatter, plotted at the subplots or the hectare scales is estimated by averaging ground projected UAVSAR pixels such that

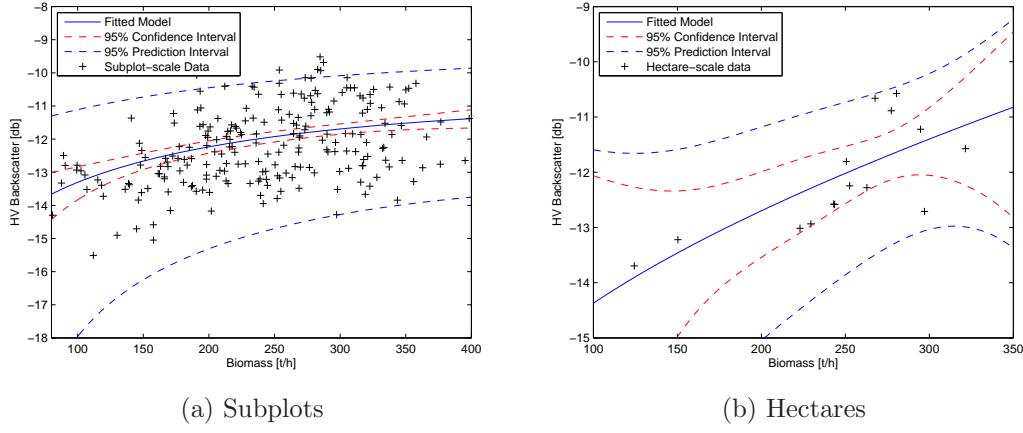


Figure 4.2: Cross-pol backscatter from UAVSAR plotted against field biomass estimates from the Ter-Mikaelian allometry at Harvard.

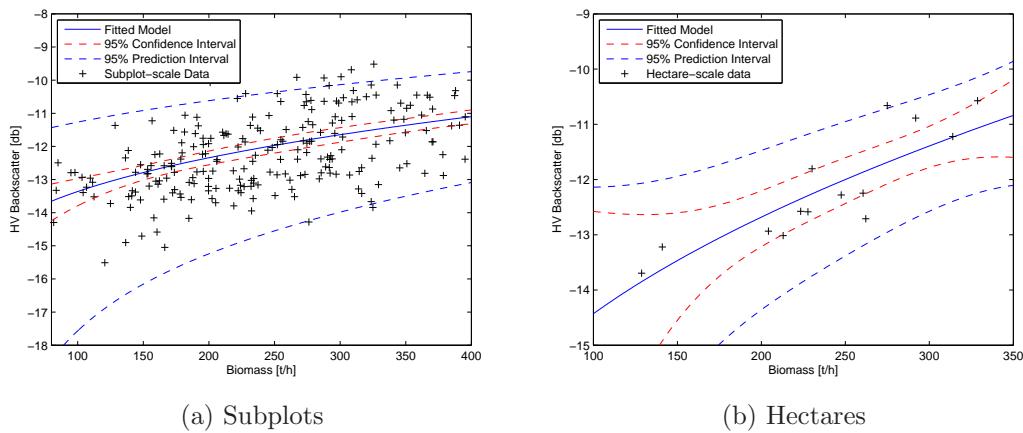


Figure 4.3: Cross-pol backscatter from UAVSAR plotted against field biomass estimates from the Jenkins allometry at Harvard.

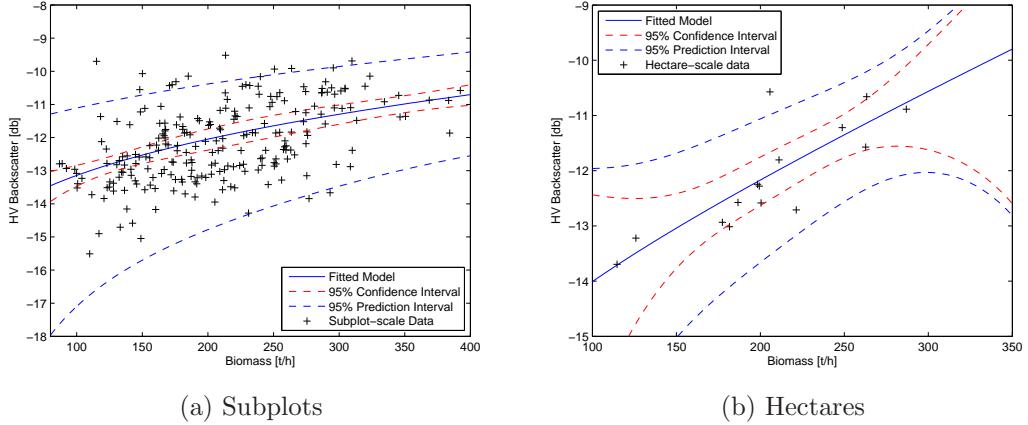


Figure 4.4: Cross-pol backscatter from UAVSAR plotted against field biomass estimates from the Lambert allometry at Harvard.

$$\bar{\sigma} = \frac{1}{N} \frac{1}{N_{ot}} \sum_p^N \sum_q^{N_{ot}} \sigma_{pq} \quad (4.43)$$

where N is the number of spatial looks and N_{ot} is the number of repeat observations of the single-look backscatter, σ_{pq} . Since the backscatter data plotted here is chosen from the ground projected imagery which is already spatially multi-looked, the number of ground projected pixels averaged is not the same as the number of looks, N given by (4.22). Since the incidence angle at the Harvard forest for all the scenes was designed to be around 40° , the number of looks, N for subplots of area $25\text{m} \times 25\text{m}$, with a native UAVSAR single-look resolution of $1.6\text{m} \times 0.66\text{m}$ and broadening factors of 1.45 in range and azimuth (for a Hamming window) is approximately 187. While at hectare scales, $N \approx 2790$. The ten scenes from the lines listed in Table 4.2 were all collected from almost the same look geometry (for repeat-pass interferometry), they are considered to be speckle identical, therefore, here $N_{ot} = N_{oi} = 10$. The confidence and prediction intervals are calculated from (4.3) and (4.7) respectively with $t_{\alpha/2} = 1.96$. The covariance matrix of coefficients, $\hat{\Sigma}_\beta$, is estimated from (4.6) while the error in equation, needed for calculating the prediction intervals, is estimated

from mean square residuals of the non-linear fit to the backscatter-biomass forward model given in (4.1).

4.3.1.2 Howland Forest

Figures 4.5, 4.6 and 4.7 plot field biomass estimates from the 368 subplots and 23 hectares at the Howland Forest against UAVSAR data from ten scenes processed from lines listed in Table 4.2. Because of the similarity in radar observations and field campaigns, the number of looks, N and number of repeat observations N_{ot} remain unchanged. Similar to the Harvard Forest, only cross-polarized (HVHV) data from

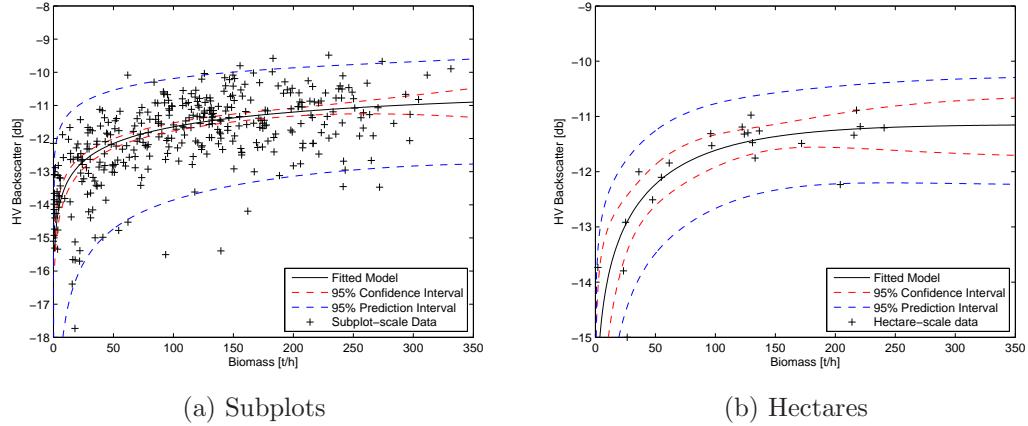


Figure 4.5: Cross-pol backscatter from UAVSAR plotted against field biomass estimates from the Ter-Mikaelian allometry at Howland.

UAVSAR is chosen for analysis here. Because of the presence of low-biomass sites in the Howland Forest dataset, the fitted curve is more similar to results from published studies [63, 72]. The results at Harvard Forest were unconventional because of the absence of any field sites with biomass values less than 80 tons/ha. The confidence and prediction intervals however are not that different in either case, with the width of the interval mostly dependent on the number of observations around the particular biomass value.

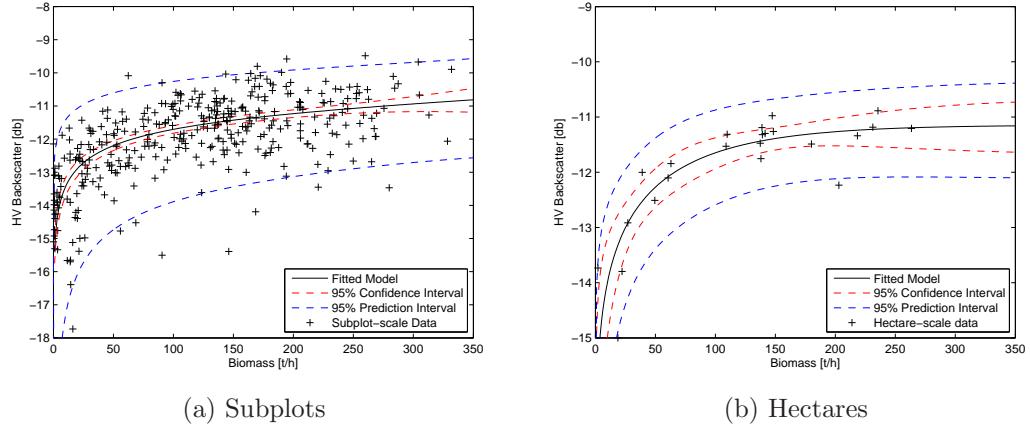


Figure 4.6: Cross-pol backscatter from UAVSAR plotted against field biomass estimates from the Jenkins allometry at Howland.

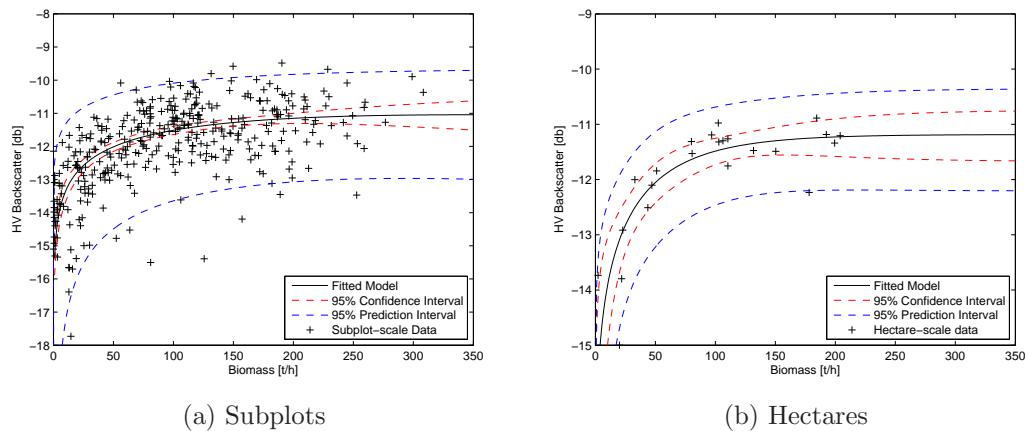


Figure 4.7: Cross-pol backscatter from UAVSAR plotted against field biomass estimates from the Lambert allometry at Howland.

4.3.1.3 Combining Harvard and Howland Forest datasets

The combined biomass values at the Harvard and Howland Forests span a much wider range, from 0-500tons/ha for subplots and 0-350tons/ha for hectares. Because of the similarity of radar observations and field data, it is possible to simply combine the two datasets so that the intervals can be computed over larger ranges of biomass values. Figures 4.8, 4.9 and 4.10 plot mean UAVSAR backscatter against field biomass estimates from the three allometric equations over both the Harvard and Howland Forest datasets. The backscatter-biomass curves from the three al-

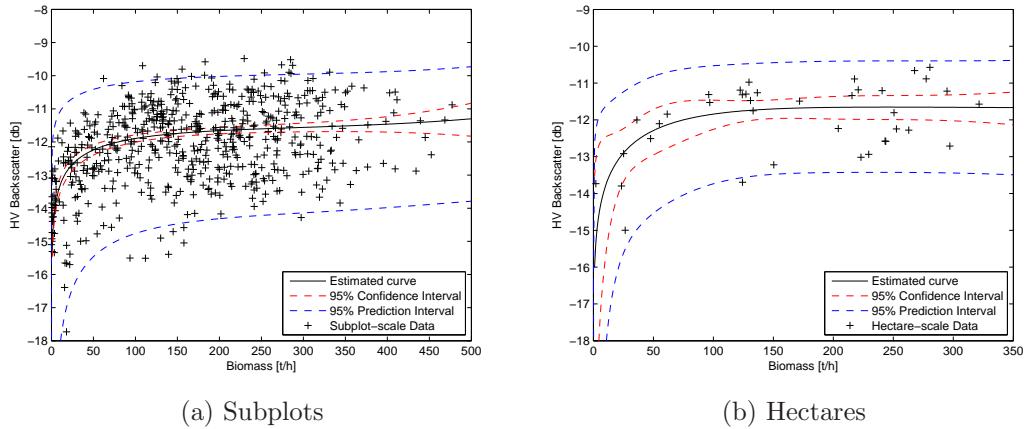


Figure 4.8: Backscatter from UAVSAR plotted against field biomass estimates from the Ter-Mikaelian allometry for data combined from the two sites.

lometries over either sites individually or combined, all have the familiar saturation trend in radar backscatter for biomass values higher than approximately 100tons/ha. The confidence intervals are wider at hectare scales than at subplot-scales primarily due to the larger number of samples used for regression. The prediction intervals are dominated mostly by the model error (or error in equation), which is not perfectly known, rather estimated from fit residuals. The fits for subplot-scale combined data using Ensemble (Jenkins) and BLUE (Lambert) allometries shows a deviation from the saturation trend at high biomass values, where backscatter apparently increases with increase in biomass. This, however, is an artifact of the lack of samples at such

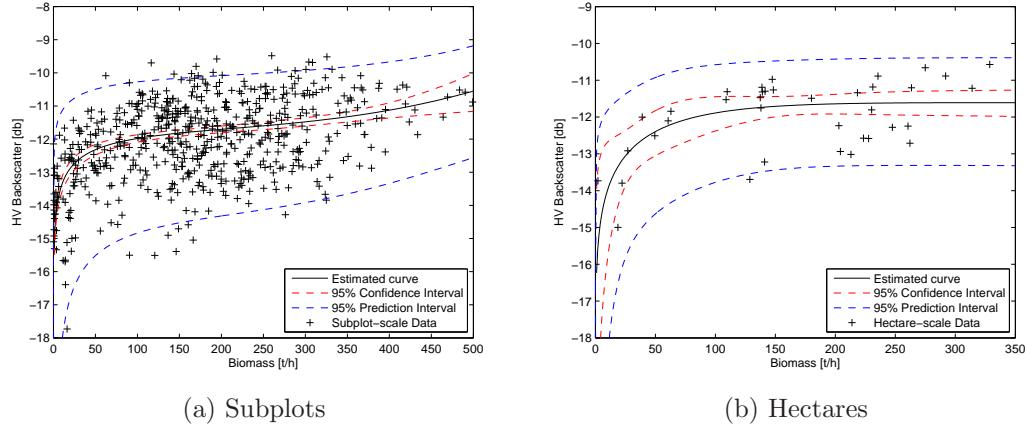


Figure 4.9: Backscatter from UAVSAR plotted against field biomass estimates from the Jenkins allometry for data combined from the two sites.

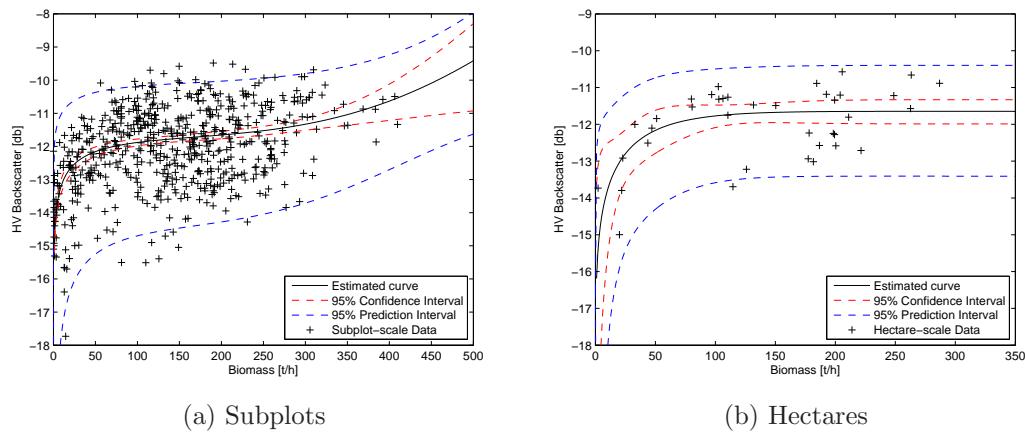


Figure 4.10: Backscatter from UAVSAR plotted against field biomass estimates from the Lambert allometry for data combined from the two sites.

large biomass values, indicated by the substantial increase in the width of the confidence intervals. The width of these intervals would seem to suggest that subplot-level estimates are more accurate. Figure 4.11 plots the inverted confidence intervals for

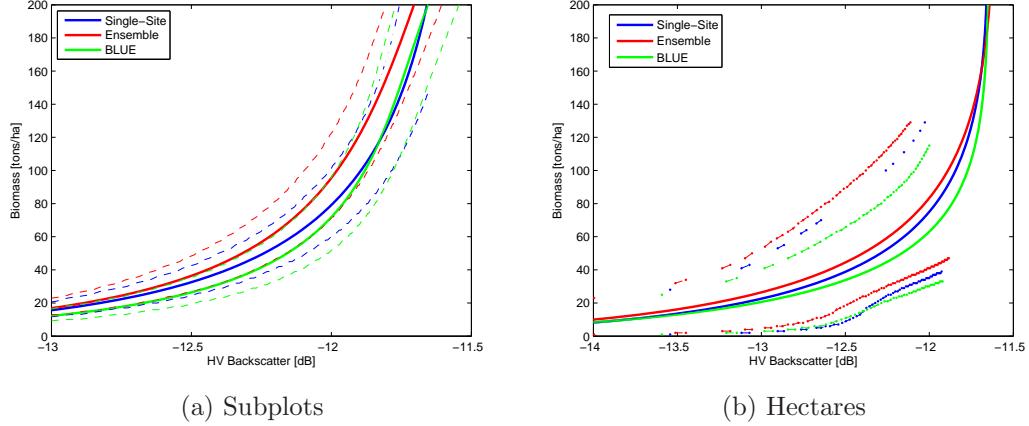


Figure 4.11: Confidence intervals on biomass estimates using data from both Harvard and Howland Forests.

the three allometries for the combined dataset. As such, the confidence intervals on subplot-level estimates are much tighter and because of the non-saturated nature of the fits the intervals do exist for larger biomass values. At hectare scales, however, the confidence intervals are much wider. Furthermore due to the saturated nature of the curves at these spatial scales, the intervals are infinitely large at biomass values larger than 80tons/ha. The width of prediction intervals for either of the sites or allometries is large enough that a numerical inversion leads to infinitely large biomass-prediction intervals over all biomass values.

A first glance at the results displayed in Figure 4.11 would suggest that subplot-scale biomass estimates from radar backscatter are more accurate, however such a conclusion seems counter intuitive. Averaging should reduce error in biomass estimates, which does not seem to be the case. The width of the confidence intervals seems only to be guided by the number of samples. The reason for this discrepancy is the exclusion of measurement error from this discussion. The covariance matrix, $\hat{\Sigma}_\beta$ assumes

no difference between the errors in subplot or hectare scales measurements of either biomass or backscatter. Any realistic confidence or prediction interval must therefore include a treatment of measurement error.

4.3.2 Accounting for measurement error

With a detailed discussion of error in measuring field biomass already presented in chapter 2, this section is devoted to estimation of the error in radar backscatter using the theoretical framework presented in Section 4.2 and [56]. Other than temporal variability, the components of backscatter error can be calculated from information about the instrument and knowledge of observation geometry. For UAVSAR data, most of these variables can be obtained from annotation files distributed with the data, and are listed in Tables 4.3 and 4.4. Table 4.3 lists the parameters with which UAVSAR data was acquired and processed over the Harvard and Howland Forest sites during the two field campaigns, while Table 4.4 lists parameters specific to the field sites and the scenes chosen for this analysis. Since the DEM distributed along with UAVSAR data is derived from SRTM, the DEM related parameters such as posting and accuracy are from [34]. The two parameters in Table 4.3 not readily available from the annotation files, AMB_t (total ambiguity ratio) and $\Delta\sigma_{dB}$ (random calibration errors) are based on values in [56].

4.3.2.1 Analyzing temporal variability

A definitive model that describes temporal variation in radar backscatter does not yet exist, mainly due to its complex nature. However, it could be a significant error source, so it must be characterized. It is possible to do so for a specific data set, such as that over the Harvard and Howland Forests, where data collection happened over short enough time scales so that factors such as seasonal and anthropogenic change can be ignored. Physical factors that are expected to still impact backscatter include weather events such as rain or wind, moisture variations in the soil or the

Table 4.3: UAVSAR hardware and processing specific variables needed for computing error in backscatter measurements.

Parameter	Value	Description [units]
B	80	Radar Bandwidth [MHz]
$\bar{\sigma}_0$	-40	Noise equivalent sigma nought [dB]
b	8	Quantization level [bits]
η_ρ	0.7	Range impulse weighting factor
η_s	0.7	Azimuth impulse weighting factor
L_{az}	1.5	Antenna azimuth length [m]
k	1/.88	beam shaping factor
$\theta_{1/2,s}$	8	Azimuth 3-dB beamwidth [degrees]
$\theta_{1/2,e}$	39	Elevation 3-dB beamwidth [degrees]
AMB_t	20	Total Ambiguities [dB] [56]
$\Delta\sigma_{dB}$	0.05	Random calibration errors [dB] [56]
$\Delta\theta$	0.8	Pointing uncertainty [dB]
Δh_{dem}	3	DEM accuracy [m]
L_{dem}	30	DEM posting [m]
$\Delta\rho$	1.875	Range resolution [m], Equation (4.23)
Δs	0.75	Azimuthal resolution [m], Equation (4.24)
$\kappa_{\rho,s}$	1.1909	Beam broadening factor, Equation (4.25)
A_{pix}	1.4062	SLC pixel area ($\Delta s \times \Delta\rho$) [m^2]
A_{dem}	900	DEM pixel area (L_{dem}^2) [m^2]

Table 4.4: Observation specific variables needed for computing error in backscatter measurements.

Parameter	Value	Description [units]
N_{oi}	10	number of speckle identical observations
N_{os}	1	number of speckle diverse observations
A_g	10000	Area of field site (hectares) [m^2]
θ_i	40	local incidence angle (mean) [degrees]
τ_ρ	3	Mean cross track slope [degrees]
τ_s	3	mean azimuth slope [degrees]
N	3223	number of spatial looks, Equation (4.22)
ρ	16322	range to target [m]
$\Delta\theta_{dem}$	0.183	Pointing error form DEM [mm], Equation (4.35)
$\Delta\tau_x$	0.0997	Error in range, azimuth slopes [m], Equation (4.38)

canopy during the day. Over the Harvard forest, UAVSAR acquired data over five days in August of 2009, the 6th, 8th, 13th, 16th and 17th. Figure 4.12 plots average cross-polarized backscatter data over two regions of evergreens and deciduous trees (roughly 3 hectares each) near the Harvard Forest from twenty ascending lines (5° heading). The two regions, chosen from optical imagery in *leaf-off* conditions, consist of roughly 8000 UAVSAR looks, enough to significantly reduce speckle while allowing the regions to remain homogeneous. For all of the acquisition dates except the 13th, there is a general trend of increase in radar backscatter with time for both types of forests. Incidence angle variation could be the reason since these lines were acquired with very varying baselines for repeat-pass interferometric observations. However, not only are the incidence angle variations insignificant enough to cause this trend, the lines were designed such that the incidence angle would increase with each line (hence time) which should theoretically cause a decrease in backscatter intensity [125]. The absence of trend on the 13th could be attributed to the fact that there was significant

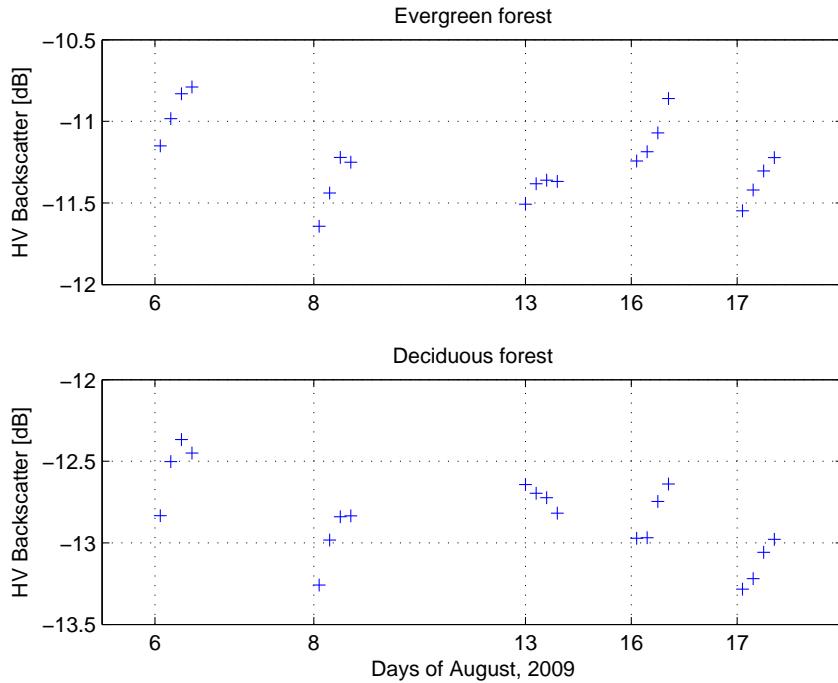


Figure 4.12: Cross-polarized backscatter UAVSAR data from twenty ascending lines over two regions of mostly evergreens and deciduous trees near the Harvard Forest.

precipitation for most of that day with rain as high as 13dBz at the time of acquisition.

Change in moisture of either the soil, trunk or the canopy seems to be the most likely reason for the backscatter trend observed over the other days.

Since the look geometry doesn't vary significantly for these scenes, the contribution to the variation seen in Figure 4.12 from pointing errors is negligible. Furthermore, with a large number of looks, the contribution of speckle and thermal noise is also minimal. In the absence of a well-defined model of temporal variation, once the errors due to DEM uncertainties have been removed the sample variance seen over these twenty lines serves as an estimate of error induced from temporal variations in cross-polarized UAVSAR backscatter data from this particular campaign. Here, this error term is approximated by

$$\Delta\sigma_t = 0.002 \text{ } [m^2/m^2]. \quad (4.44)$$

With this approximation, the error in radar backscatter measurements can be com-

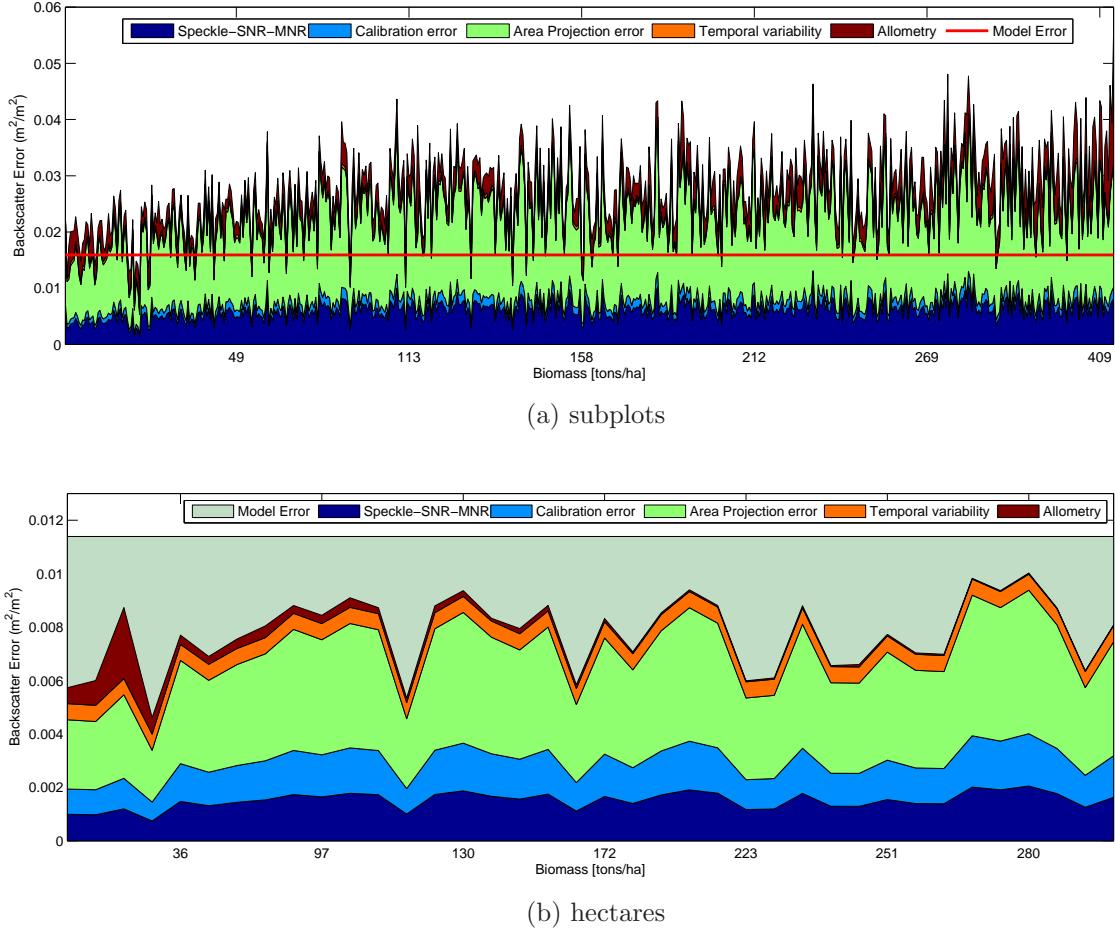


Figure 4.13: Subplot and hectare scale errors in backscatter for the combined dataset over Harvard and Howland Forests. Also plotted, are model errors and allometric errors (in field biomass using the Single-site [118] equations) projected into units of radar backscatter using fitted coefficients.

puted for the particular viewing geometry and field sites over the Harvard and Howland forests. Figure 4.13 plots the estimated total error in backscatter and its components, *Speckle-SNR-MNR* ($\Delta\sigma_s$), *Calibration error* ($\Delta\sigma_c$), *Area projection error* ($\Delta\sigma_a$) and *temporal variability* ($\Delta\sigma_t$) for the field sites at both the hectare and subplot scales. Also plotted are the allometric errors, converted to backscatter units using

$$\Delta\sigma_{mi} = \frac{\partial\sigma}{\partial M} \Big|_{M_i} \Delta M_i \quad i = 1, \dots, n \quad (4.45)$$

where σ is the backscatter modeled as a function of biomass given in (4.1) and M is field biomass. The number of measurements, n , vary with spatial scales. At hectares, $n = 38$ for a total of 38 hectares from the combined dataset over the Harvard and Howland Forests, while $n = 608$ for subplot-scales (with 16 subplots per hectare). Figure 4.13 also plots the model error, estimated from the RMS value of the residuals from the naive fits (where measurement error is not considered). The absolute error in radar backscatter at subplot scales is almost an order of magnitude higher, primarily due to the decrease in the number of looks by a factor of 16. However, note that the backscatter measurement error at subplot scales is almost always larger than the estimated model error, while model error dominates at hectares scales. Furthermore, the backscatter measurement error at subplot scales is comparable to the dynamic range of the backscatter values over the field sites.

As discussed in Section 4.1.2, there are two aspects to obtaining measurement error corrected confidence bounds; estimating bias in regression coefficients and bootstrapping estimates of confidence intervals using the unbiased coefficients. Figure 4.14 shows results of the SIMEX analysis conducted on the coefficients of the backscatter-biomass model in (4.1) using the modeled measurement errors in radar backscatter and field biomass for hectare scale plots. The SIMEX method estimates unbiased coefficients by evaluating the quadratic fits to $\hat{\beta}_i(\lambda)$ at $\lambda = -1$. Figure 4.14 shows the trend in each of the regression variables as a function of λ and their respective quadratic fits that are extended to $\lambda = -1$. Predicted backscatter values from both the biased and unbiased set of coefficients, shown in the bottom-right panel of Figure 4.14 indicate that measurement error does not introduce a significant bias to the predicted backscatter. At subplot scales, with measurement errors in radar backscatter that are comparable to the dynamic range of cross-polarized data over the two field sites, the SIMEX starts to falter because of the failure of iterative fitting algorithms to

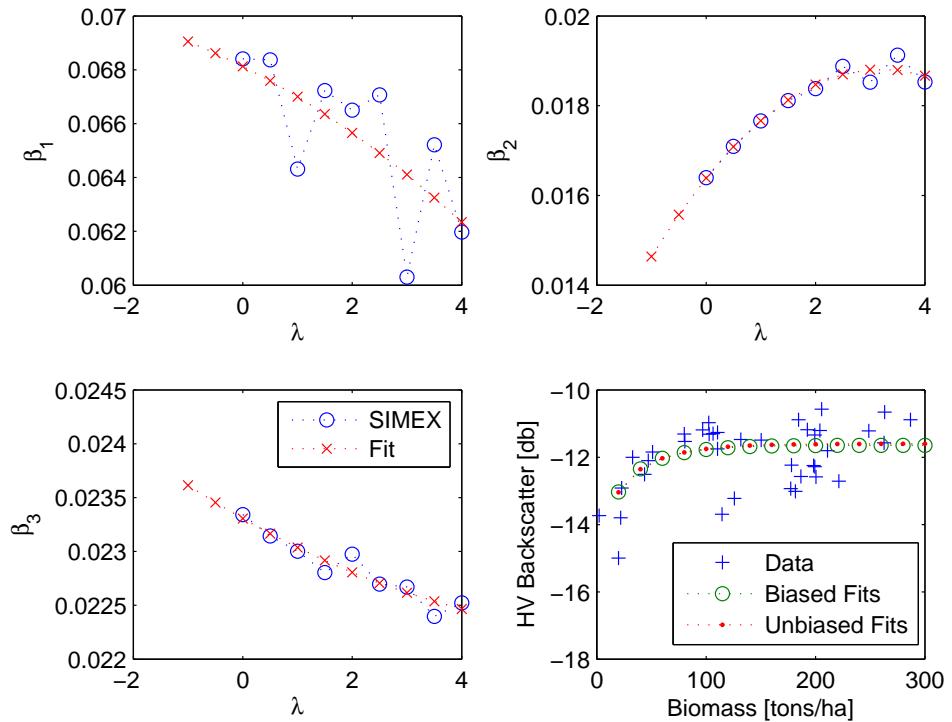


Figure 4.14: Estimation of bias introduced by measurement error on the fit coefficients using SIMEX. Fits using biased and unbiased coefficients are plotted as well.

converge frequently. With enough such failures, SIMEX results become untrustworthy, which is the case with subplot-scale data over the Harvard and Howland Forests. Figure 4.15 shows measurement-error corrected confidence intervals estimated using

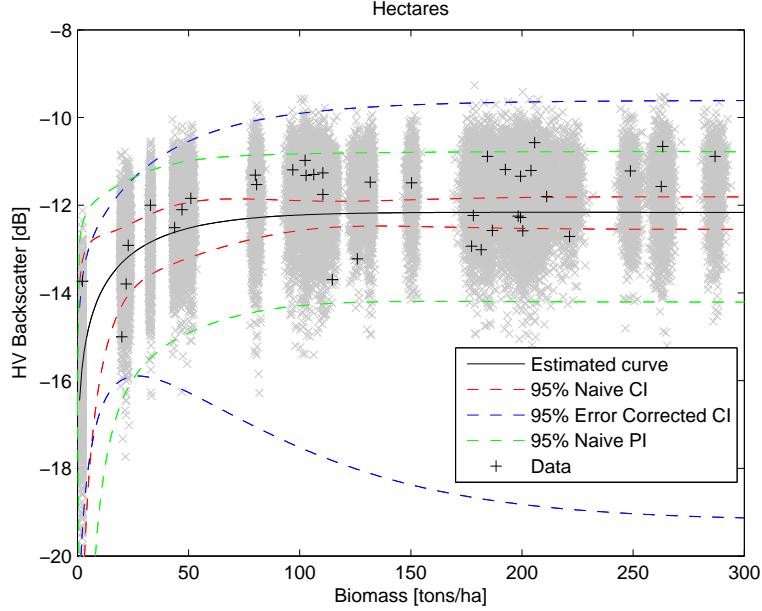


Figure 4.15: Cross-pol backscatter from UAVSAR plotted against field biomass estimates from the BLUE allometry [71] for data combined from the two sites. Confidence intervals from the naive fits as well as the ones corrected for modeled measurement error are shown.

the parametric bootstrap method for hectare-scale data over the two sites. The grey cloud of points are the simulated datasets (for $B = 1000$). The width of the confidence interval estimated from the bootstrap method is much larger than the width of the naive confidence interval (where measurement error is not considered). In fact the confidence interval is wider than the naive prediction interval as well, indicating that the impact of measurement error is fairly significant. The confidence intervals from the two other allometries are similarly wide. These bounds are large enough such that they cannot be inverted to obtain confidence intervals on biomass estimates from the radar. This would suggest that if the bootstrapped intervals are to be trusted, the combination of large model error and measurement error in radar backscatter reduce

the confidence in biomass estimates from radar backscatter to the point that such estimates over the North-Eastern United States would be of very low quality.

4.3.3 Analyzing model error

Figure 4.13 shows that at hectare scales model error is larger than the total measurement error, and is consequently the dominant contributor to the width of the measurement-error corrected confidence intervals. Unlike the species specific regres-

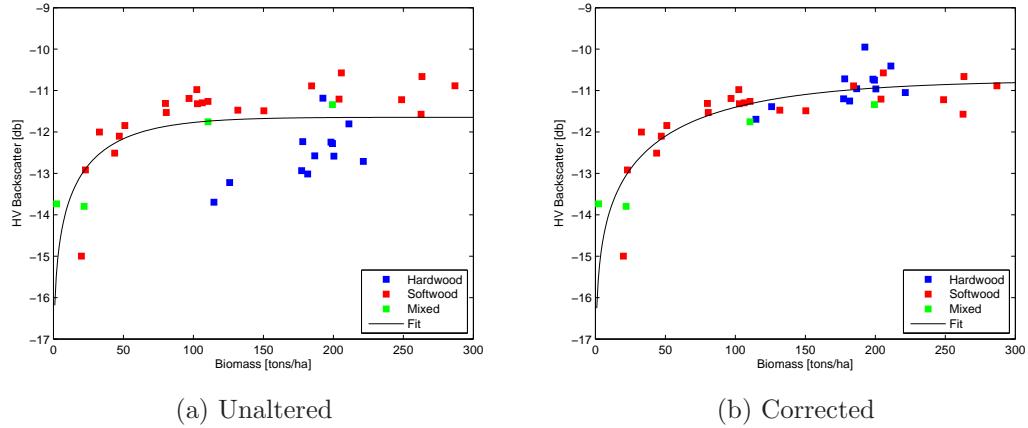


Figure 4.16: Impact of species on the relationship between backscatter and field biomass over the Harvard and Howland Forests.

sion equations used in Dobson et. al. [27] the backscatter-biomass model in (4.1) makes no such distinction. The UAVSAR backscatter data over the Harvard and Howland forests, however, is distinctly different between areas dominated by deciduous trees and areas of mostly coniferous trees. Areas of deciduous trees appear darker than the coniferous trees in the cross-polarized UAVSAR imagery, most likely due to the amount of attenuation caused by the larger moisture content in the deciduous tree canopies, by as much as one dB.

Figure 4.16 shows the difference in backscatter between the two forest types for the hectare scale field sites. The hectare plots are classified as softwoods or hardwoods based on the biomass contribution of each type of tree to the total biomass of the

hectare, with a majority defined as 60% or more. Field sites where the biomass contribution of either forest type is not more than 60% of the total biomass are classified as mixed forest. The grouping of hardwoods and softwoods is quite apparent in Figure

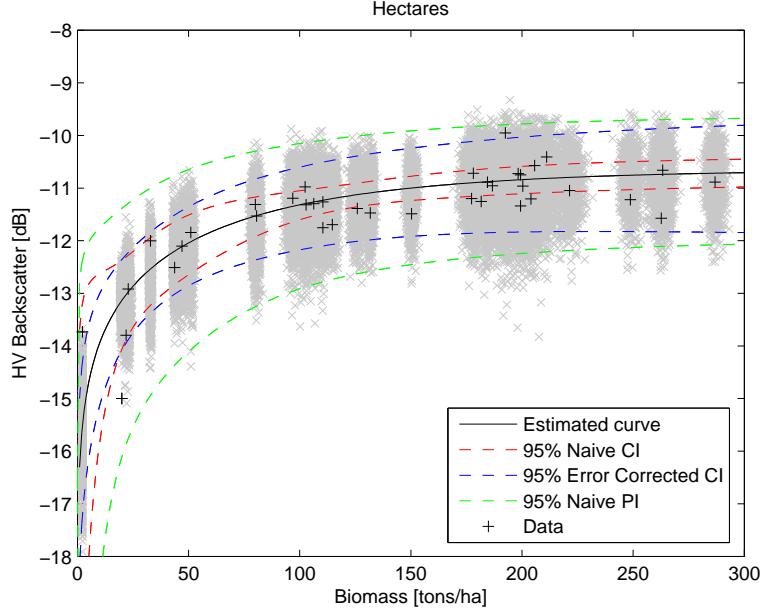


Figure 4.17: Measurement error corrected confidence intervals using the BLUE [71] allometry.

4.16a, with the hardwood dominated sites lower in backscatter value than softwood sites with similar biomass by a dB or so. Figure 4.16b plots UAVSAR data where the backscatter from hardwoods has been offset by a backscatter coefficient value of $0.025[m^2/m^2]$. This particular bias value has been chosen by looking at the average offset between regions of hardwoods and softwoods. This offset isn't guided by scattering models, rather it was chosen to minimize the impact of species dependence of backscatter on model error. Because of this correction, the RMS residuals reduce significantly, from a value of $0.0113[m^2/m^2]$ to $0.007[m^2/m^2]$ a factor of two reduction. Figure 4.17 plots the measurement error corrected intervals at hectare scales using the BLUE [71] allometry with the backscatter intensity values of hardwood dominated hectares offset by $0.025[m^2/m^2]$. The error corrected intervals shown here

are much narrower compared to those seen in Figure 4.15, highlighting the impact of model error on the confidence intervals. Figure 4.18 plots the backscatter-biomass

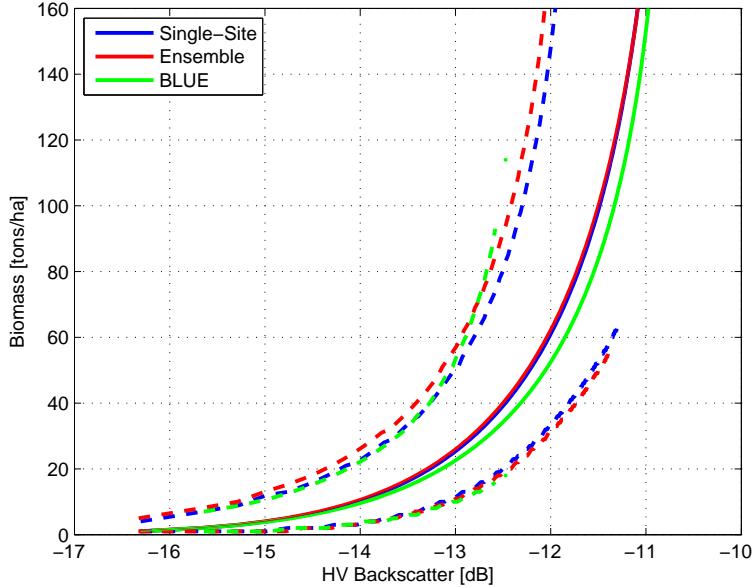


Figure 4.18: Measurement error corrected confidence intervals on biomass estimates from radar backscatter for the three allometries; single-site [118], ensemble [65] and BLUE [71] allometric equations.

model and numerically inverted confidence bounds for the three allometric equations using the corrected data. The asymmetric confidence intervals are not invertible after 60 tons/ha, primarily due to the saturation in radar backscatter at higher biomass values. However, even where these confidence intervals are invertible they are quite wide, suggesting that even though backscatter can predict low biomass, the quality of these estimates would still be low. Figure 4.19 shows the width of the 95% confidence interval on biomass estimates as a function of the radar derived mean biomass. The width of the confidence interval is always larger than the mean value of the biomass estimate itself, highlighting further the low quality of radar derived backscatter estimates over the Harvard and Howland Forests even at low biomass levels where radar

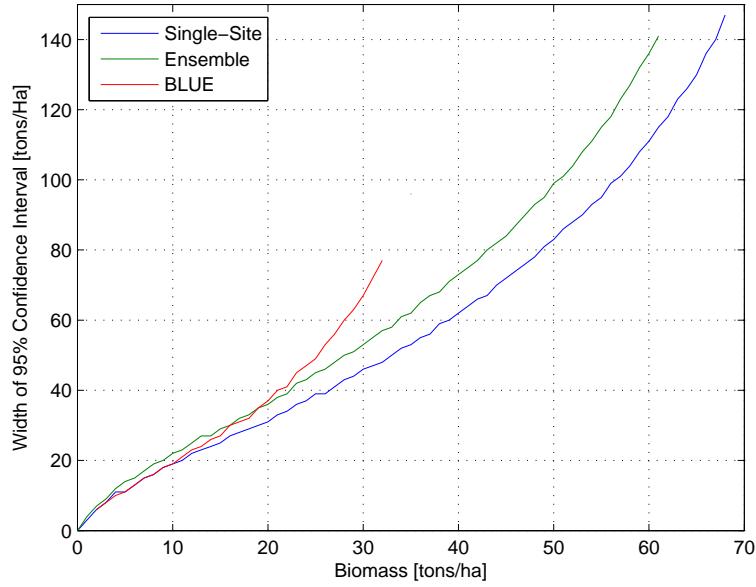


Figure 4.19: Width of the confidence intervals on biomass estimates obtained from radar backscatter using the three allometries; single-site [118], ensemble [65] and BLUE [71] allometric equations.

backscatter is expected to perform best.

In this chapter, a statistical framework for estimating the quality of biomass estimates obtained from radar backscatter incorporating measurement errors in both field biomass and radar backscatter was established. Techniques for calculating confidence intervals in a non-linear regression scenario with measurement error, such as the case when estimating biomass from radar backscatter, were discussed. A forward model that predicts radar backscatter from biomass was chosen for the analysis presented here and a model for estimating the error in radar measurements was outlined. Calculation of the errors in radar measurements from the UAVSAR field campaigns over the Harvard and Howland Forests based on the error model and instrument information were discussed. Results of the regression analysis between field estimates of biomass, outlined in Chapter 2 and backscatter data from the UAVSAR

field campaign over the Harvard and Howland Forests shows that measurement error corrected confidence intervals from the non-linear backscatter-biomass model are wide enough that they cannot be inverted to obtain the corresponding intervals on biomass estimates from radar backscatter at either the hectare or the subplot scales. The reasons for the large confidence interval widths are two fold, a) the saturation of the backscatter-biomass curve and b) combination of large model error and measurement error in radar backscatter. An analysis of the model error shows that a distinct bias in backscatter value between the deciduous and coniferous trees contributes to an increase in model error variance. Removing the bias in backscatter between the two forest types reduces model error and consequently allows for estimation of confidence bounds on biomass estimates. However, the confidence intervals can only be estimated at hectare scales, at subplot scales measurement error in radar backscatter tends to dominate and cause the bootstrap method to fail. The width of confidence bounds on biomass estimates at hectare scales over the Harvard and Howland forests show cross-polarized radar backscatter to be a poor estimator of forest biomass, even at low biomass levels.

CHAPTER 5

RADAR INTERFEROMETRY

In principle, a radar interferometer can obtain very accurate height estimates by measuring the path length difference of the scattered electric field received by two antennas separated by some distance, called a baseline, and relating it to scatterer height [79, 107, 108] through a simple geometric transformation. The sensitivity of the interferometer to height increases with increasing baseline lengths. Large baselines on a single platform (or on two platforms, as in a tandem mission), however, increases system cost and complexity considerably. Repeat orbit interferometers on the other hand rely on the proximity of two satellite overpasses to synthesize a baseline. Such a system utilizes a single antenna, hence a small platform, and is therefore less expensive and more realizable. In this case the interferometric pair is formed by a repeat-pass observation of the satellite, the baseline is formed by a slight change ($\approx 1\text{km}$) in the satellite orbit. If the baseline of an InSAR system is zero (or nearly so) the instrument can detect very minute changes of the target from one pass of the satellite to another. This allows for studies of physical processes such as ice dynamics and earth deformation. Large baselines, however, are more sensitive to topographic variations. The topographic sensitivity of a non-zero baseline interferometer can be used to invert for tree heights if knowledge of the true ground surface and canopy penetration characteristics are available [68, 113]. Interferometric correlation magnitude has also been shown to contain information of tree heights and stem volumes [52], [4], [111]. An interferometer then, in principle, could generate global biomass maps through accurate measurements of forest height, in essence combining the map-

ping capabilities of radar and the accurate vertical measurements of a lidar. Any

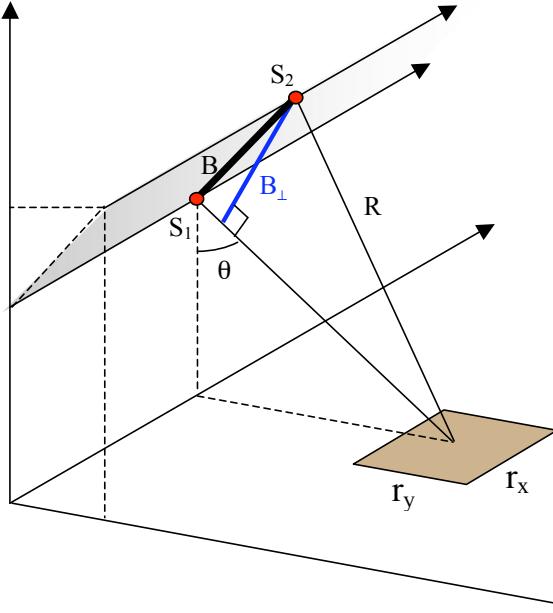


Figure 5.1: Typical configuration for a repeat orbit interferometric SAR (InSAR). S_1 and S_2 represent position of the radar for the two passes separated by the baseline, B . The InSAR maps pixels of resolution r_x , r_y in range and azimuth from a look angle of θ . Difference in phase of an electric field scattered from the pixel at the two antennas is used to derive height estimates.

repeat pass InSAR observation is susceptible to changes in the scene during the two acquisitions [96, 111]. This particular loss of coherence, or temporal decorrelation, is an important contributor to the uncertainty in forest heights and structure estimates from InSAR measurements. In the following chapter, an extensive analysis of this form of error source for repeat pass interferometry is described. This is done by first clearly defining the quantity and then demonstrating how it may be calculated from interferometric observations. We then perform this analysis over an extensive geographic region. The L-Band data-set, shown in Figure 5.4, collected by the SIR-C shuttle mission which flew over the eastern US in October 1994 is analyzed in conjunction with the National Land Classification Dataset (NLCD 1992). It is shown

that weather, wind, and seasons all play a role of varying degrees on temporal decorrelation.

5.1 Interferometric Correlation

The interferometric correlation is defined as

$$\gamma = \frac{\langle S_1 S_2^* \rangle}{\sqrt{\langle |S_1|^2 \rangle \langle |S_2|^2 \rangle}} \quad (5.1)$$

where γ is interferometric correlation, S_1 and S_2 are electric fields received by the two antennas as shown in Figure 5.1. In the case of a repeat orbit interferometer, S_1 would be the scattered field received at the first and S_2 field received at the second pass of the instrument. This observed correlation can be broken down to its components as [138]

$$\gamma_{obs} = \gamma_{geom} \cdot \gamma_{thermal} \cdot \gamma_{vol} \cdot \gamma_{temp} \quad (5.2)$$

where γ_{geom} reflects slight changes of the radar viewing geometry from both ends of the interferometer, while $\gamma_{thermal}$ and γ_{vol} are the contributions of system thermal noise and volumetric scattering respectively and γ_{temp} represents temporal decorrelation. Among these components, because of its stochastic and non-stationary nature, temporal decorrelation is the hardest to model, isolate and analyze in vegetated areas. Furthermore, to better manage tradeoffs and resources for a repeat pass InSAR design, it is important to understand this effect to better quantify this potentially dominant error source, especially as it applies to estimating tree heights and vegetation structure on global scales.

5.1.1 Components of Interferometric Correlation

In order to analyze the effect of temporal decorrelation, other components that contribute to overall observed correlation must either be absent or corrected for.

The contributions from $\gamma_{thermal}$, γ_{geom} and γ_{vol} can all be mathematically modeled [138, 79, 107]. A discussion of these effects follows.

Additive thermal noise in interferometric data reduces coherence. This is referred to as $\gamma_{thermal}$. Assuming that additive noise is incoherent with the received signal and different in both interferometric channels, it can be shown that thermal effects can be modeled as a function of the signal to noise ratios (SNR)

$$\gamma_{thermal} = \frac{1}{\sqrt{1 + \text{SNR}_1^{-1}} \sqrt{1 + \text{SNR}_2^{-1}}} \quad (5.3)$$

where SNR_1 and SNR_2 are the signal to noise ratios for the two channels. The observed correlation can be corrected for thermal effects by using $\gamma_{thermal}$ estimates obtained using (5.3) as simply

$$\gamma_{gvt} = \frac{\gamma_{obs}}{\gamma_{thermal}} \quad (5.4)$$

where γ_{gvt} is the combined effect of volumetric, spatial and temporal effects. Respective SNR estimates required in (5.3) are obtained, in this case, from the intensity images of each pass.

Geometric decorrelation, $1 - \gamma_{geom}$, sometimes referred to as baseline decorrelation, is reflective of the loss of coherence in an interferogram due to slight changes in the viewing geometry. It is intuitive that two radar returns will not be fully correlated if a scatterer is viewed from two different angles. This change of viewing angles is proportional to the projected interferometric baseline. Geometric decorrelation is broken down further into spatial, $\gamma_{spatial}$, and rotational, γ_{rot} effects, where the former is a function of the across-track component, while the latter is a function of the along-track component of the interferometric baseline. Because the orbits of spaceborne sensors are essentially parallel, when InSAR observations are processed to a common Doppler frequency, rotational effects are essentially zero (i.e. $\gamma_{rot} = 1$). Hence the geometric correlation is given by

$$\gamma_{geom} = 1 - \frac{2B_\perp r_y \cos \theta}{\lambda R} \quad (5.5)$$

where B_\perp is the perpendicular baseline in the look direction, θ is the look angle, r_y is the range resolution of the radar, λ is the wavelength and R is range to the target.

As one can see in (5.5) the geometric correlation coefficient tends to unity when the perpendicular baseline nears zero. Conversely, complete decorrelation occurs at critical baselines [138] as in

$$B_{\perp,crit} = \frac{\lambda R}{2r_y \cos \theta}. \quad (5.6)$$

The volumetric decorrelation in interferometric data is reflective of scattering of radar signals from multiple heights within each resolution element. The observed correlation signature can be modeled as the Fourier transform of the radar backscatter volume as a function of height [120, 121]

$$\gamma_{vol} = \frac{\int \sigma(z) \exp^{-j\kappa_z z} dz}{\int \sigma(z) dz} \quad (5.7)$$

where $\sigma(z)$ is the effective radar backscatter cross section per unit height, z , and the vertical wavenumber, κ_z , is given by

$$\kappa_z = \frac{4\pi B_\perp}{\lambda R \sin \theta}. \quad (5.8)$$

The shape of $\sigma(z)$ is dependent on the target. In case of a surface scatterer, such as a flat field, this can be approximated as an impulse function, $\delta(z - z_0)$ with z_0 being the height of the ground. In this case the Fourier integral collapses and no volumetric decorrelation for any κ_z is observed, as expected. On the other hand, a volumetric scattering target with uniform backscatter as a function of height, say a forest that extends up to a scattering height of h_v , would have nearly a maximum amount of observable decorrelation, thereby reflecting a “worst case” contribution of

volume scattering to the overall decorrelation. The corresponding Fourier transform gives the volumetric correlation from such a scattering model as

$$\gamma_{vol} = \frac{2\sin(\kappa_z h_v/2)}{\kappa_z h_v}. \quad (5.9)$$

This shows that it is possible to *estimate* the amount of volumetric decorrelation caused by trees of height h_v observed by a system with parameters κ_z , under the assumptions of uniform effective radar backscatter cross section, spatial homogeneity in a resolution element and no surface return. In general, this would be an upper limit to the observed correlation, even when surface and non-uniform scattering effects are taken into account. Conversely, it may be possible to estimate tree heights if the observed decorrelation is composed entirely of volumetric effects. This would be hindered by the presence of observational errors (such as temporal decorrelation and thermal noise effects) as well as the unquantified bias introduced by the physical models.

5.2 Interferometric Heights and Biomass

Lidar derived heights were shown to be related to forest biomass over the Harvard and Howland Forests in Chapter 3. Lidar-biomass relationships were stronger than those between backscatter and biomass seen in Chapter 4. Since an interferometer also measures some aspect of forest height, the relationship between interferometric heights and forest biomass would also be of interest, and one that has been studied previously [113, 68]. We explore this briefly by analyzing the relationship between SRTM [34] derived heights and field estimates of biomass over the Harvard and Howland Forests. Over forested regions, SRTM elevations that are referenced to an ellipsoid, are offset from the ‘bare-ground’ elevation due to a phase-center located somewhere within the vertical extent of the scattering forest. Figure 5.2 illustrates

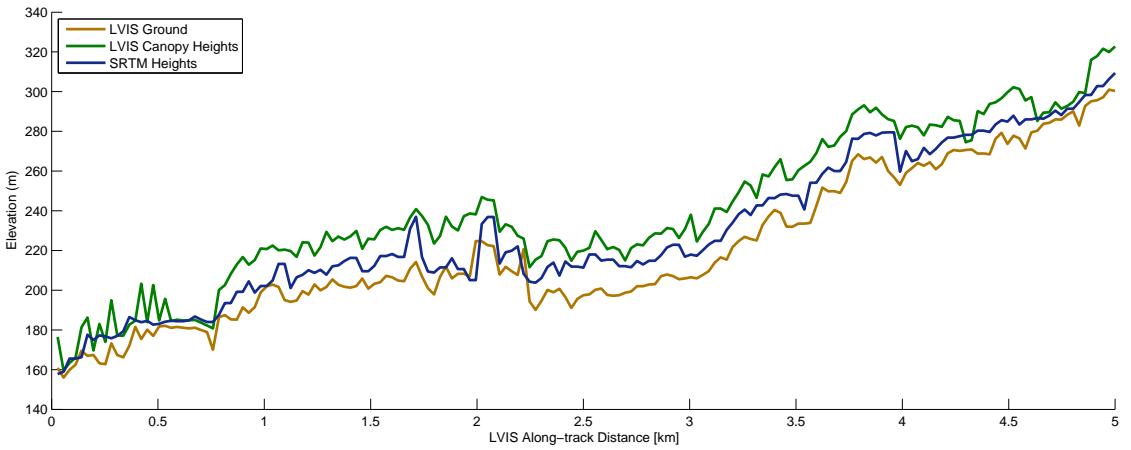


Figure 5.2: SRTM derived topography and LVIS heights. SRTM elevations are plotted with LVIS detected ground and canopy top using only nadir looking shots.

the location of SRTM phase center comparing SRTM elevations to LVIS derived canopy top (RH100) and ground elevation, all relative to the WGS84 ellipsoid. In some cases, such as bare surfaces (seen towards the beginning of the transect) both the SRTM and LVIS metrics indicate the same elevations, as expected. However, in forested areas (most of the transect) the SRTM heights are located somewhere within the vertical extent of the canopy, bounded by the LVIS ground at the bottom and RH100 at the top. The variation of SRTM elevations in the forested regions shows that the information contained in the interferometric heights is different from that in lidar derived heights, likely due to the varied impact of forest structure on lidar and radar scattering. Figure 5.3 plots the SRTM derived forest heights, obtained by subtracting the LVIS ground elevation from the SRTM elevation, against field biomass estimates over the Harvard and Howland forests. SRTM derived heights are generally higher for the Harvard Forest sites, similar to what was observed in the lidar data in Chapter 3. There seems to be a clear relationship between SRTM heights and field biomass. Table 5.1 summarizes the regression statistics from linear fits between SRTM heights and field biomass in the form of R^2 and RMSE (in units of tons/ha)

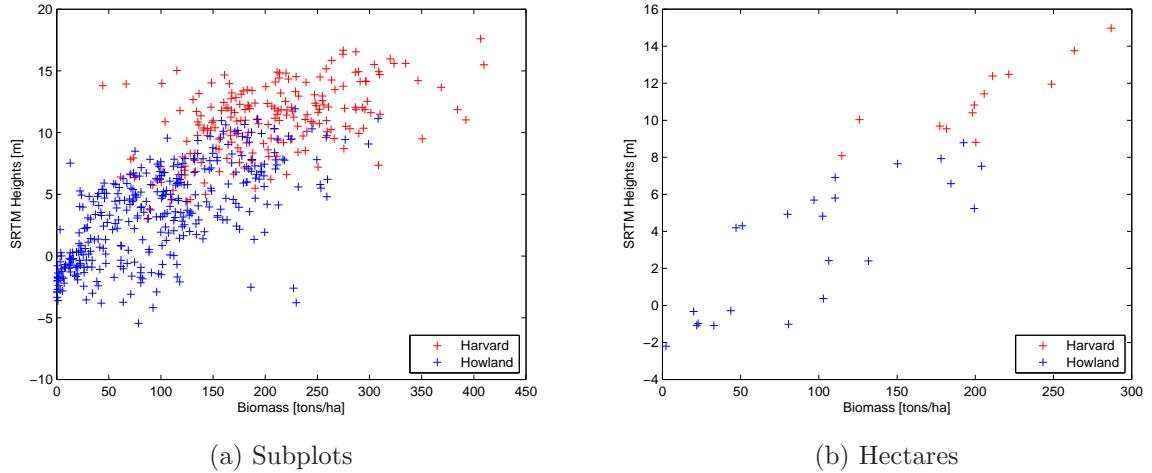


Figure 5.3: Heights derived from SRTM DEM and LVIS ground plotted against biomass from Harvard and Howland Forests at the subplot and hectare scales.

Site	Subplots		Hectares	
	R ²	RMSE	R ²	RMSE
Harvard Forest	0.27	60.03	0.72	25.66
Howland Forest	0.42	51.37	0.63	38.36
Combined	0.58	54.83	0.79	34.55

Table 5.1: Regression statistics for linear fits between SRTM derived heights and biomass data from the Harvard and Howland Forest sites. The RMS errors are in units of tons/ha.

for the Harvard and Howland Forest sites individually and combined at the subplot and hectare scales. For this analysis, field biomass was estimated using the BLUE allometric equations, described in Section 2.1.2.3. The correlations between SRTM heights and field biomass are particularly strong at the hectare scales, with R^2 values of up to 0.79 and RMSE of 34 tons/ha, as good or even better than the relationship between lidar metrics and field biomass for these sites.

The fit statistics summarized in Table 5.1 do not include measurement errors in either SRTM heights or field biomass. SRTM, however, was a single-pass instrument operated at C-band, an accuracy analysis that includes SRTM measurement errors would not entirely be applicable for a repeat-pass L-band mission scenario. In fact, of critical importance in the parametrization of errors for such a mission is the characterization of temporal decorrelation. The following sections are devoted to an analysis of temporal decorrelation.

5.3 Observations of Temporal Decorrelation

In this analysis we use SIR-C L-Band repeat pass data collected over the eastern United States on October 9 and October 10, 1994. We analyzed a composite swath comprised of eleven individual scene pairs described by their processing run numbers: pr52102-pr52126 for pass 1 (October 9) and pr52136-pr52160 for pass 2 (October 10). Figure 5.4 shows the individual SIR-C scene boundaries overlaid on a map of this area. The SIR-C L-Band radar collected this fully polarimetric dataset from a nominal altitude of 225km with a spatial resolution of approximately 5m and a look angle of 25 degrees. The perpendicular baseline, B_{\perp} , for this particular interferometric pair varies from 30m to less than a meter. The critical perpendicular baseline, $B_{\perp,crit}$ for this pair is 5km. The maximum baseline of 30m is such a small percentage of the critical baseline that there is a negligible spatial component to the overall interferometric decorrelation. Equation 5.7 shows the highest expected decorrelation

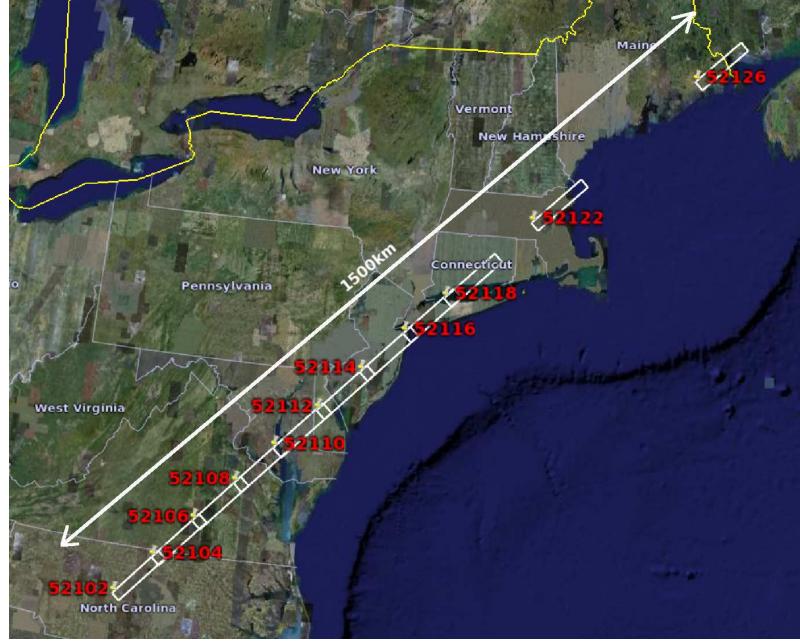


Figure 5.4: The SIR-C swath over eastern US. Each of the 11 scenes comprising this swath is represented by a processing run number (52102 to 52126), and maps the area highlighted by the white polygon.

due to scattering from a volume, such as, trees. In the case of this particular SIR-C data, if we use an average tree height of 50m (about twice the height of a tall tree in the Eastern U.S., clearly an extreme), the maximum volumetric decorrelation would be 0.03, or 3%. Since most trees are not this tall, and the baseline is less than 30m, the effect of volumetric decorrelation for this data set is essentially non-existent. Hence, one can claim that the observed decorrelation is dominated mainly by thermal and temporal effects. Once corrected for $\gamma_{thermal}$, the observed decorrelation is in most part caused by temporal changes, which makes this dataset ideal for an analysis of temporal decorrelation.

To estimate interferometric correlation γ given in (5.1), the sample estimate of correlation was used. Sample correlation is obtained by spatial averaging. For this data analysis we averaged 30 pixels, 3 in range and 10 in azimuth, which amount to an estimated 25 independent looks [49]. These estimates of the interferometric

correlation are not ideal and are biased for low coherence values [119]. This effect is observed and discussed later.

To correct for thermal effects using (5.3) knowledge of noise levels is required. Power levels seen in dark areas of images were compared with published values of the typical noise-equivalent sigma-zero (NESZ) of -36dB for SIR-C [40]. The two were in agreement. A NESZ of -36dB was used throughout this analysis as an estimate of noise levels. Due to natural statistical variations in calibration and noise power estimates, the correction for thermal effects inevitably leads to pixels with coherence larger than unity. The coherence of such pixels was set to 1, leading to the conclusion that no temporal decorrelation was observed for those pixels. The SIR-C data was processed from reformatted signal data products using the SAR processor from GAMMA remote sensing.

Ancillary data needed for this analysis include digital elevation models (DEMs) and land-cover classification data. The Shuttle Radar Topography Mission (SRTM) DEM and the National Elevation Dataset (NED) are available with high spatial resolutions (30m) over the entire US. SRTM data is derived from C-band radar interferometry [34], while NED is based on United States Geological Survey digital topographic data originally compiled from ground survey and aerial imagery [48]. Vertical accuracy of both NED and SRTM is approximately 5-10m. The SRTM DEM was used to remove topography from the interferometric phase of the SIR-C data. Even though the SRTM C-band DEM and SIR-C L-band data may have different scattering phase centers, the gradients observed by both frequencies, however, are expected to be similar, and hence the differences will have little effect on the estimate correlation magnitude, which was the focus of this study.

The National Land Cover Database (NLCD) [2, 128] was used as the primary land-cover data. This U.S. wide dataset was classified and compiled primarily from 30m spatial resolution Landsat Thematic Mapper (TM) imagery from 1991 to 1993. In the

NLCD, land-cover is classified into nine major (Level I) categories and up to 21 minor (Level II) categories. The Level I categories are water, barren, shrubland, natural grassland, wetland, developed (urban), forest, orchard, and cultivated herbaceous (agriculture). Agriculture is further classified into pasture, row crops, small grain crops and fallow and forest is further classified into deciduous, evergreen or mixed forest. A study of the accuracy of the NLCD over the eastern U.S. [115] showed an overall accuracy of 70-80% for Level I data and up to 66% for Level II data.

SRTM, NED and NLCD datasets have been co-registered with the SIR-C data presented here to sub-pixel level accuracies. Meteorological data from about 1500 weather stations around the eastern US was also collected through the National Climate Data Center (NCDC) archives [51]. This data includes hourly precipitation levels, wind speed and direction before and at the time of the SIR-C data takes in October 1994.

5.4 Analysis of Temporal Decorrelation

There are several physical mechanisms that could cause temporal decorrelation in interferometric data. It has been observed [5, 110] that occurrence of weather events such as precipitation and wind during or between the two passes decreases coherence. A slow decrease in coherence over time on the order of weeks and months has also been observed [138]. Another decorrelation mechanism may be human-driven land-cover change, although this is minimal and restricted to small geographic scales for short InSAR repeat times. The amount of temporal decorrelation might also be related to seasons. This includes vegetation phenology (annual changes), for example the lack of leaves on deciduous forests during the fall season, or the presence of a snow layer on the ground. Furthermore, temporal effects also have a dependence on the type of surface. Water, for example, is expected to have a decorrelation scale on the order of seconds.

The predominant NLCD Level I classes in the eastern U.S. swaths are forest, agriculture, and water. Because these may have different signatures for temporal decorrelation for this reason it is essential to separate pixels based on type of terrain in order to create meaningful statistics on large spatial scales. The NLCD land-cover data allows us to separate pixels in SIR-C interferometric data based on land-cover class. There are, however, some potential error sources associated with using the NLCD data that could influence the present analysis. First, there are classification errors of omission and commission, in particular where similar classes were not completely separable by the Landsat TM. These error rates have been documented [115] and show the overall accuracy over the eastern US to be 80% and accuracy for forest and agriculture to be 82% and 74% respectively. Secondly, land-cover changes can occur; however the fact that the NLCD was compiled from Landsat TM imagery taken from early-mid 1990s, the same time frame as the 1994 SIR-C data, reduces the risk of this type of error. Data processing artifacts introduced by resampling and interpolation for data co-registration are assumed to be relatively small given the high spatial resolutions. Therefore we expect the classification accuracy statistics to represent the dataset error for this analysis.

5.4.1 Temporal decorrelation as a function of land-cover

Figures 5.5 and 5.6 show histograms and density functions of observed temporal decorrelation in SIR-C data for four predominant land-cover types in the region and classified by the NLCD at Level I or II: water, agricultural land, evergreen forests and deciduous forests. We chose to use the two Level II forest classes of evergreen and deciduous in order to separate SIR-C pixels that might have signatures of seasonal effects on in interferometric correlation. Because of the potential of greater misclassification in these finer Level II classes, this diversity comes at the cost of reduced

classification accuracy, down to 55% for each subclass from 82% for the Level I Forest class.

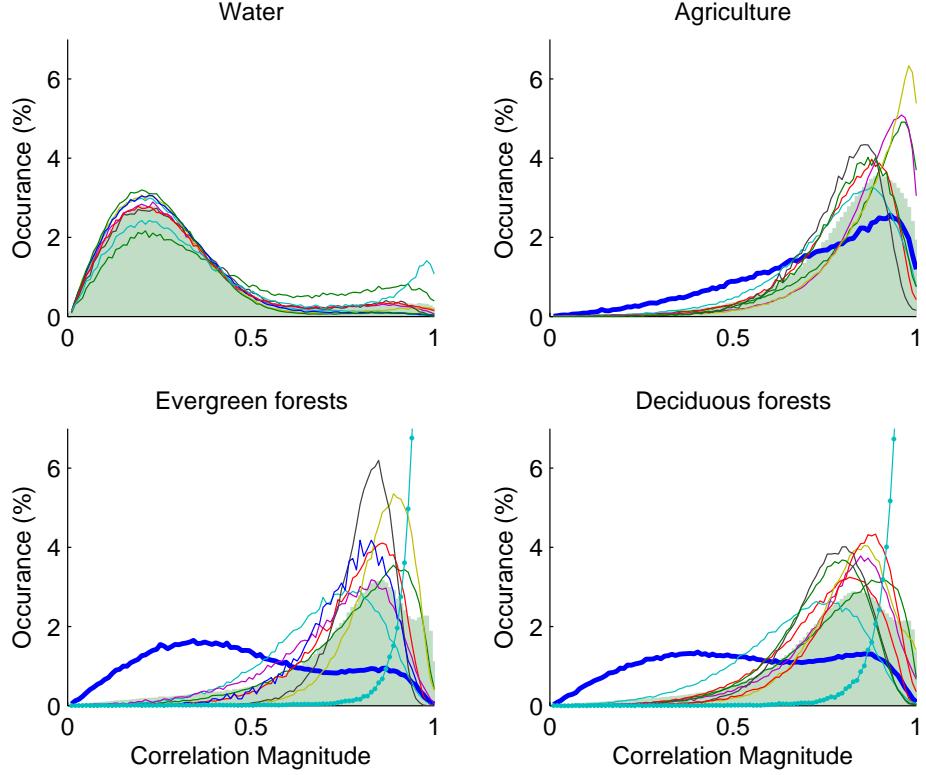


Figure 5.5: Probability density functions of correlation magnitude are shown for four different land cover types, water, agricultural land, evergreen and deciduous forests. Land-cover classification data (NLCD 92) is used to separate pixels into these particular categories. Each solid line represents a single scene, while the shaded region is the average statistics for the entire swath. The bold blue line and the dotted line represent the two scenes 52102 and 52116 that suffer the most and least temporal decorrelation respectively. A high temporal decorrelation for water pixels is evident for all scenes.

Each solid line in Figures 5.5 and 5.6 represents the density or distribution functions from a single scene, while the shaded curves are the respective statistics for each land-cover class for the entire swath. The most apparent signature in these statistics is the low coherence of water pixels. This is expected. In fact coherence in water pixels of an interferogram with a one day temporal baseline should be zero, so the mode of 0.2 then may seem too high. This non-zero value of correlation magnitude is due

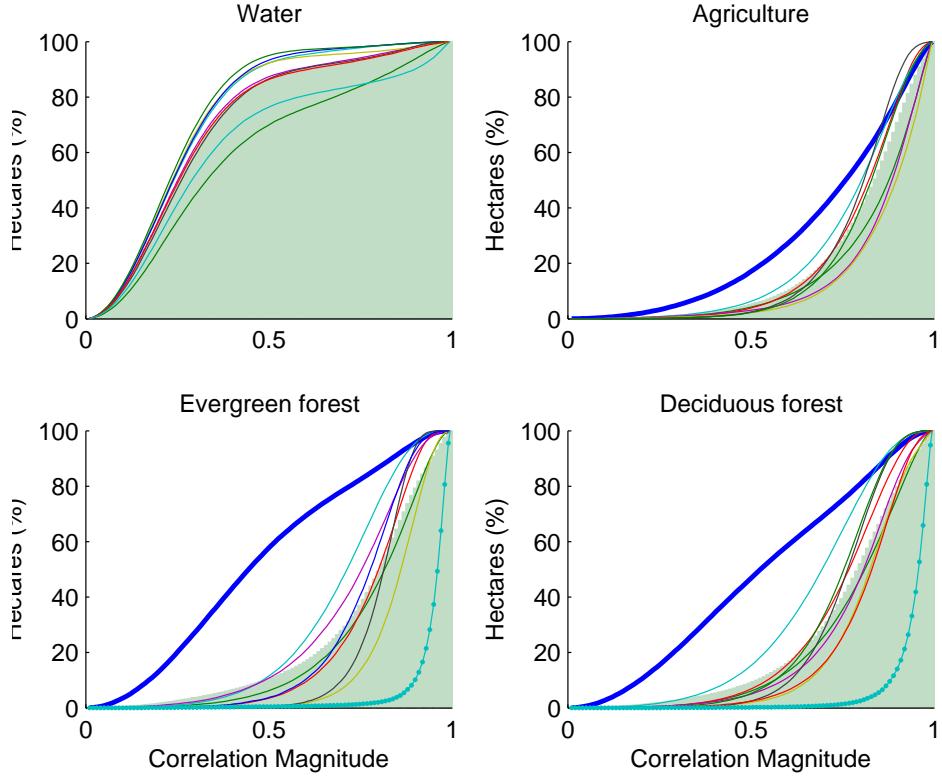


Figure 5.6: Cumulative distribution functions of correlation magnitude derived from the histograms shown in Figure 5.5. Each solid line represents a single scene from the swath while shaded regions are mean statistics. Deviations from mean behavior for agricultural and forested areas are seen in scenes 52102 (highest decorrelation) and 52126 (lowest decorrelation).

to a natural bias in estimating correlation magnitude [119] and in short reflects the inability of the observed correlation magnitude to be less than zero or greater than unity. The amount of bias depends on the number of independent looks averaged (in this case 25). For this number of looks a bias of around 0.19 is expected for a true coherence value of 0.0. The coherence, then corrected for thermal effects is seen to have mode of 0.2 in Figure 5.5, in good agreement with this bias estimate. Another feature seen in the correlation statistics for water pixels is the secondary mode at coherences closer to unity. This is due slight inaccuracies in land classification. Open water pixels are not expected to be misclassified, however pixels near the

water's edge often include coherent scatterers. A large portion of the area mapped in some scenes of this swath includes coastlines. These pixels add high coherence points to the histograms, and explain the anomalies seen in these statistics.

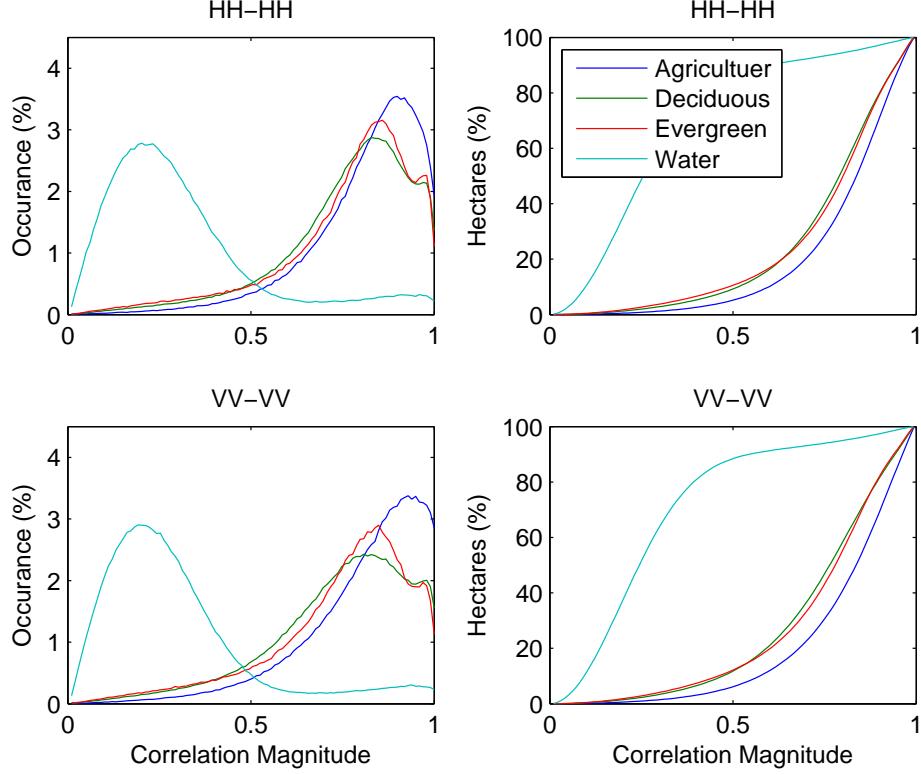


Figure 5.7: Coherence statistics, density (left) and distribution (right) functions of four land-classes, water, agricultural area, deciduous and evergreen forest are plotted as a function of the polarizations HH-HH, and VV-VV. These statistics include pixels from the entire swath. Differences between these distributions are statistically significant.

Although the difference in correlation statistics between agricultural land and forested areas is not as apparent as the difference between statistics of these land-cover classes and water, the differences, however, are statistically significant in all cases. The Mann-Whitney U-test [85] was used to test for significance. This is a classic two sample non-parametric hypothesis test, where the null hypothesis, i.e. that two observations are from the same population and therefore have the same

statistics, is rejected at the 5% significance level (or p-value less than 0.05). The highest p-value calculated for all pairs based on land-cover classification was 0.034, showing statistical significance in each case. The large amount of pixels gathered for each land-cover class is one of the primary reasons for such high confidence in these coherence statistics.

From a mission planning perspective, the distribution functions in Figure 5.6 are an important resource. According to these curves, one can claim for example that 60% of predominantly evergreen forests will have interferometric coherence less than 0.8 compared to only 40% of agricultural land with coherence less than 0.8.

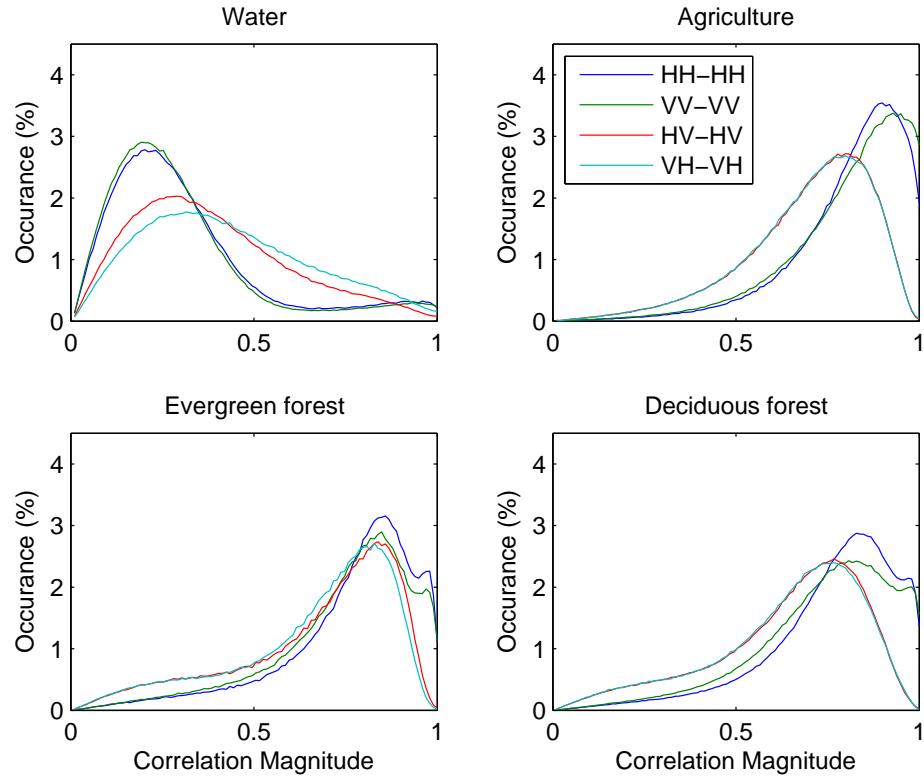


Figure 5.8: Probability density functions of coherence obtained from pixels from the entire SIR-C swath are shown for four different land-classes and polarizations. Agricultural land is better correlated in VV-VV polarization, while forested areas seem to have better coherence for HH-HH polarization.

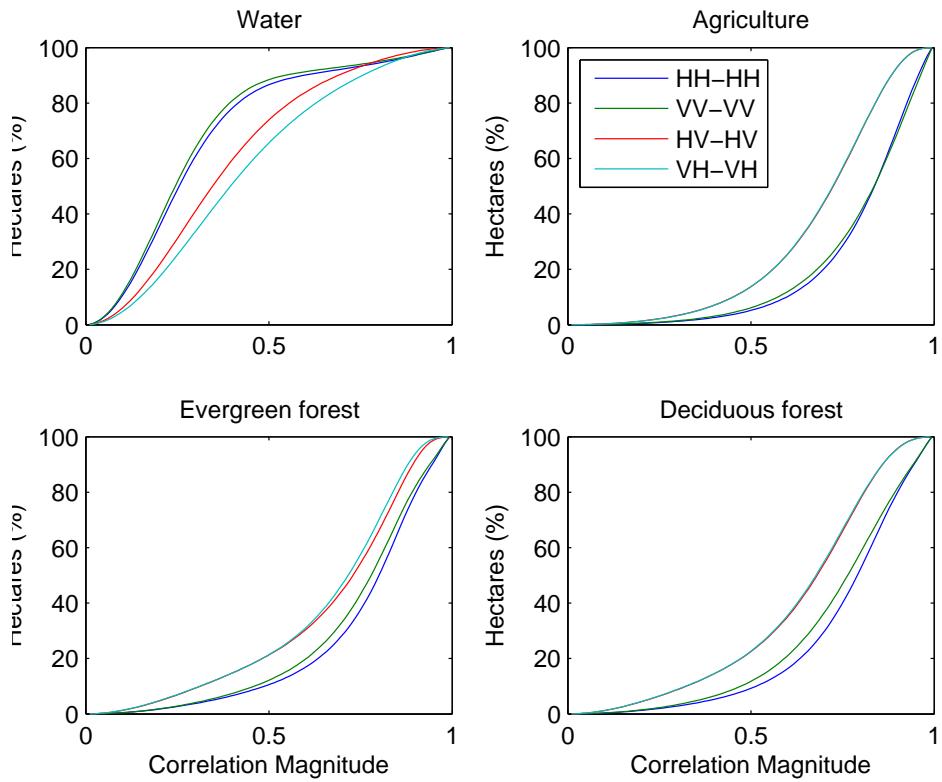


Figure 5.9: Distribution functions of coherence are shown for four land-classes and four polarizations. Difference between HV-HV and VH-VH pixels is statistically insignificant. Abnormally high coherence in water pixels for these polarizations is explained by lack of SNR.

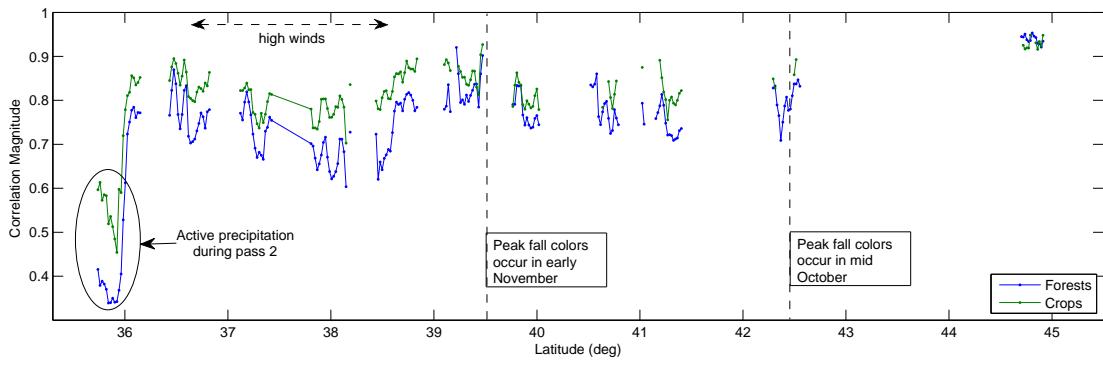


Figure 5.10: Correlation magnitude of agricultural and forested land is plotted as a function of latitude. Each point represents mean coherence of respective pixels from 100 cross-track InSAR lines. Weather and seasonal events such as active precipitation, high winds and fall foliage times are annotated.

The SIR-C data analyzed in this study is fully polarimetric. The results presented in Figures 5.5 and 5.6 are from the HH-HH polarization, i.e. transmission and reception of horizontally polarized waves for both passes. Figure 5.7 shows the coherence statistics of HH-HH and VV-VV polarizations for the four land-cover classes discussed so far. The difference in temporal decorrelation between land-cover classes for both polarizations is quite apparent here. Agricultural land suffers the least amount of temporal decorrelation in either polarization, as evident by the fact that histograms of this land-cover class are more biased towards unity than any other class. As a scalar measure of the amount of temporal decorrelation, one can consider the mode of each density function. The distribution mode for interferometric coherence in agricultural land lies between 0.96-0.97 for the two polarizations while mode of coherence for pixels from forested terrain is between 0.82-0.84.

Figures 5.8 and 5.9 show observed correlation statistics for the four main polarizations. Water pixels in the cross-polarization data seem to have high coherence. This is the effect of particularly low SNR for cross-polarization data compared to co-polarized data over water, and may be suggestive of imaging ambiguities which are normal but typically insignificant in regions with higher reflected power. Vertical polarization data seems to be better correlated than data from horizontal polarization in agricultural areas while the opposite can be seen for either of the forest classifications, an interesting if not slight signature, that may warrant future investigation.

5.4.2 Extent of atmospheric effects on interferometric coherence

The large scale statistics seen in Figures 5.5-5.9 are inclusive of pixels from all the area mapped by this SIR-C swath. The statistics are inclusive of pixels that may or may not have variations resulting from weather events or other phenomenon discussed in Section 5.4.1. Coherence statistics from some individual scenes in Figures 5.5 and 5.6 exhibit deviations from the mean behavior. The density functions of southern

scenes 52102 to 52108 for forests have significant contributions from low correlation pixels. Scene 52102 in particular (marked in Figures 5.5 and 5.6 by the thicker dark blue line) has a near bimodal distribution for forested areas, indicating a high concentration of decorrelated pixels. An inspection of the scene shows patterns of a weather event apparent from a sharp gradient in coherence across the scene. Similar patterns are observed from scenes 52104 to 52108, though they are not as pronounced. On the other hand, there is very high coherence in northern scenes, particularly 52126 (in Figures 5.5 and 5.6 this is represented by the cyan line with dotted markers) with mean coherences as high as 0.92 for forested areas.

To further investigate the effects of weather phenomenon, the coherence for agricultural and forested areas along the entire SIR-C swath was analyzed as a function of latitude. The result is shown in Figure 5.10. Each point on this curve is the averaged correlation magnitude of forested or agricultural land for strips of 100 cross-track In-SAR lines, representing roughly 9600 hectares for each point and plotted as a function of the latitude measured at the center of each strip. To assure that the averaged coherence value per strip is representative, mean values were not considered for sample populations of less than 50 hectares (200 samples) for each land-cover classification. This was particularly necessary for strips dominated by water.

The first thing to notice in Figure 5.10 is that irrespective of latitude there is always some amount of temporal decorrelation for either land-cover type. Secondly, agricultural areas almost always suffered less temporal decorrelation than forests. The difference seen between mean coherence for each land-cover class is statistically significant at almost every latitude except for the points from the very last scene 52126 (around 45° latitude). There also seems to be a general upward trend in coherence as a function of latitude, with the highest coherence seen farthest north. The highest temporal decorrelation is seen at lower latitudes, as much as 70% at one point.

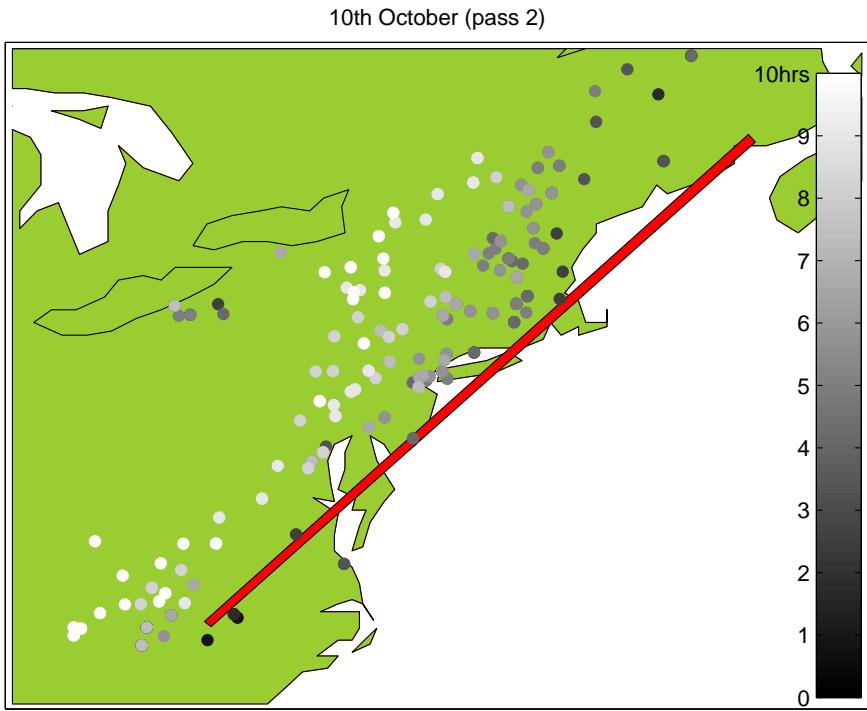


Figure 5.11: Hourly precipitation leading up to pass 2 with zero indicating the actual time of SIR-C overpass (3am on the 10th). The radar swath is approximated by the red polygon, each dot represents a station that reported some amount of precipitation during this time. The color of each dot represents time before the second pass (white is precipitation ten hours before the shuttle overpass, i.e. 5pm on the 9th, black represents rain at 3am on the 10th).

As discussed earlier, and indicated in Figure 5.10, the reasons for such high temporal decorrelation at lower latitudes could be weather events. The sharp gradient visible in coherence scene 52102, shows up as a sudden jump in coherence at a latitude of 36° . This latitude is at the center of scene 52102. Although the severity of temporal decorrelation depends on land-cover class, the similarity in trends for both classifications also points towards a weather event as a possible source of this decorrelation. Similarly, between 37° to 39° of latitude there appears a region of increased decorrelation, where the amount of decorrelation peaks near 40%, and seems to correspond to gradients observed in scenes 52104 to 52108.

The weather events that may have caused this decorrelation over a short period of 24 hours are likely precipitation and high winds. To investigate further we look at the NCDC archives of hourly surface climatological data from some 1500 weather stations in relative proximity to the area mapped by this swath. Hourly precipitation and wind data collected around the time of the SIR-C overpass from these stations was analyzed and presented in Figures 5.11 and 5.12. Figure 5.11 shows precipitation data from weather stations for the second pass of SIR-C. In the figure, a solid red polygon approximates the SIR-C swath, and each circle represents the weather station that reported a non trivial amount (more than 0.5 mm/hr) of precipitation. The color of each circle represents how many hours before the second SIR-C pass that the station recorded precipitation. White, for example represents precipitation 10 hours before the second pass while a black dot represents precipitation observed at the time of the second SIR-C overpass. No precipitation is reported up to 12 hours before the first SIR-C data take on October 9th from weather stations surrounding the swath in the eastern U.S. One can see from Figure 5.11 that a large weather system develops around 2pm (EST) on the 9th (12 hours prior to the second data take) and moves northeast with some stations reporting moderate rain (up to 8mm/hr). By the time of the second data take (3am EST on the 10th) most of the area around the SIR-C swath has no active precipitation except at the very southern tip. The amount of precipitation recorded by weather stations in this region at this time is fairly high, up to 13mm/hr, indicative of moderate to heavy rainfall. Hence, these results show that the high temporal decorrelation seen in Figure 5.10 around 36° is very likely to be caused by active precipitation.

Similarly, wind data at the time of both the SIR-C overpasses is analyzed and plotted in Figure 5.12. High winds are seen between latitudes 37° and 39° , sometimes as high as 14m/s. These winds are strong enough to cause branches to move, and may be the reason for the decorrelation seen in Figure 5.10 for this range of latitudes.

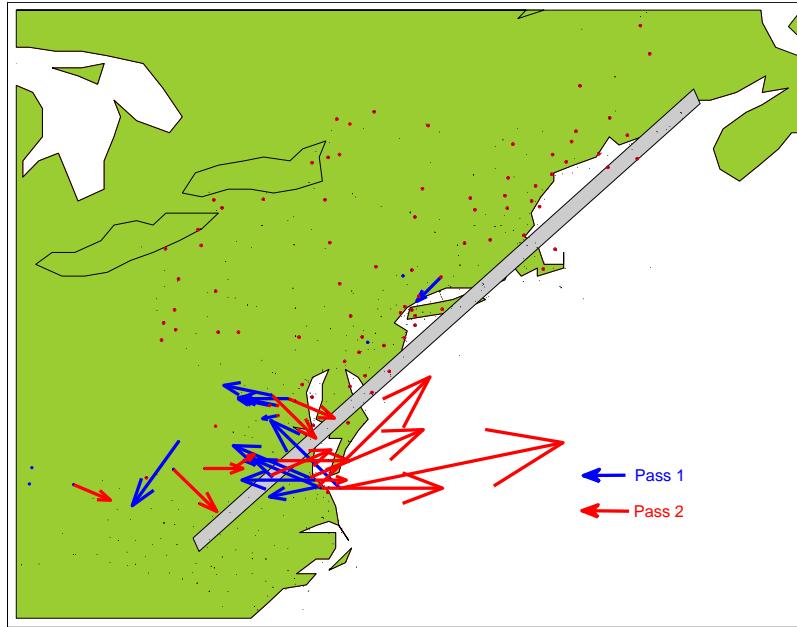


Figure 5.12: Wind data from weather stations for the two shuttle passes over the eastern US is plotted. Radar swath is approximated by the gray polygon. Difference between times at which wind is measured at the weather stations and SIR-C shuttle overpass is around 3 minutes. Arrow length indicates wind speed with the largest arrow representing a speed of around 14m/s.

The northern most scene (pr52126) is characterized by low temporal decorrelation, only about 8%. There is a lack of wind and precipitation for this area at the time of the SIR-C overpasses. As indicated on Figure 5.10, peak fall color and leaf drying occurs sometime during mid-October for latitudes 43° and above with leaf-off conditions occurring two weeks later. Hence, in the northern-most region, most deciduous trees would be either dry or in leaf-off conditions. Temperatures for the Washington county (the location of scene) weather stations at the time of SIR-C overflight suggest temperature ranges between 40° and 50° F. A combination of stable weather conditions and dry or no leaves on deciduous forests seem a likely contributor for the low values of temporal decorrelation.

In conclusion, it is evident that some amount of temporal decorrelation is present throughout this dataset, that it varies with land-cover type and that the degree of temporal decorrelation is likely weather (precipitation and wind) and seasonally (i.e. phenology and cultivation cycles) dependent. Active precipitation seems to cause up to 70% loss in coherence, wind also seems to decrease coherence, although by not as much. Stable weather and fall foliage conditions may lead to high coherence, as seen in some of the scenes over northern latitudes analyzed here. Estimates of tree heights or forest biomass from a repeat-pass interferometric system with a large temporal separation between its two passes would likely suffer from impacts of weather driven temporal decorrelation that are hard to detect and even harder to model and correct for.

CHAPTER 6

POLARIMETRIC INTERFEROMETRY

Polarimetric radar backscatter is sensitive to orientation, shape and dielectric properties of the scattering medium [73, 19]. The ability of polarimetric radars to switch between different polarimetric transmit and receive states is exploited to highlight various scattering mechanisms and obtain information about the physical characteristics of the scatterer. Because of this ability, polarimetry has been extremely useful for land-classification applications [127, 41, 18]. Interferometry on the other hand is more sensitive to height, vertical structure and density of scatterers, consequently it has been used to generate very accurate elevation models [34]. The combination of polarimetry and interferometry, referred to as PolInSAR (for Polarimetric Interferometric SAR), allows for the estimation of vertical structure by using both polarimetry and interferometry in a complementary manner. In the following chapter we discuss polarimetric interferometry and its application to forest parameter inversion.

6.1 Mathematical Formulation

We start by considering target scattering vectors in the Pauli basis [18], defined as

$$\vec{k}_i = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{hh} + S_{vv} & S_{hh} - S_{vv} & \sqrt{2}S_{hv} \end{bmatrix}^T \quad (6.1)$$

with $i = 1, 2$ for either channels of an interferometer and S_{xx} the scalar received field at the i^{th} channel for a specific transmit-receive linear polarization pair. The interfero-

metric measurement of a single-baseline fully polarimetric instrument is encapsulated in the 6×6 generalized coherency matrix, Λ_2 , defined by [16, 19]

$$\Lambda_2 = \left\langle \begin{bmatrix} \vec{k}_1 \\ \vec{k}_2 \end{bmatrix} \begin{bmatrix} \vec{k}_1^\dagger & \vec{k}_2^\dagger \end{bmatrix} \right\rangle = \begin{bmatrix} \mathbf{T}_{11} & \boldsymbol{\Omega}_{12} \\ \boldsymbol{\Omega}_{12}^\dagger & \mathbf{T}_{22} \end{bmatrix} \quad (6.2)$$

where $\langle \rangle$ represents the expected value (approximated by spatial averaging in practice), ‘ \dagger ’ denotes a complex conjugate, \mathbf{T}_{ii} is the 3×3 complex Hermitian correlation matrix that contains polarimetric information while $\boldsymbol{\Omega}_{12}$ is the 3×3 non-Hermitian complex matrix that contains the polarimetric and interferometric information of the scatterer. The complex-valued interferometric coherence for a particular polarization combination is given by [16, 95]

$$\gamma(\vec{w}_1, \vec{w}_2) = \frac{\langle \vec{w}_1^\dagger \boldsymbol{\Omega}_{12} \vec{w}_2 \rangle}{\sqrt{\langle \vec{w}_1^\dagger \mathbf{T}_{11} \vec{w}_1 \rangle \langle \vec{w}_2^\dagger \mathbf{T}_{22} \vec{w}_2 \rangle}} \quad (6.3)$$

where \vec{w}_i , (with subscript i denoting the interferometric channel) is a 3×1 complex unitary vector that enables the selection of a particular transmit or receive polarization. Because a change of polarization basis can be used to highlight particular scattering mechanisms (e.g. dipole scattering, double-bounce, etc.), \vec{w}_i is often referred to as the scattering basis vector. For instance, to select the interferometric coherence between the first Pauli basis ($S_{hh} + S_{vv}$) at either end of the baseline, the scattering vectors are set to $\vec{w}_{1,2} = [1 \ 0 \ 0]^T$. This formulation is extended to include a change of basis that would allow coherence from a polarization combination on the Poincaré sphere to be synthesized by transforming the generalized coherency matrix itself, such that [19]

$$\Lambda'_2 = \begin{bmatrix} \mathbf{U}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_3 \end{bmatrix} \begin{bmatrix} \mathbf{T}_{11} & \boldsymbol{\Omega}_{12} \\ \boldsymbol{\Omega}_{12}^\dagger & \mathbf{T}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{U}_3^\dagger & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_3^\dagger \end{bmatrix} \quad (6.4)$$

where Λ'_2 is the transformed generalized coherency matrix, and \mathbf{U}_3 is a 3×3 unitary rotation matrix. This rotation matrix is parameterized by the two Deschamps parameters α_w and δ_w that define any polarization state on the Poincaré sphere, as shown in Figure 6.1. This rotation matrix is given by $\mathbf{U}_3 = \mathbf{U}_3^L \mathbf{U}_{LP3}$, where the matrix \mathbf{U}_3^L

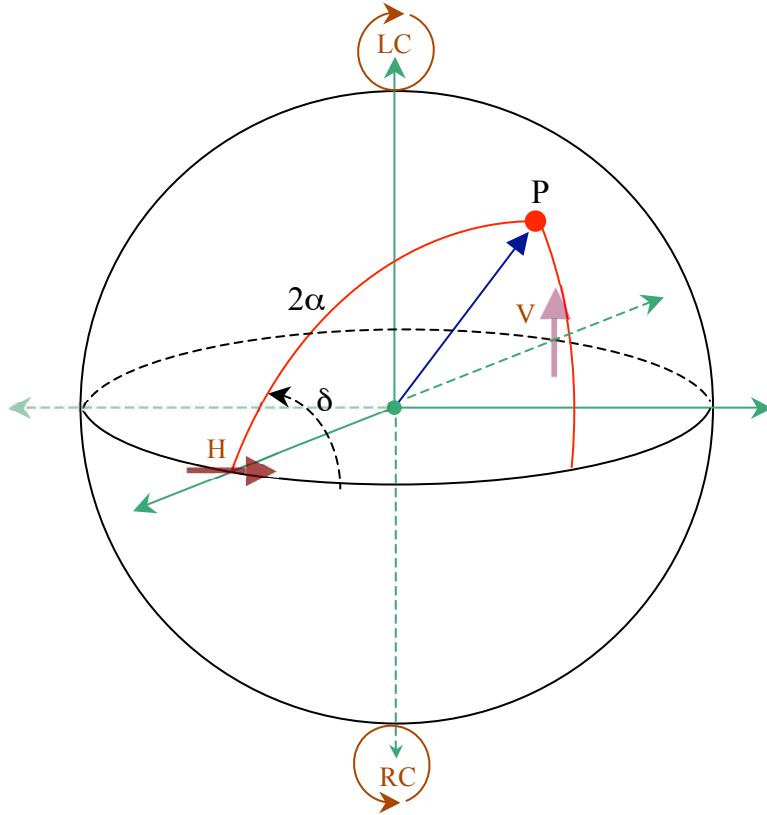


Figure 6.1: Parameterization of the Poincaré sphere using the Deschamps parameters. The location of a point, P , on the Poincaré sphere is illustrated using the Deschamps parameters α_w and δ_w .

describes the rotation of the linear (H , V) basis to the arbitrary point defined by the two Deschamps parameters α_w and δ_w , and is given by [19]

$$\mathbf{U}_3^L = \begin{bmatrix} \cos^2 \alpha_w & -\sqrt{2} \cos \alpha_w \sin \alpha_w e^{-i\delta_w} & \sin^2 \alpha_w e^{-i2\delta_w} \\ \sqrt{2} \cos \alpha_w \sin \alpha_w e^{-i\delta_w} & \cos^2 \alpha_w - \sin^2 \alpha_w & -\sqrt{2} \cos \alpha_w \sin \alpha_w e^{-i\delta_w} \\ \sin^2 \alpha_w e^{i2\delta_w} & \sqrt{2} \cos \alpha_w \sin \alpha_w e^{i\delta_w} & \cos^2 \alpha_w \end{bmatrix} \quad (6.5)$$

while the rotation matrix \mathbf{U}_{LP3} describes the transformation from the Pauli basis to the linear (H,V) basis, given by [18]

$$\mathbf{U}_{LP3} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & -1 & 0 \end{bmatrix}. \quad (6.6)$$

The complex coherence of an arbitrary polarization combination can then be synthesized using elements of the transformed generalized coherency matrix, Λ'_2 in (6.3). In general the choice of the scattering vector, \vec{w}_i is differentiated into two classes [90, 19], the so called *Single Scattering Mechanism* (SSM), characterized by the restriction; $\vec{w}_1 = \vec{w}_2$ and the *Multiple Scattering Mechanism* (MSM), characterized by the less restrictive condition; $\arg(\vec{w}_1^\dagger \vec{w}_2) = 0$. Once the generalized coherency matrix has been transformed to the arbitrary choice of polarization using (6.5) and (6.6) the SSM condition reduces to selecting only the diagonal elements of the sub-matrices of Λ'_2 , namely \mathbf{T}'_{ii} and $\boldsymbol{\Omega}'_{12}$, while MSM allows selection of some off-diagonal elements as well.

6.1.1 The coherence region

Of crucial importance to polarimetric interferometry is the spread of the set of synthesized coherences in the complex plane. A larger spread offers a better chance of inverting for physical characteristics of the target using scattering models, on the other hand, if coherences do not vary much with polarizations, then polarimetry is of little use. A useful tool for visualizing the spread of polarimetric coherences is the boundary of the region in the complex plane that contains all possible polarimetric combinations for an arbitrary generalized coherency matrix, Λ_2 . For the constrained SSM ($\vec{w}_1 = \vec{w}_2$) scenario, the estimation of this boundary region is based on the

eigenvalue decomposition of linear combinations of the sub-matrices of the coherency matrix, given by [36, 21]

$$\mathbf{T}^{-1}\boldsymbol{\Omega}_H\vec{w} = \lambda\vec{w} \left\{ \begin{array}{l} \boldsymbol{\Omega}_H = \frac{1}{2} \left(\boldsymbol{\Omega}_{12}e^{i\phi} + \boldsymbol{\Omega}_{12}^\dagger e^{-i\phi} \right) \\ \mathbf{T} = \frac{1}{2} (\mathbf{T}_{11} + \mathbf{T}_{22}) \end{array} \right. \quad (6.7)$$

where \mathbf{T} and $\boldsymbol{\Omega}_H$ are both 3×3 complex matrices and therefore yield three complex eigenvectors. The boundary of the coherence region is estimated by varying the free phase parameter ϕ in (6.7) over the range $0 \leq \phi \leq \pi$ and estimating the complex coherence for the maximum and minimum eigenvalue for each ϕ , such that

$$\begin{aligned} \lambda_{max}, \vec{w}_{max} \rightarrow \gamma_{max}(\phi) &= \frac{\vec{w}_{max}^\dagger \boldsymbol{\Omega}_{12} \vec{w}_{max}}{\vec{w}_{max}^\dagger \mathbf{T} \vec{w}_{max}} \\ \lambda_{min}, \vec{w}_{min} \rightarrow \gamma_{min}(\phi) &= \frac{\vec{w}_{min}^\dagger \boldsymbol{\Omega}_{12} \vec{w}_{min}}{\vec{w}_{min}^\dagger \mathbf{T} \vec{w}_{min}} \end{aligned} \quad (6.8)$$

where λ_{max} , \vec{w}_{max} , and λ_{min} , \vec{w}_{min} are the maximum and minimum eigenvalue/eigenvector pairs of the matrix $\mathbf{T}^{-1}\boldsymbol{\Omega}_H$ respectively.

It has been shown [18, 19] that for a pure surface scatterer without SNR or temporal effects, the entire coherence region collapses to a point on the unit circle at its topographic phase angle. For a random volume scatterer, the region again collapses to a point. However, unlike a surface scatterer, the location of this point is no longer on the unit circle, rather it lies inside the circle at a location determined by its complex volumetric decorrelation. In either case, polarimetry proves to be of little use. Polarimetric interferometry, however, is potentially beneficial when there is a combination of surface and volume scattering in the return signal. The so called *Random Volume over Ground* (RVoG) two layer scattering model [120, 121], defines the observed complex coherence as a combination the surface and volumetric coher-

ences modulated by a polarization dependent parameter, $\mu(\vec{w})$, the ratio of effective surface-to-volume scattering. The RVoG model is given by [121, 95, 17]

$$\tilde{\gamma}(\vec{w}) = \tilde{\gamma}_{sur} \frac{\tilde{\gamma}_{vol} + \mu(\vec{w})}{1 + \mu(\vec{w})} \quad (6.9)$$

where $\tilde{\gamma}_{sur}$, the surface coherence is modeled as

$$\tilde{\gamma}_{sur} = e^{i\phi(z_0)} \quad (6.10)$$

and $\phi(z_0)$ is the topographic phase. The volumetric coherence in (6.9) is modeled as

$$\tilde{\gamma}_{vol} = e^{ik_z z_0} \frac{\int_0^{h_v} f_v(z') e^{ik_z z'} dz'}{\int_0^{h_v} f_v(z') dz'} \quad (6.11)$$

where k_z is the vertical wavenumber, z_0 is the ground height, h_v is the height of the volume and $f_v(z)$ is the vertical structure function, given by

$$f_v(z) = \exp\left(\frac{2\sigma_e z}{\cos \theta_0}\right). \quad (6.12)$$

In the above, σ_e is the two way scalar extinction coefficient and θ_0 is the incidence angle. The RVoG model of the observed coherence in (6.9) reduces to volume-only scattering when $\mu = 0$ and surface scattering as μ approaches infinity. For a general scenario where varying degrees of surface and volume scattering exist, the RVoG model is an equation for a straight line in a complex plane. The slope of this line depends on interferometric baseline, vegetation height and extinction, while the length of the line additionally depends on the amplitude of the ground scattering. Theoretically then, all synthesized polarimetric coherences, obtained for a particular pixel using (6.3) should lie along a straight line in the complex plane. Figure 6.2 illustrates the RVoG model as a straight line in the complex coherence plane with a ground phase

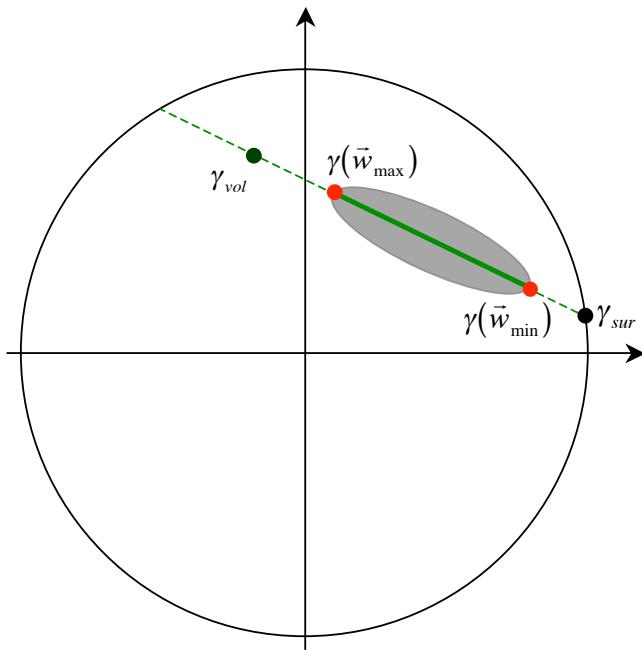


Figure 6.2: Illustration of the RVoG model. The two layer scattering model, is shown in the complex coherence plane with ground coherence $\gamma_{sur} = e^{-\phi(z_0)}$, and two *optimal* coherences $\gamma(\vec{w}_{\max})$ and $\gamma(\vec{w}_{\min})$.

$\phi(z_0)$, and two *optimal* coherences $\gamma(\vec{w}_{\max})$ and $\gamma(\vec{w}_{\min})$ [15, 95], that are extrema of the boundary region $\{\gamma_{\max}(\phi), \gamma_{\min}(\phi) \mid 0 \leq \phi \leq \pi\}$ described in (6.8). The *visible* region of the RVoG model, shown by the solid green line, describes the coherences that can be synthesized with the data. The *surface-only* coherence, γ_{sur} and the *volume-only* coherence aren't always visible in the data due to signal extinction, so they are shown on the theoretical dashed green line. The grey shaded area is a more practical example of the coherence region which is no longer a perfectly straight line due to measurement uncertainties.

6.2 PolInSAR Parameter Inversion

The two layer RVoG model is based on physical parameters of a scatterer such as a forest. Of those parameters, the two most important are topographic phase and volumetric height. The PolInSAR formulation thus enables us to estimate forest height from a polarimetric interferometric measurement alone, without relying on an external DEM or simplistic assumption about forest structure. The algorithm for such a parameter inversion follows a three stage process outlined in [17]. The first stage involves synthesis of the polarimetric coherences as described in (6.3) and calculation of the boundary region using (6.8). In the second stage topographic phase is estimated from the synthesized polarizations and the boundary region based on the RVoG model. The third stage involves estimation of forest height using a polarization channel with volume-only coherence and the topographic phase from stage two. The first stage of this inversion process has already been described in Section 6.1, a discussion of the other two stages follows.

6.2.1 Surface topography estimation

The easiest method for estimating surface topography would be to choose a polarimetric coherence, such as the HH+VV channel, that is known to be dominated by

surface scattering. The topographic phase would then simply be $\phi_{sur} = \arg(\tilde{\gamma}_{HH+VV})$. Unfortunately, coherence from this (or in practice any other) channel is contaminated by a complex volumetric decorrelation due to a significant thickness of the volume scattering layer, that biases the topographic phase. In practice [95, 53, 90], estimation of topographic phase is based on the assumption in the RVoG model that coherences from a two layer scatterer all lie on a straight line in the complex plane. Theoretically, the surface-only coherence, with $\mu = 0$, will have a magnitude of unity as it experiences no volumetric decorrelation, lying therefore at the intersection of the line with the unit circle. To find that intersection point, a straight line is fit to the set of coherences, $\tilde{\Gamma}$, comprised of the synthesized polarizations (or the boundary region points, or both), such that

$$\begin{aligned} & \left\{ \tilde{\Gamma} : \tilde{\Gamma} \in \tilde{\gamma}(\alpha_w, \delta_w), \quad \forall \quad 0 \leq \alpha_w \leq \pi/2, \quad -\pi \leq \delta_w \leq \pi \right\} \\ & \left\{ \tilde{\Gamma} : \tilde{\Gamma} \in \tilde{\gamma}_{max}(\phi), \tilde{\gamma}_{min}(\phi), \quad \forall \quad 0 \leq \phi \leq \pi \right\}. \end{aligned} \quad (6.13)$$

The line is defined with the slope M and intercept C , such that

$$\begin{aligned} \Im\{\tilde{\Gamma}\} &= M \Re\{\tilde{\Gamma}\} + C \\ y &= Mx + C \end{aligned} \quad (6.14)$$

where $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote the real and imaginary parts respectively, while the slope, M is estimated by

$$\hat{M} = \frac{-c_1 \pm \sqrt{c_1^2 - 4c_2c_0}}{2c_2} \quad (6.15)$$

where the parameters c_0 , c_1 and c_2 are given by

$$\begin{aligned}
c_0 &= - \sum_i (x_i - \bar{x})(y_i - \bar{y}) \\
c_1 &= \sum_i \{(x_i - \bar{x})^2 - (y_i - \bar{y})^2\} \\
c_2 &= \sum_i (x_i - \bar{x})(y_i - \bar{y})
\end{aligned} \tag{6.16}$$

where $\bar{x} = \frac{1}{N} \sum_i x_i = \frac{1}{N} \sum_i \Re\{\tilde{\Gamma}_i\}$ and $\bar{y} = \frac{1}{N} \sum_i y_i = \frac{1}{N} \sum_i \Im\{\tilde{\Gamma}_i\}$ for $N = n(\tilde{\Gamma})$, the number of synthesized coherences (or boundary region points or both). The intercept C is estimated by

$$\hat{C} = \bar{y} - \hat{M}\bar{x}. \tag{6.17}$$

The intersection of this line with the unit circle is the extinction corrected topographic phase. This point can be found by simultaneously solving the two equations, $x^2 + y^2 = 1$ and $y = \hat{M}x + \hat{C}$, giving

$$\left. \begin{array}{l} x_p = \frac{-\hat{M}\hat{C} \pm \sqrt{\hat{M}^2 - \hat{C}^2 + 1}}{1 + \hat{M}^2} \\ y_p = \hat{M}x_p + \hat{C} \end{array} \right\} \rightarrow e^{i\phi} = x_p + iy_p. \tag{6.18}$$

The line defined by (\hat{M}, \hat{C}) intersects the unit circle at two points, either of which could be a plausible solution. Different methods exist for resolving this ambiguity. Figure 6.3 illustrates two common approaches [19]. The first approach, shown in Figure 6.3a, relies on the assumption that the complex coherence from channel HV is similar to the *volume-only* coherences thus a smaller value of μ compared to coherences from either HH+VV or HH-VV channels that are similar to the *surface-only* channel with higher μ -values. Rank ordering these coherences based on the μ -value along the best fit line suggests coherence, γ_{s_1} , as the channel with surface only scattering and thus the best choice for estimating topographic phase. The second method, shown in Figure 6.3b, is based on interferometry and relies on the fact that if the volume layer is above the surface layer then the phase differences between the volume-only and

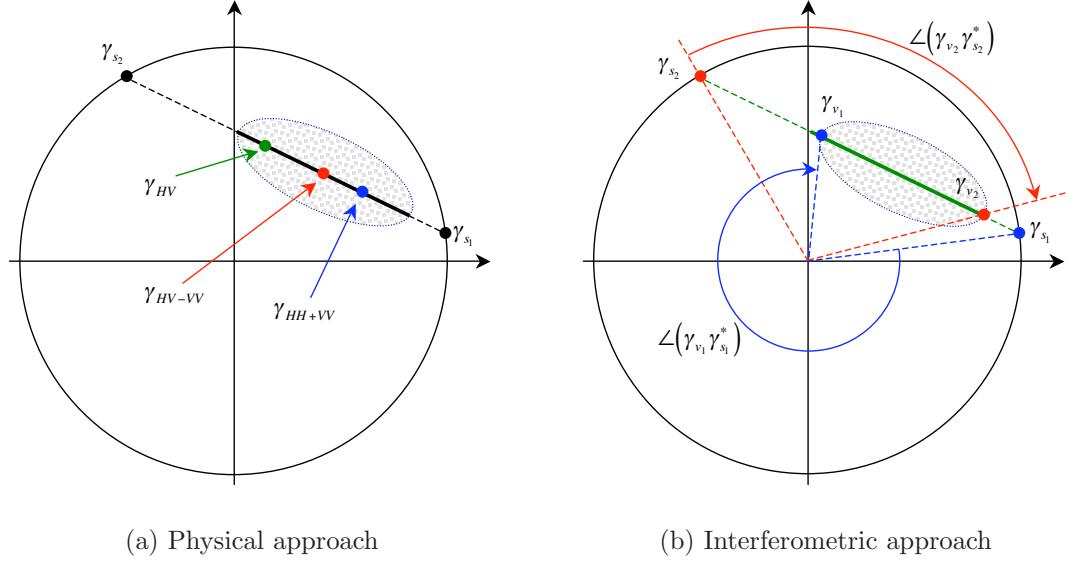


Figure 6.3: Illustration of methods for resolution of topographic phase ambiguity.

surface-only coherences must be measured in a clockwise direction. The two possible topographic coherences, γ_{s_1} and γ_{s_2} , have two corresponding volume-only coherences that belong to $\tilde{\Gamma}$ with the lowest μ -value relative to either topographic coherence. In this case, these two coherences are shown here by γ_{v_1} and γ_{v_2} , and their phase differences by $\angle(\gamma_{v_1}\gamma_{s_1}^*)$ and $\angle(\gamma_{v_2}\gamma_{s_2}^*)$ (measured clockwise) respectively. The larger phase difference, in this case $\angle(\gamma_{v_1}\gamma_{s_1}^*)$, can be rejected, leaving γ_{s_2} as the logical choice for the topographic coherence.

6.2.2 Height inversion

The final step in the parameter inversion algorithm [17] involves estimating the height of the scattering volume. In addition to the topographic phase estimated previously, this requires an estimate of the complex coherence with a phase center located near the top of the volume, or volume-only coherence $\tilde{\gamma}_v$. It is not guaranteed that such a coherence could be synthesized with the available data because of mean wave extinction and variation in the vertical structure of the vegetation itself. Under

these circumstances the best approximation is to use a polarimetric coherence from the set $\tilde{\Gamma}$ with the highest μ -value. This would be akin to finding the coherence farthest from the topographic phase along the best fit line, i.e. $\max \|\tilde{\gamma}_s - \tilde{\Gamma}\|$. If the interferometric method was used in estimating the topographic phase as described in the previous section, such a coherence is already estimated. Forest height is then simply given by

$$h_v = \frac{\arg(\tilde{\gamma}_v \tilde{\gamma}_s^*)}{k_z} \quad (6.19)$$

where γ_s and $\tilde{\gamma}_v$ are the surface- and volume-only coherences respectively while k_z , the vertical wavenumber, is given by

$$k_z = \frac{4\pi B_n}{\lambda R \sin \theta} \quad (6.20)$$

where B_n is the perpendicular interferometric baseline assuming a repeat-pass or double baseline configuration, R is the range to target, λ the radar wavelength and θ the incidence angle. This height, however, is underestimated because the phase center of the volume-only coherence is not guaranteed to be at the top of the volume. The structure-dependent bias must therefore be accounted for by model based correction methods. More specifically, this is accomplished by using the two parameter full-RVoG model to estimate the extinction corrected height and the extinction coefficient for a particular forest pixel [95, 17]. The extinction coefficient is calculated by minimizing the difference between the observed volume-only coherence $\tilde{\gamma}_v$ and the model for volumetric coherence of the form outlined in (6.11). This can be written mathematically as [19]

$$\min_{h_v, \sigma} \left\| \tilde{\gamma}_v - e^{i\phi_0} \frac{p}{p_1} \frac{e^{p_1 h_v} - 1}{e^{p h_v} - 1} \right\| \quad (6.21)$$

where, ϕ_0 is the phase of the surface-only coherence, i.e. $\phi_0 = \angle(\tilde{\gamma}_s)$, and

$$p = \frac{2\sigma}{\cos \theta} \quad (6.22)$$

$$p_1 = p + ik_z \quad (6.23)$$

where σ is the extinction coefficient and θ is the look angle. This is usually implemented by simple iterative searching algorithms such as the simplex method.

6.3 UAVSAR PolInSAR Data

Part of the polarimetric dataset collected by UAVSAR during the Harvard Forest campaign of 2009 (described in section 4.3) has been processed interferometrically at the Jet Propulsion Laboratory. Table 6.2 summarizes thirteen of the forty lines that were processed as a stack of polarimetric interferograms, such that an interferogram can be formed between any polarization and baseline combination by a simple correlation. Each UAVSAR line can be uniquely identified by its ‘Track Name’, but for this particular subset each scene is given a simpler ‘Track Number’. Track Number 6, from a scene collected on the 8th of August, is used as the reference image for the entire stack. Each of the thirteen scenes in Table 6.2 have been coregistered to the slant range coordinates of the reference image and have gone through the process of range-spectral filtering [102, 46] and removal of flat-earth and topographic phases [54, 67]. Table 6.2 summarizes the cross-track baselines for all possible pairs from the stack of thirteen scenes. The flight-tracks of all the lines flown over the Harvard Forest were at the same altitude and the spatial-separation was designed to be only in the horizontal (or cross-track) direction. That, coupled with the UAVSAR platform’s ability to fly very precise tracks [55] meant that the vertical component of all baselines was negligible. The baseline information in Table 6.2, therefore, sufficiently summarizes the relative positioning of the antennas. Table 6.3 summarizes the temporal separation between each of the possible interferometric pairs. For each scene, the polarimetric UAVSAR data is distributed as four single look complex (SLC) images

Track Number	Track Name	Cross-track Baseline (m)	Vertical Baseline (m)	Baseline Length (m)	Baseline Angle (deg)	Date	Time of Day	Delta Time (hrs)
1	harvrd_00501_005_001	35.5	0	35.50	0.0	August 6, 2009	13:59:57	-49.7
2	harvrd_00502_005_003	25.3	-0.1	25.30	-0.2	August 6, 2009	14:31:27	-49.2
3	harvrd_00504_005_005	10.1	-0.6	10.12	-3.4	August 6, 2009	15:01:56	-48.7
4	harvrd_00508_005_007	-20	0	20.00	180.0	August 6, 2009	15:32:54	-48.2
5	harvrd_00503_005_003	16.3	0.4	16.30	1.4	August 8, 2009	15:13:26	-0.5
6	harvrd_00505_005_005	0	0	0.00	0.0	August 8, 2009	15:43:51	0.0
7	harvrd_00509_005_007	-30.2	0.1	30.20	179.8	August 8, 2009	16:14:49	0.5
8	harvrd_00510_005_004	-44.9	0.2	44.90	179.7	August 13, 2009	19:15:09	123.5
9	harvrd_00511_005_006	-73.3	0.2	73.30	179.8	August 13, 2009	19:48:39	124.1
10	harvrd_00507_005_003	-14.7	-0.3	14.70	-178.8	August 16, 2009	13:51:11	190.1
11	harvrd_00506_005_007	-4.7	-0.2	4.70	-177.6	August 17, 2009	14:57:43	215.2
12	harvrd_00512_005_003	40.5	0.2	40.50	0.3	August 17, 2009	13:58:03	214.2
13	harvrd_00513_005_001	50.3	-0.4	50.30	-0.5	August 17, 2009	13:28:53	213.8

Table 6.1: Summary for the thirteen UAVSAR lines collected over a period of eleven days in August 2009 over the Harvard Forest that have been processed as a stack of interferograms. The reference track (track number 6) is highlighted in yellow.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0	10.2	25.4	55.5	19.2	35.5	65.7	80.4	108.8	50.2	40.2	-5	-14.8
2		0	15.2	45.3	9	25.3	55.5	70.2	98.6	40	30	-15.2	-25
3			0	30.1	-6.2	10.1	40.3	55	83.4	24.8	14.8	-30.4	-40.2
4				0	-36.3	-20	10.2	24.9	53.3	-5.3	-15.3	-60.5	-70.3
5					0	16.3	46.5	61.2	89.6	31	21	-24.2	-34
6						0	30.2	44.9	73.3	14.7	4.7	-40.5	-50.3
7							0	14.7	43.1	-15.5	-25.5	-70.7	-80.5
8								0	28.4	-30.2	-40.2	-85.4	-95.2
9									0	-58.6	-68.6	-113.8	-123.6
10										0	-10	-55.2	-65
11											0	-45.2	-55
12												0	-9.8
13													0

Table 6.2: Cross-track baselines (in meters) between all pairs for the thirteen UAVSAR lines. These lines were collected over a period of eleven days in August 2009 over the Harvard Forest. The number identifying each line in this table is the field ‘Track Number’ from Table 6.1.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.0	-0.5	-1.0	-1.5	-49.2	-49.7	-50.2	-173.3	-173.8	-239.9	-265.0	-264.0	-263.5
2		0.0	-0.5	-1.0	-48.7	-49.2	-49.7	-172.7	-173.3	-239.3	-264.4	-263.4	-263.0
3			0.0	-0.5	-48.2	-48.7	-49.2	-172.2	-172.8	-238.8	-263.9	-262.9	-262.4
4				0.0	-47.7	-48.2	-48.7	-171.7	-172.3	-238.3	-263.4	-262.4	-261.9
5					0.0	-0.5	-1.0	-124.0	-124.6	-190.6	-215.7	-214.7	-214.3
6						0.0	-0.5	-123.5	-124.1	-190.1	-215.2	-214.2	-213.8
7							0.0	-123.0	-123.6	-189.6	-214.7	-213.7	-213.2
8								0.0	-0.6	-66.6	-91.7	-90.7	-90.2
9									0.0	-66.0	-91.2	-90.2	-89.7
10										0.0	-25.1	-24.1	-23.6
11											0.0	1.0	1.5
12												0.0	0.5
13													0.0

Table 6.3: Temporal separation (in hours) between all pairs for the thirteen UAVSAR lines. These lines were collected over a period of eleven days in August 2009 over the Harvard Forest. The number identifying each line in this table is the field ‘Track Number’ from Table 6.1.

for the four polarimetric channels HH, HV, VH and VV. In addition, a LOS (line of sight) file describing position of the UAVSAR platform relative to each single-look pixel, given by a look vector in the SCH coordinates [54], is distributed for each scene. This look vector information can be used to derive the exact baseline for any of the interferometric pairs and their corresponding vertical wavenumber, k_z , necessary for height estimation.

6.3.1 Estimating the vertical wavenumber

The vertical wavenumber, k_z defined as the rate of change of interferometric phase, ϕ , with respect to the height, h , is given by

$$k_z = \frac{\partial \phi}{\partial h} = \frac{2\pi\rho}{\lambda} \frac{1}{\langle \hat{l}, \hat{n} \rangle} \left\langle \vec{b}, -\hat{l} + \frac{\hat{g}}{\langle \hat{l}, \hat{g} \rangle} \right\rangle \quad (6.24)$$

where, ρ , the range to a pixel at location (x, y) is given by

$$\rho = \left| \vec{l}_r(x, y) \right| \quad (6.25)$$

where \vec{l}_r is the look vector for the reference image pointing from the radar to the pixel. The unit look vector, \hat{l} is then given by

$$\hat{l} = \frac{\vec{l}_r(x, y)}{\rho} \quad (6.26)$$

and the baseline vector, \vec{b}_{jr} , describing the spatial separation between the reference line and an arbitrary j^{th} line is given by

$$\vec{b}_{jr}(x, y) = \vec{l}_r(x, y) - \vec{l}_j(x, y) \quad (6.27)$$

where \vec{l}_j is the look vector of the j^{th} line. Finally, \hat{g} , the unit ground vector tangent to the surface is given by

$$\hat{g} = \frac{\hat{l} - \langle \hat{l}, \hat{n} \rangle \hat{n}}{\left| \hat{l} - \langle \hat{l}, \hat{n} \rangle \hat{n} \right|}. \quad (6.28)$$

The unit surface normal vector, \hat{n} , for a flat surface can simply be considered as $\hat{n} = [0, 0, 1]$ in the SCH coordinates, that the LOS files are distributed in.

6.3.2 Polarization synthesis and coherence optimization

The next step in the PolInSAR processing of UAVSAR data is the synthesis of polarimetric coherences using (6.3) and estimation of the boundary region based on coherence optimization from (6.7) and (6.8). Figure 6.4 shows synthesized polarizations and the coherence region of two pixels from interferograms using UAVSAR lines 2 and 3 (see Table 6.2). The set of synthesized polarimetric coherences, shown here by the pale green markers, are estimated by varying α_w and δ_w over their respective ranges in (6.5). For each pair of α_w and δ_w , the Λ'_2 matrix is computed using (6.4) and finally the scalar polarimetric coherence is estimated from (6.3) with 144 looks (using a 6×24 averaging window). Additionally, the boundary of the region encompassing all possible coherences is estimated from (6.8) by varying the free phase

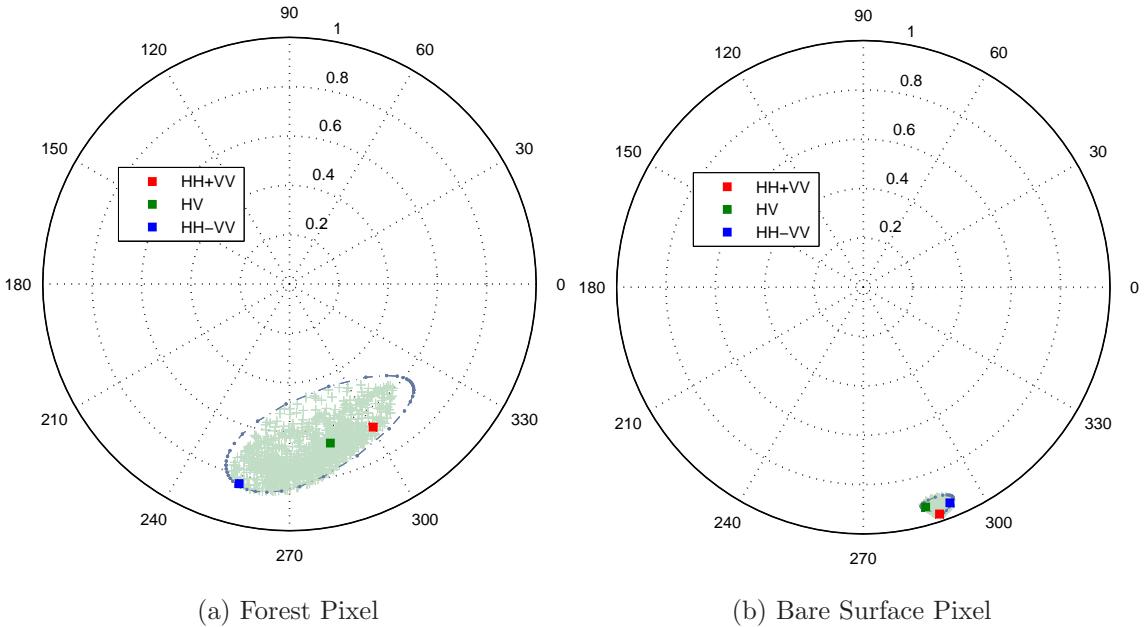


Figure 6.4: Polarization synthesis and coherence optimization for UAVSAR data. Examples of polarization synthesis and the boundary region from two UAVSAR pixels one a forest and the other a bare-surface pixel.

parameter, ϕ between 0 and 2π . Figure 6.4a shows the synthesized coherences and boundary region from a forest pixel. This particular pixel seems to have both surface and volume scattering indicated by a large variation in the synthesized coherences. The three Pauli-basis coherences are also shown. No synthesized polarimetric coherence, including the HH+VV and HH-VV coherences, approach the unit circle. With sufficient SNR in this pixel and a small temporal separation of one half of an hour (see Table 6.3) between the two UAVSAR passes, it is safe to assume that the observed decorrelation is mostly volumetric in nature. Furthermore, it is also evident that volumetric decorrelation is present in every polarimetric channel, necessitating the use of the RVoG model to estimate the topographic phase. Figure 6.4b shows the synthesized coherences and the boundary region for a bare-surface pixel (from a swampy area with no forest cover). The polarimetric coherences in this case, are tightly grouped and much closer to the unit circle, exhibiting very little volumetric,

thermal or temporal decorrelation. Unlike this particular pixel (which had some signal likely due to double bounce scattering from the interface of small vegetation and water) most bare-surface pixels are dominated by noise. Coherence for pixels dominated by thermal noise or observing ambiguities (likely ISLR or azimuth ambiguities), generally biased low, requires correction.

6.3.3 Multi-polarization SNR correction

Thermal noise is well known to cause decorrelation in interferometric echoes [138, 108]. To extract quantitative estimates from the volumetric coherence, it is important to correct for thermal effects by normalizing the observed coherence by expected thermal decorrelation, which is modeled as

$$\gamma_{thermal} = \frac{1}{\sqrt{1 + \text{SNR}_1^{-1}} \sqrt{1 + \text{SNR}_2^{-1}}} \quad (6.29)$$

where the signal to noise ratio at each end of the baseline, SNR_1 and SNR_2 , is given by

$$\text{SNR} = \frac{\sigma}{\sigma_0} \quad (6.30)$$

where σ is the scattered intensity in a particular pixel and σ_0 is the noise power, also known as the ‘noise equivalent sigma zero’ (NESZ). The framework for SNR correction, however, is not well established for the multi-polarization scenario. Here two methods are explored in an attempt to better understand the problem.

The first method relies on the assumption that noise is independent between the two polarimetric channels p and q , such that the correlation between the scattered fields is given by

$$\langle S_p S_q \rangle = \begin{cases} \sigma_{pq} & p \neq q \\ \sigma_{pq} + \sigma_0 & p = q \end{cases}. \quad (6.31)$$

Similarly, it can also be assumed that noise is independent between two interferometric channels. Theoretically then, noise should decorrelate everywhere except along the

diagonal of the generalized correlation matrix, Λ_2 if the linear lexicographic basis set (HH, HV, VH, VV) is used. It should therefore be possible to subtract the scalar NESZ from the diagonal of the Λ_2 matrix and correct for thermal noise. Figure 6.5

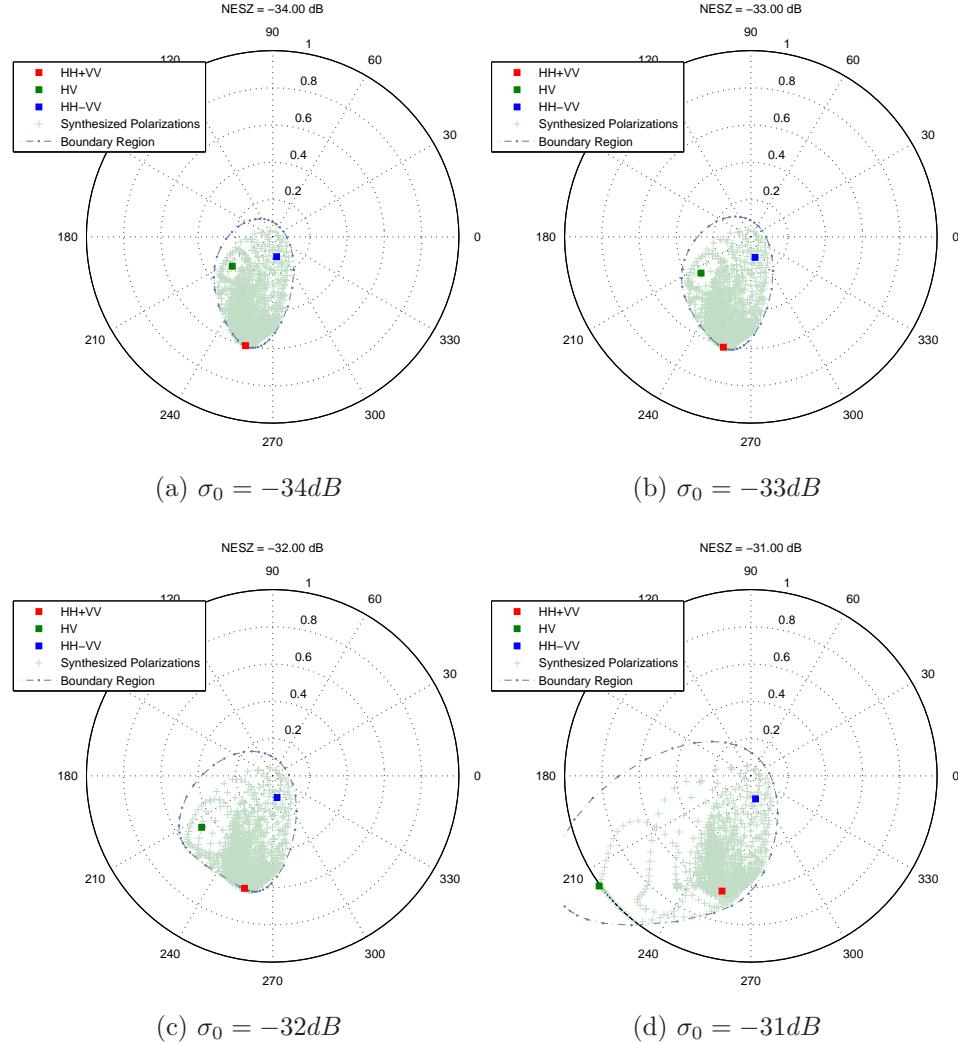


Figure 6.5: Examples of SNR correction with varying NESZ. Four cases of SNR correction using varying NESZ of -34dB, -33dB, -32dB and -31dB. An NESZ of -31dB NESZ seems to be over-estimated, evident by the coherence region boundary and some synthesized coherences located outside the unit circle.

shows this particular correction method applied to a pixel over a flat area in the UAVSAR imagery with a backscatter coefficient of -27dB. Since the maximum NESZ for UAVSAR is -40dB [55], there should be at least 14dB of signal to noise ratio in

this pixel. Here the interferometric correlation is estimated using 144 samples (from a 6×24 spatial average). Correcting the Λ_2 matrix with a scalar NESZ of -40dB makes no difference to the synthesized polarizations. Figure 6.5 assess the potential impact of a higher noise levels in the radar by varying the NESZ between -35dB and -31dB. The coherence region expands as NESZ is increased, to the point that some coherences become larger than unity (indicative of an overestimated NESZ). Since this is a bare-surface pixel, SNR correction should have moved all coherences to the unit circle, ideally to the same phase angle as well. The fact that the region expands, instead of contracting and translating to the unit circle shows that this method may not prove to be the most feasible option.

Another possible method for multi-polarization noise correction would be to divide each synthesized polarization by its respective modeled thermal coherence, such that

$$\gamma_{corr}(\vec{w}_1, \vec{w}_2) = \frac{\gamma(\vec{w}_1, \vec{w}_2)}{\gamma_n(\vec{w}_1, \vec{w}_2)} \quad (6.32)$$

where the coherence due to the combined effects of thermal and multiplicative noise, γ_n , is modeled as

$$\gamma_n(\vec{w}_1, \vec{w}_2) = \frac{\text{SNR}(\vec{w}_1, \vec{w}_2)}{\text{SNR}(\vec{w}_1, \vec{w}_2) (\text{MNR}(\vec{w}_1, \vec{w}_2)^{-1} + 1) + 1} \quad (6.33)$$

where the polarization dependent Multiplicative Noise Ratio, $\text{MNR}(\vec{w}_1, \vec{w}_2)$ is given by

$$\frac{1}{\text{MNR}} = \frac{1}{\text{QNR}} + \text{ISLR}_\rho + \text{ISLR}_s + \frac{1}{\text{AMB}_t} \quad (6.34)$$

where QNR is the quantization noise ratio, ISLR is the integrated sidelobe ratio in the range, ρ , and azimuth, s , direction and AMB_t is the total (range and azimuthal) ambiguity. For the sake of simplicity QNR, ISLR and the range ambiguity are ignored,

leaving azimuth ambiguity the leading contributor to multiplicative noise. The true azimuthal ambiguity for a pixel with intensity, σ_p , is given by

$$\text{AMB}_z = \frac{G_t G_r (\sigma_u + \sigma_l)}{\sigma_p} \quad (6.35)$$

where G_t and G_r are the transmit and receive antenna gains, while σ_u , σ_l are the backscatter intensities of the upper and lower ambiguity pixels respectively. The nominal radar ambiguity (a term commonly estimated while characterizing radar systems) is given by

$$\widetilde{\text{AMB}} = 2G_t G_r. \quad (6.36)$$

The true ambiguity can thus be written as a function of the nominal ambiguity as

$$\text{AMB}_z = \left(\frac{\sigma_u + \sigma_l}{2\sigma_p} \right) \widetilde{\text{AMB}}. \quad (6.37)$$

The location of the first azimuthal ambiguity with respect to the pixel in question at a particular range and azimuth (ρ, s) is given by [26]

$$\Delta s = \pm \frac{f_p V_{st}}{f_R} \quad (6.38)$$

where f_p is the pulse repetition frequency of the radar, V_{st} is the velocity of the aircraft relative to the target and f_R is the Doppler rate used in the processor. Figure 6.6 shows an example of a pixel from UAVSAR imagery that has been corrected for SNR and MNR effects using a nominal ambiguity ratio of 19dB and a noise equivalent sigma zero of -35dB. The intensities of the upper and lower ambiguities for each polarimetric channel are chosen as the mean background intensities of the entire image instead of the actual ambiguity pixel for simplicity. The effect of correcting for SNR and MNR in such a manner seems to yield better results than those seen in Figure 6.5 for the

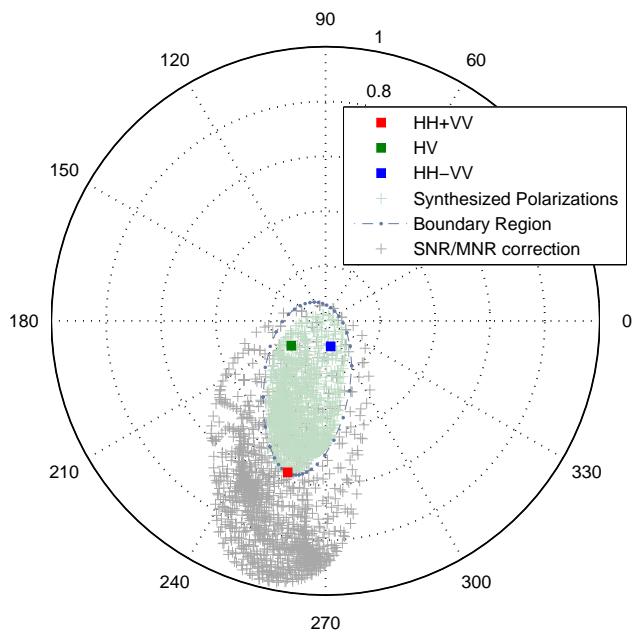


Figure 6.6: Example of SNR and MNR correction for a *dark* pixel. A nominal ambiguity of 19dB and a noise equivalent sigma zero (NESZ) of -35dB are used. Colored symbols indicate the observed coherences while greyed ones indicate the corrected coherences.

corresponding NESZ value of -35dB. The result seems to suggest that the *darker* UAVSAR pixels may be dominated by multiplicative noise rather than actual signal. However, a more detailed analysis and better models of multi-polarization SNR/MNR coherence are necessary.

6.3.4 Estimation of topographic phase

For forest pixels that generally have high SNR, noise correction is less of a problem. The next stage in the PolInSAR inversion process [17] is the estimation of the surface-only polarization channel leading to estimation of topographic phase and estimation of the corresponding volume-only channel. The topographic phase estimation process, described in Section 6.2.1, relies on the two layer RVoG model given in (6.9) where a straight line is fit to either the synthesized coherences, the boundary region points, or both. The intersection of this line with the unit circle is chosen as the topographic phase. Figure 6.7 shows examples of the RVoG line fits to four different forest pixels. The best fit line (using the boundary region coherences only) is shown by the black line that intersects the unit circle at two distinct locations. The interferometric approach from Figure 6.3b is used to resolve this ambiguity (using clockwise-only estimation of phase differences). In each of the cases shown in Figure 6.7, the best choice for topographic phase and the corresponding volume-only channel, as the coherence farthest from the topographic phase along the best fit line, are shown by the black squares (labeled as ‘chosen pair’). The angular difference between the chosen surface-only and volume-only coherence, highlighted by thick black line, when normalized by the vertical wavenumber leads to PolInSAR derived height estimates.

6.4 PolInSAR Initial Results

The phase difference between the volume- and surface-only coherences can be used to estimate heights using (6.19). Even though forest heights estimated using

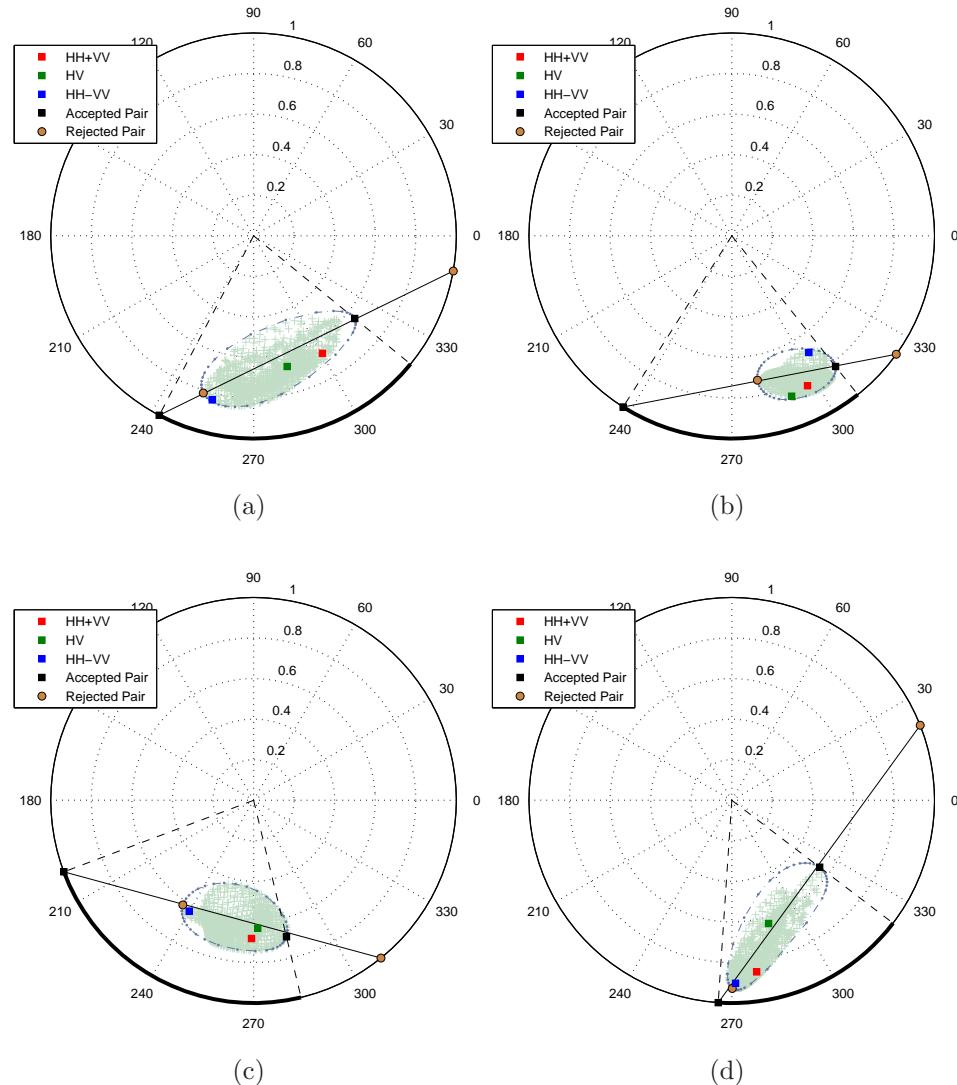


Figure 6.7: Examples of straight line fits to polarimetric coherences (using points from the boundary region only) and selection of the pair of ground and volume coherence channels from four forest pixels with high SNR.

this method are likely to be underestimated due to the penetration of radar waves into the canopy, it is often employed for its simplicity and practicality. For this very reason, this method of height estimation was applied to the stack of polarimetric UAVSAR interferograms over a small region near the Harvard Forest site. Figure



(a) Aerial photograph



(b) Radar backscatter image

Figure 6.8: Aerial photograph and radar image of a region measuring roughly $1\text{km} \times 0.5\text{km}$ near Petersham, MA chosen for analysis of PolInSAR performance.

6.8a shows an aerial photograph of the patch while Figure 6.8b shows the backscatter intensity from one of the scenes from the UAVSAR stack, multilooked using a 3×12 averaging window. The multilooked intensity image is in slant range radar coordinates so it is not perfectly coregistered with the geolocated aerial photograph.

This patch, measuring roughly $1\text{km} \times 0.5\text{km}$, is in an area of relatively small topo-

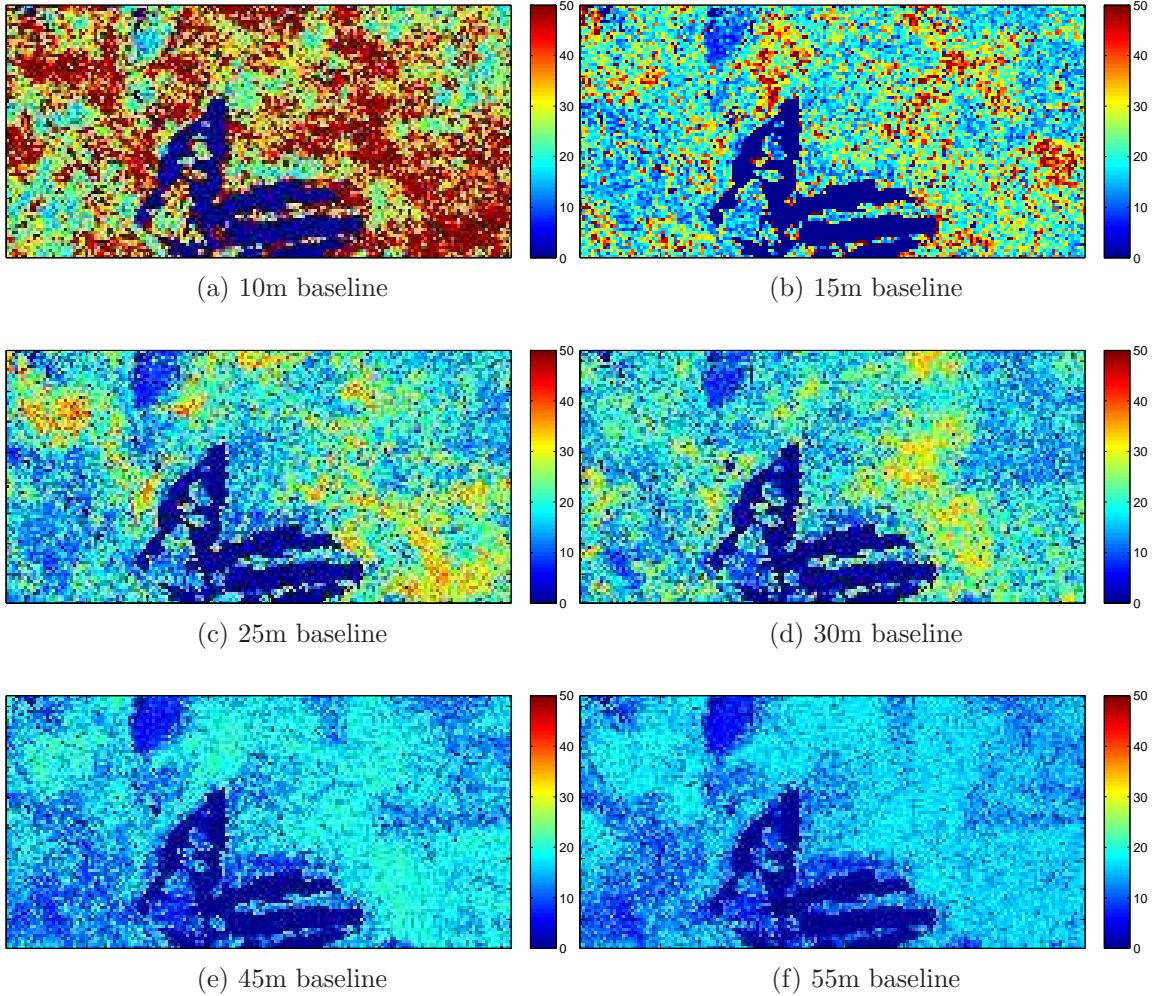


Figure 6.9: PolInSAR height estimates obtained from six baselines varying between 10m and 55m over a region near the Harvard Forest site.

graphic relief just south of the Harvard Forest test site. The region has different types of land-cover, comprising mostly of mixed forest with a swamp at the top-left of the image and a golf-course in the center. PolInSAR heights were estimated for each pixel in the patch using (6.19). Figure 6.9 shows the results of the PolInSAR height inversion method applied to six pairs with baselines varying from 10m to 55m. These pairs, combinations of tracks 1 through 4 (see Table 6.1), are chosen because they have minimal temporal separation with the maximum repeat-time of 90 min-

utes. Even though effects of temporal decorrelation cannot be ruled out completely, it is expected that such small temporal separation between the two UAVSAR passes would minimize the impact of temporal decorrelation on PolInSAR height estimates. The low SNR regions, such as the golf-course pixels, are masked out. Even though

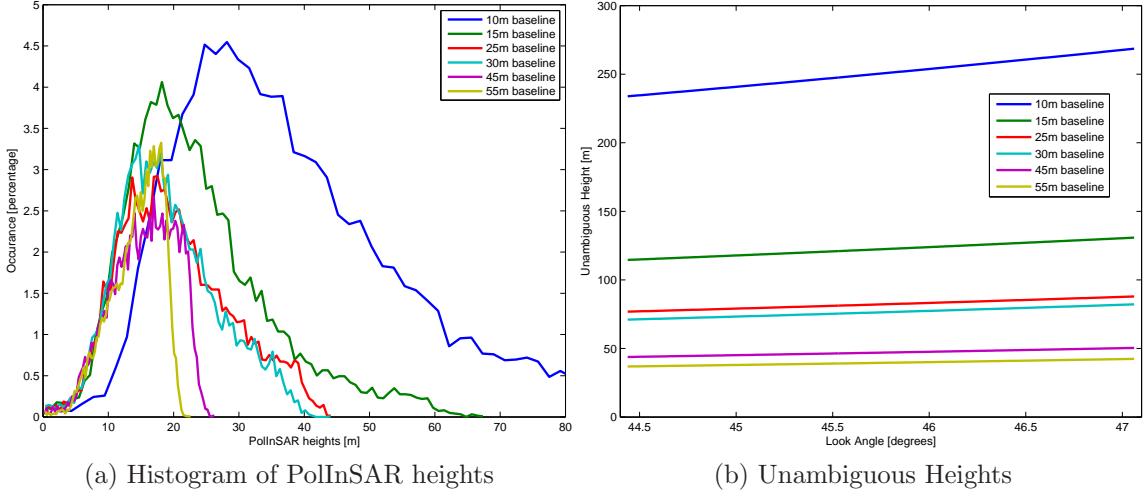


Figure 6.10: Histograms of PolInSAR height estimates from six UAVSAR baselines. These baselines vary from 10 to 55 meters in length and their corresponding unambiguous heights as a function of the varying look angle across the image section are also shown.

UAVSAR NESZ is extremely low (-45dB in this region) the coherences in these pixels are low, indicating that the signals from such pixels are dominated by noise, likely multiplicative in nature (such as azimuthal ambiguities, or integrated sidelobe ratios). Without an adequate method for correcting these errors, height estimates from these pixels are biased and unreliable. Other than the 10 and 15m baseline pairs, formed by correlating data from tracks 1 with 2 and 2 with 3 respectively, height estimates are consistent for the image, albeit different for different baselines. Heights for pixels over the swamp (top left of the image) are estimated to be below 5m, consistently, by each pair. These pixels have high SNR, likely due to a large double bounce scatter from the interface of water and low lying vegetation. Height estimates from the 10 and 15m baselines are unnaturally high (with height estimates of up to 100m) in large

part due to the small k_z values of these baselines, essentially serving to magnify small errors in phase-difference estimates. Figure 6.10a plots the histogram of all heights estimated in this image section, while Figure 6.10b shows the unambiguous heights for each of the interferometric pairs, given by $2\pi/k_z$, as a function of look angle that varies between 44 and 47 degrees across this section. In addition to the overestimated heights for the 10 and 15m baselines, the saturation of PolInSAR heights for large baselines of 45 and 55 meters at 20 and 25 meters is apparent. These saturation heights correspond to half the ambiguity height. This is a direct consequence of adopting the interferometric approach to resolve the ground phase ambiguity as discussed in Section 6.3.4. Since the ground-volume pair with the smallest angular difference is chosen to estimate volumetric extent, the phase difference never exceeds π , saturating heights at π/k_z instead of $2\pi/k_z$. With the exception of the 10, baseline, the majority of PolInSAR height estimates, however, are consistent as evident by the coincident modes of the height distributions.

6.4.1 Temporal stability of PolInSAR height estimates

Statistics of the PolInSAR height estimates in Figure 6.10a are assumed to be unaffected by temporal decorrelation because all the scenes were acquired within a short time span, with the largest repeat time of 90 minutes. With large temporal separations between the two passes of a repeat-pass interferometer (a more likely mission scenario) the effects of temporal decorrelation cannot be assumed to be absent. The impact of temporal decorrelation on PolInSAR heights has been demonstrated before [96] for different eco-regions and forests. Figure 6.11 shows the impact of temporal decorrelation on PolInSAR heights from UAVSAR repeat-pass data for the region near the Harvard Forest test site. Figure 6.11a shows the histograms of PolInSAR heights estimated using four different 15m baselines with repeat-times varying from one half of an hour up to ten days (263 hours), while Figure 6.11b plots the distri-

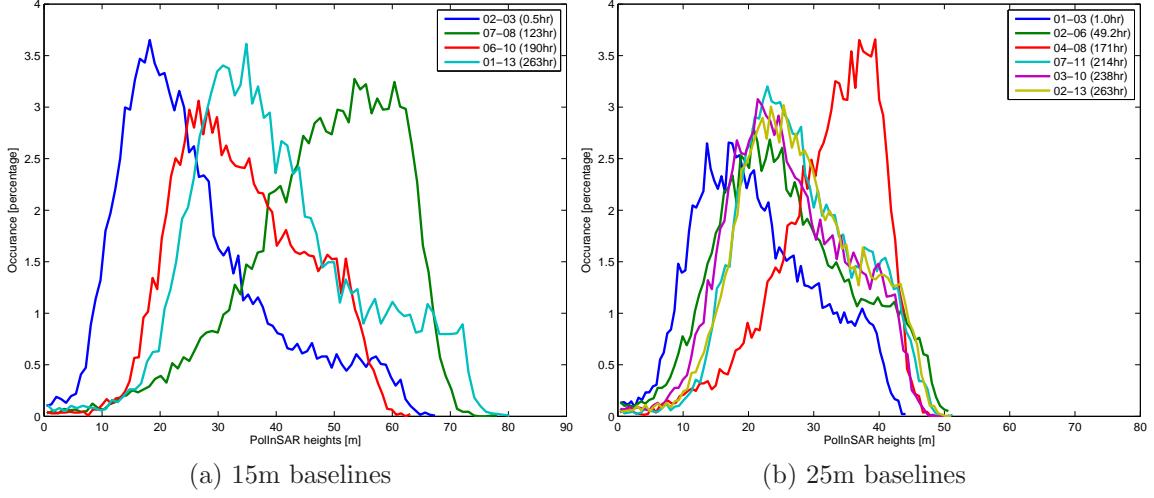


Figure 6.11: Effect of temporal decorrelation on UAVSAR PolInSAR height estimates. These heights are estimated from from 15 and 25m baselines over a variety of different repeat-times. The impact of precipitation before the acquisition of data on track 8 is evident by the large bias in height estimates from baselines that include track 8.

bution of heights derived from six different 25m baselines with repeat-times varying around the same range. In each case, pairs with shortest temporal separations lead to smallest heights, with a progressive increase in mean height as the repeat-times increase, highlighting the general overestimation of PolInSAR heights as temporal decorrelation increases with greater repeat-times. There is a marked deviation, however, from this trend in each case. During the UAVSAR data acquisition campaign, there was a precipitation event on the 13th of August, affecting data from tracks 8 and 9. Rain and wind were shown in Chapter 5 to cause significant temporal decorrelation in repeat-pass data. The rain event immediately prior to the acquisition of tracks 8 and 9 effects the PolInSAR heights dramatically as well. The heights are markedly over-estimated form the pair of tracks 7 and 8 shown in Figure 6.11a by the dark green line. Similarly, PolInSAR heights from the pair between tracks 4 and 8 are also visibly over estimated than the other 25m baselines, even though some have longer repeat-times.

6.4.2 Preliminary comparison with lidar heights

A comparison between PolInSAR derived heights and lidar heights is of obvious interest. Lidar measurements (in particular RH100) are often seen to be best correlated with field measurements of canopy tops [30, 60, 75] and are consequently considered good estimators of forest height. LVIS lidar metrics should therefore serve as a means of judging the quality of PolInSAR height estimates. However, the interaction of radar waves with a scattering volume is physically different from the interaction of light pulses with the same object. This complicates a one-to-one comparison of PolInSAR heights and lidar RH metrics. Figure 6.12 highlights this particular differ-

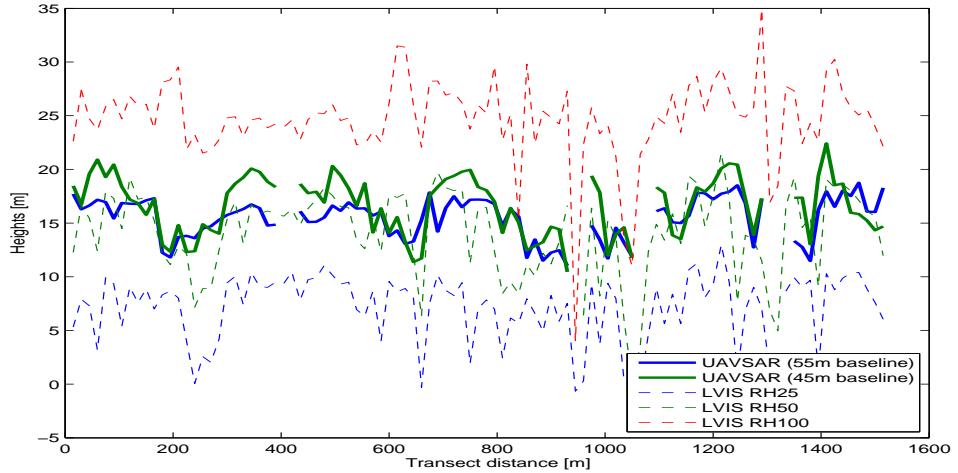


Figure 6.12: Comparison of PolInSAR and lidar heights. PolInSAR heights are estimated using UAVSAR data and lidar heights are chosen from a transect through the region near the Harvard Forest site.

ence, by plotting PolInSAR heights and LVIS RH metrics for a transect through the image section shown in Figure 6.8. The PolInSAR heights (from baselines between tracks 01-04 and 02-04) are on average around 15 meters, almost the same height as the RH50 metric. This similarity is likely due to the penetration of the radar waves through the canopy, however, the effect of the π ambiguity height for these baselines (at 25 meters) cannot be ruled out. Figure 6.13 plots PolInSAR heights from four different UAVSAR baselines against the LVIS RH50 metric for all pixels in the image

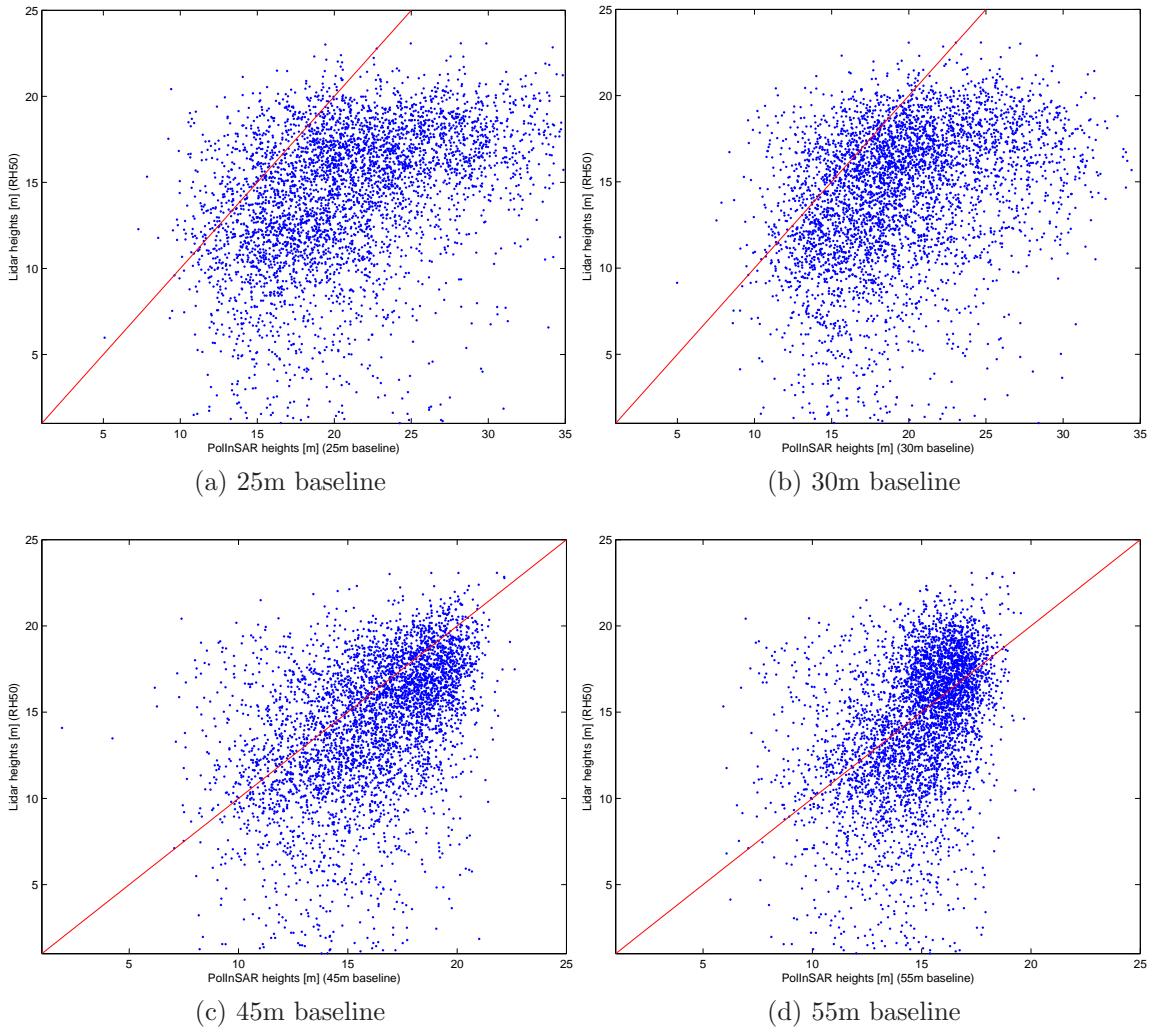


Figure 6.13: Comparison of PolInSAR heights from 25, 30, 45 and 55m baselines and the LVIS RH50 metric that seems to behave most similar to the PolInSAR heights.

section from Figure 6.8. The overestimation of PolInSAR heights from the 25m and 40m baselines is quite apparent in Figures 6.13a and 6.13b. A similar analysis of PolInSAR heights form these two baselines and the RH75 or RH100 metric does not yield any better results. It is likely that the PolInSAR estimates are biased. PolInSAR heights from the 45m baseline and the LVIS RH50 metric shown in Figure 6.13c seems to have the best correlation, while heights from the 55m baseline seem to be saturated at 19 meters (the π ambiguity height for this baseline).

In the absence of a PolInSAR error model it is hard to ascertain how much of the error in the relationship between lidar heights and PolInSAR estimates is due to physical scattering differences and how much is due to error in PolInSAR estimates themselves. A parametric PolInSAR error model is therefore a necessity. For the purposes of this dissertation, such an error model is not included, however it is noted as an important future task. The ability to process UAVSAR data and obtain reasonable PolInSAR height estimates has, however, been established, and as such this chapter serves as a stepping stone towards future research where questions such as the accuracy of these estimates and their relationship with forest biomass will be explored thoroughly.

CHAPTER 7

CONCLUSIONS AND FUTURE WORK

7.1 Summary

This dissertation sought to assess the accuracy of biomass and structure estimates from remotely sensed data. The analysis presented here focused on lidars and radars, the two most commonly used remote sensing instruments for estimating forest parameters. With forests of the northeastern United States of primary focus, two study sites, the Harvard Forest in Petersham, Massachusetts and the Howland Forest in central Maine were chosen. During field campaigns in the summer of 2009, data including tree diameters and species information, among others, were inventoried from 15 hectares in the Harvard Forest and 23 hectares in the Howland Forest. This diameter data was used to estimate individual tree weights using allometry, and aggregated to generate hectare-scale biomass estimates. Three different sets of allometric equations were used to estimate field biomass. The accuracies of these biomass estimates including allometric and measurement errors were quantified. Full-waveform lidar data, from the NASA/GSFC LVIS instrument flown over the two study sites in August 2009 was analyzed to assess the performance of lidar metrics in estimating forest biomass. Regression statistics in the form of R^2 and RMSE from linear fits between field data and lidar metrics were used to ascertain the quality of lidar estimates of biomass. Radar backscatter data, from the NASA/JPL UAVSAR instrument collected over the Harvard and Howland Forests in August 2009 as well, was used to explore the relationship between field biomass and radar backscatter. The accuracy of backscatter-biomass relationships were explored using a radar backscatter error model. The two measure-

ment errors, in radar backscatter and field estimates of biomass, were included in a statistical framework allowing for the estimation of measurement-error corrected confidence intervals on estimates of biomass from radar backscatter. The impact of temporal decorrelation on the accuracy of height and consequently biomass estimates from repeat-pass InSAR systems was characterized. The accuracy of biomass estimates from InSAR heights, explored using SRTM data were shown to be on par with lidar estimates, however, those estimates were not effected by temporal decorrelation. In order to assess the impact of a repeat-pass configuration on height and biomass estimates, temporal decorrelation was analyzed using single-day repeat pass SIR-C data from 1994. Finally, polarimetric-interferometric data from the 2009 UAVSAR campaign over the Harvard Forest was used to estimate forest height. Multiple baseline combinations were used to assess the consistency of PolInSAR heights and the impact of varying repeat-times were discussed. Finally, PolInSAR heights were compared with lidar heights for a small region near the Harvard Forest test site.

7.2 Conclusions

The analysis of allometric relationships and field estimates of biomass suggest that mean biomass estimates vary significantly with the type of allometric equations used. Accuracy statistics suggests that the BLUE (Best Linear Unbiased Estimator) allometric equations [71] contain the least error at both the hectare and subplot level spatial resolutions for equations that were developed from data spanning a large spatial extent. At subplot scales, the errors in the BLUE allometric equations are around the 8 tons/ha average, while at hectare-scales this average reduces to 5 tons/ha. The Ensemble equations of [65] contain the most error, although the error statistics are not rigorous and therefore unreliable. Single-site allometry [118] tend to have the least error, however this error is underestimated due to inaccurate statistics and the narrow spatial extent of the datasets used to generate these allometric coefficients.

Preliminary analysis of the LVIS data suggested good correlations over the Howland Forest and poor performance over the Harvard Forest. Metrics other than the simple energy quartiles were also tested over the Harvard Forest to see if there was some other metric better related to biomass without much success. RH100 was shown to still be the best predictor of biomass. Combinations of LVIS metrics did not attain the high correlations reported in the literature either. It was observed that using different types of allometric equations to predict biomass from diameter data has a significant impact on lidar performance. Use of simplified equations and excluding low biomass data from the Howland Forest dataset seemed to make the lidar-biomass relationships over the two sites comparable, with RMSE values of 50% of the biomass standard deviation the best result and corresponding R^2 values between 0.6 and 0.7 at hectare scales. At subplot scales the relationships between height and biomass seemed weaker, with the best R^2 values between 0.3 and 0.4 and RMSEs between 60 and 70% of the biomass variation. Combining the two datasets using simplified allometries suggest a possible saturation trend in the height-biomass relationship, explaining the poor performance of lidar over the Harvard Forest where generally higher biomass values were recorded.

Results of the regression analysis between field estimates of biomass, outlined in Chapter 2 and backscatter data from the UAVSAR field campaign over the Harvard and Howland Forests shows that measurement error corrected confidence intervals from the non-linear backscatter-biomass model are wide enough that they cannot be inverted to obtain the corresponding intervals on biomass estimates from radar backscatter at either the hectare or the subplot scales. The reasons for the large confidence interval widths are two fold; a) the saturation of the backscatter-biomass curve and b) combination of large model error and measurement error in radar backscatter. An analysis of the model error shows that a distinct bias in backscatter value between the deciduous and coniferous trees contributes to an increase in model error variance.

Removing the bias in backscatter between the two forest types reduces model error and consequently allows for estimation of confidence bounds on biomass estimates. However, the confidence intervals can only be estimated at hectare scales because at subplot scales, measurement error in radar backscatter (primarily due to speckle) tends to dominate and cause the bootstrap method to fail. The width of confidence bounds on biomass estimates at hectare scales over the Harvard and Howland forests show cross-polarized radar backscatter to be a poor estimator of forest biomass, even at low biomass levels with the width exceeding 200 percent for every biomass value.

While the relationship between interferometric heights, estimated from SRTM and LVIS elevations, and field biomass proves to be as strong, if not stronger than lidar metrics, with RMSE values as low as 34tons/ha, this result does not take into account errors due to temporal decorrelation. The analysis of single day repeat-pass data from SIR-C showed that some amount of temporal decorrelation seems to always be present with large repeat-times, that it varies with land-cover type and that the degree of temporal decorrelation is likely weather (precipitation and wind) and seasonally (i.e. phenology and cultivation cycles) dependent. Active precipitation seems to cause up to 70% loss in coherence, wind also seems to decrease coherence, although by not as much. Stable weather and fall foliage conditions may lead to high coherence values. Estimates of tree heights or forest biomass from a repeat-pass interferometric system with a large temporal separation between its two passes would likely suffer from impacts of weather driven temporal decorrelation that are hard to detect and even harder to model and correct for.

Initial analysis of PolInSAR results using UAVSAR data showed promise. The three step PolInSAR inversion process was run successfully on six UAVSAR baselines with small repeat-times (90minutes or less) over a region near the Harvard Forest test site. Height estimates from the 10 baselines were largely unreliable, however, other baselines yielded realistic and consistent results, with mean height estimates

over forested region between 15 to 20 meters. Analysis of multi-temporal 15m and 25m baselines showed an increased bias in PolInSAR height estimates with increasing temporal separation between the two passes, with the exception of one particular track that is contaminated by a precipitation event. Height estimates in that case are severely biased.

7.3 Unique Contributions

The unique contributions of this work are listed below

- Characterization of errors in field estimates of biomass over the Harvard and Howland Forests, that haven't been analyzed before.
- A consistent analysis of the performance of LVIS lidar metrics in estimating forest biomass over the Harvard and Howland Forest sites.
- Indication of a possible saturation in the lidar-biomass relationship above 300 tons/ha in this ecoregion, not seen before.
- Estimation of the errors in biomass estimates from radar backscatter over the Harvard and Howland Forests.
- Impact of measurement errors in radar backscatter and field biomass estimates were included for the first time in an analysis of the accuracy of backscatter-biomass relationships.
- An analysis of temporal decorrelation in single-day repeat-pass data that quantified the impacts of weather events on interferometric measurements.
- First analysis of PolInSAR height estimates over the Harvard Forest region from UAVSAR data.

7.4 Future Work

- Synthesizing full lidar waveforms into scalar energy metrics and using them or their combinations is the most common method for analyzing the biomass prediction capability of lidars. This was the primary focus of lidar analysis presented in Chapter 3 as well. However, lidar waveform decompositions into Fourier or Legendre basis sets for example, are not explored as frequently. Exploring waveform decompositions for forest structure and biomass estimation may hold the key to better utilization of information in lidar returns [122]. With the availability of full-waveform LVIS data over both the Harvard and Howland sites, a thorough and systematic analysis could prove to be of significant value.
- Even though comparing PolInSAR heights to lidar metrics is an often used method for assessing the accuracy PolInSAR estimates, an important component to such an analysis, the effect of measurement error on PolInSAR heights, is overlooked. The development of a PolInSAR measurement error model, where errors in measuring the scattered fields are translated through the PolInSAR algorithm to calculate the errors in heights is necessary.
- UAVSAR derived PolInSAR height estimates were not used to derive biomass over the Harvard and Howland sites. Such an analysis would be of immediate interest. Furthermore a complete analysis of the accuracy of PolInSAR derived estimates of forest biomass including parametric measurement errors would be of significance.
- With the availability of multi-baseline UAVSAR data, an analysis of tomographic structure estimates, their accuracies and ability to predict forest biomass are all questions that can be answered, and ones that could be of interest to the remote sensing community.

- Polarization coherence tomography (PCT), a recently established technique [14], exploits the information in polarimetric-interferometric data differently than the traditional PolInSAR technique. An assessment of the accuracy PCT over the Harvard and Howland forests using UAVSAR data could prove to be of significant value.

APPENDIX

VARIANCE-COVARIANCE MATRICES

The variance-covariance matrices, Σ_{ab} , used in (2.22) to estimate the error in tree weights for each of the species encountered during the field campaigns are given here.

Table A.1: Variance-Covariance matrices Σ_{ab} used for calculating the error in tree biomass estimates for the species summarized in [71]

a_{stem}	b_{stem}	a_{bark}	b_{bark}	$a_{branches}$	$b_{branches}$	$a_{foliage}$	$b_{foliage}$
Striped Maple (<i>Acer pennsylvanicum</i>)							
2.135E-06	-7.349E-06	8.980E-08	-1.051E-06	1.784E-07	-3.364E-06	9.006E-08	-8.134E-07
-7.349E-06	2.660E-05	-2.913E-07	3.795E-06	-6.489E-07	1.318E-05	-3.478E-07	3.583E-06
8.980E-08	-2.913E-07	8.311E-07	-1.047E-05	-8.198E-08	1.482E-06	-4.864E-08	3.256E-07
-1.051E-06	3.795E-06	-1.047E-05	1.356E-04	1.064E-06	-2.055E-05	6.218E-07	-4.617E-06
1.784E-07	-6.489E-07	-8.198E-08	1.064E-06	4.479E-07	-8.481E-06	2.968E-07	-2.527E-06
-3.364E-06	1.318E-05	1.482E-06	-2.055E-05	-8.481E-06	1.719E-04	-5.754E-06	5.548E-05
9.006E-08	-3.478E-07	-4.864E-08	6.218E-07	2.968E-07	-5.754E-06	8.504E-07	-7.060E-06
-8.134E-07	3.583E-06	3.256E-07	-4.617E-06	-2.527E-06	5.548E-05	-7.060E-06	6.882E-05
Red Maple (<i>Acer rubrum</i>)							
2.708E-05	-8.680E-05	3.634E-06	-3.659E-05	1.683E-06	-3.208E-05	2.590E-06	-1.487E-05
-8.680E-05	2.998E-04	-1.042E-05	1.155E-04	-6.039E-06	1.205E-04	-7.768E-06	4.514E-05
3.634E-06	-1.042E-05	7.300E-06	-8.375E-05	-1.305E-06	2.334E-05	6.529E-08	2.105E-07
-3.659E-05	1.155E-04	-8.375E-05	1.024E-03	1.279E-05	-2.382E-04	-2.657E-07	-6.604E-06
1.683E-06	-6.039E-06	-1.305E-06	1.279E-05	1.156E-05	-2.021E-04	1.308E-05	-7.916E-05
-3.208E-05	1.205E-04	2.334E-05	-2.382E-04	-2.021E-04	3.632E-03	-2.311E-04	1.456E-03
2.590E-06	-7.768E-06	6.529E-08	-2.657E-07	1.308E-05	-2.311E-04	4.259E-05	-2.548E-04
-1.487E-05	4.514E-05	2.105E-07	-6.604E-06	-7.916E-05	1.456E-03	-2.548E-04	1.601E-03
Sugar Maple (<i>Acer saccharum</i>)							
6.989E-05	-1.507E-04	6.910E-05	-3.093E-04	9.645E-08	-1.492E-06	-3.867E-06	2.702E-05
-1.507E-04	3.324E-04	-1.483E-04	6.734E-04	-1.922E-07	2.725E-06	8.160E-06	-5.775E-05
6.910E-05	-1.483E-04	1.053E-04	-4.724E-04	9.387E-07	-7.191E-06	-3.305E-06	2.549E-05
-3.093E-04	6.734E-04	-4.724E-04	2.154E-03	-3.768E-06	2.147E-05	1.557E-05	-1.284E-04
9.645E-08	-1.922E-07	9.387E-07	-3.768E-06	1.445E-05	-1.274E-04	2.751E-06	-2.019E-05
-1.492E-06	2.725E-06	-7.191E-06	2.147E-05	-1.274E-04	1.164E-03	-2.531E-05	2.071E-04
-3.867E-06	8.160E-06	-3.305E-06	1.557E-05	2.751E-06	-2.531E-05	5.807E-06	-4.519E-05
2.702E-05	-5.775E-05	2.549E-05	-1.284E-04	-2.019E-05	2.071E-04	-4.519E-05	3.860E-04
Yellow Birch (<i>Betula alleghaniensis</i>)							
3.733E-04	-5.164E-04	1.497E-06	-1.765E-05	8.990E-06	-8.868E-05	6.010E-05	-1.494E-04
-5.164E-04	7.232E-04	-2.147E-06	2.785E-05	-1.322E-05	1.352E-04	-8.449E-05	2.135E-04
1.497E-06	-2.147E-06	2.772E-06	-3.987E-05	-7.015E-07	5.762E-06	-3.255E-06	7.725E-06
-1.765E-05	2.785E-05	-3.987E-05	5.916E-04	9.501E-06	-7.803E-05	4.624E-05	-1.140E-04

<i>a_{stem}</i>	<i>b_{stem}</i>	<i>a_{bark}</i>	<i>b_{bark}</i>	<i>a_{branches}</i>	<i>b_{branches}</i>	<i>a_{foliage}</i>	<i>b_{foliage}</i>
8.990E-06	-1.322E-05	-7.015E-07	9.501E-06	5.606E-06	-5.278E-05	1.726E-05	-4.328E-05
-8.868E-05	1.352E-04	5.762E-06	-7.803E-05	-5.278E-05	5.269E-04	-1.601E-04	4.229E-04
6.010E-05	-8.449E-05	-3.255E-06	4.624E-05	1.726E-05	-1.601E-04	1.453E-04	-3.654E-04
-1.494E-04	2.135E-04	7.725E-06	-1.140E-04	-4.328E-05	4.229E-04	-3.654E-04	9.804E-04
Black Birch (<i>Betula lenta</i>)							
2.221E-04	-3.871E-04	3.899E-05	-3.055E-04	2.459E-06	-9.731E-05	-5.319E-06	3.581E-05
-3.871E-04	7.113E-04	-6.636E-05	5.411E-04	-5.084E-06	2.087E-04	7.969E-06	-3.776E-05
3.899E-05	-6.636E-05	2.582E-05	-2.012E-04	2.637E-06	-9.387E-05	-3.782E-06	3.165E-05
-3.055E-04	5.411E-04	-2.012E-04	1.626E-03	-2.236E-05	7.978E-04	2.847E-05	-2.309E-04
2.459E-06	-5.084E-06	2.637E-06	-2.236E-05	7.959E-06	-2.743E-04	1.760E-06	-1.211E-05
-9.731E-05	2.087E-04	-9.387E-05	7.978E-04	-2.743E-04	9.593E-03	-6.146E-05	4.672E-04
-5.319E-06	7.969E-06	-3.782E-06	2.847E-05	1.760E-06	-6.146E-05	8.382E-06	-7.140E-05
3.581E-05	-3.776E-05	3.165E-05	-2.309E-04	-1.211E-05	4.672E-04	-7.140E-05	7.201E-04
Paper Birch (<i>Betula papyrifera</i>)							
4.169E-06	-2.282E-05	1.243E-06	-3.042E-05	3.203E-07	-7.987E-06	7.550E-07	-5.068E-06
-2.282E-05	1.298E-04	-6.871E-06	1.738E-04	-1.761E-06	4.286E-05	-4.170E-06	2.670E-05
1.243E-06	-6.871E-06	5.919E-07	-1.461E-05	1.700E-07	-4.224E-06	4.991E-07	-3.180E-06
-3.042E-05	1.738E-04	-1.461E-05	3.739E-04	-4.207E-06	1.040E-04	-1.227E-05	7.715E-05
3.203E-07	-1.761E-06	1.700E-07	-4.207E-06	8.434E-07	-2.059E-05	1.129E-06	-7.119E-06
-7.987E-06	4.286E-05	-4.224E-06	1.040E-04	-2.059E-05	5.317E-04	-2.863E-05	1.985E-04
7.550E-07	-4.170E-06	4.991E-07	-1.227E-05	1.129E-06	-2.863E-05	7.748E-06	-4.957E-05
-5.068E-06	2.670E-05	-3.180E-06	7.715E-05	-7.119E-06	1.985E-04	-4.957E-05	3.564E-04
Grey Birch (<i>Betula populifolia</i>)							
1.881E-05	-9.439E-05	1.107E-06	-3.097E-05	-4.049E-06	1.583E-04	-1.037E-06	3.299E-05
-9.439E-05	4.972E-04	-7.091E-06	2.267E-04	1.807E-05	-6.618E-04	3.799E-06	-7.750E-05
1.107E-06	-7.091E-06	1.280E-06	-3.232E-05	3.277E-07	-1.642E-05	-8.160E-08	9.821E-07
-3.097E-05	2.267E-04	-3.232E-05	9.237E-04	-8.525E-06	5.167E-04	6.680E-07	8.746E-05
-4.049E-06	1.807E-05	3.277E-07	-8.525E-06	7.060E-06	-3.090E-04	4.594E-06	-1.841E-04
1.583E-04	-6.618E-04	-1.642E-05	5.167E-04	-3.090E-04	1.410E-02	-2.048E-04	8.630E-03
-1.037E-06	3.799E-06	-8.160E-08	6.680E-07	4.594E-06	-2.048E-04	3.872E-06	-1.564E-04
3.299E-05	-7.750E-05	9.821E-07	8.746E-05	-1.841E-04	8.630E-03	-1.564E-04	6.716E-03
American Chestnut (<i>Castanea dentata</i>)							
2.135E-06	-7.349E-06	8.980E-08	-1.051E-06	1.784E-07	-3.364E-06	9.006E-08	-8.134E-07
-7.349E-06	2.660E-05	-2.913E-07	3.795E-06	-6.489E-07	1.318E-05	-3.478E-07	3.583E-06
8.980E-08	-2.913E-07	8.311E-07	-1.047E-05	-8.198E-08	1.482E-06	-4.864E-08	3.256E-07
-1.051E-06	3.795E-06	-1.047E-05	1.356E-04	1.064E-06	-2.055E-05	6.218E-07	-4.617E-06
1.784E-07	-6.489E-07	-8.198E-08	1.064E-06	4.479E-07	-8.481E-06	2.968E-07	-2.527E-06
-3.364E-06	1.318E-05	1.482E-06	-2.055E-05	-8.481E-06	1.719E-04	-5.754E-06	5.548E-05
9.006E-08	-3.478E-07	-4.864E-08	6.218E-07	2.968E-07	-5.754E-06	8.504E-07	-7.060E-06
-8.134E-07	3.583E-06	3.256E-07	-4.617E-06	-2.527E-06	5.548E-05	-7.060E-06	6.882E-05
American Beech (<i>Fagus grandifolia</i>)							
4.073E-04	-8.046E-04	6.525E-06	-1.731E-04	1.916E-05	-1.684E-04	2.709E-05	-2.212E-04
-8.046E-04	1.610E-03	-1.345E-05	3.707E-04	-3.960E-05	3.590E-04	-5.414E-05	4.519E-04
6.525E-06	-1.345E-05	9.602E-07	-2.443E-05	-4.082E-07	2.400E-06	-1.362E-06	9.705E-06
-1.731E-04	3.707E-04	-2.443E-05	6.560E-04	7.544E-06	-3.436E-05	3.086E-05	-2.162E-04
1.916E-05	-3.960E-05	-4.082E-07	7.544E-06	1.770E-05	-1.447E-04	1.273E-05	-1.035E-04
-1.684E-04	3.590E-04	2.400E-06	-3.436E-05	-1.447E-04	1.246E-03	-1.045E-04	9.049E-04
2.709E-05	-5.414E-05	-1.362E-06	3.086E-05	1.273E-05	-1.045E-04	3.018E-05	-2.385E-04
-2.212E-04	4.519E-04	9.705E-06	-2.162E-04	-1.035E-04	9.049E-04	-2.385E-04	1.979E-03
White Ash (<i>Fraxinus americana</i>)							
2.025E-04	-3.456E-04	3.632E-05	-2.779E-04	3.003E-05	-2.040E-04	1.717E-04	-4.589E-04
-3.456E-04	6.310E-04	-6.066E-05	4.945E-04	-5.411E-05	3.758E-04	-2.868E-04	7.762E-04
3.632E-05	-6.066E-05	1.573E-05	-1.185E-04	-3.263E-06	2.171E-05	1.721E-05	-4.562E-05
-2.779E-04	4.945E-04	-1.185E-04	9.512E-04	2.248E-05	-1.551E-04	-1.301E-04	3.507E-04

<i>a_{stem}</i>	<i>b_{stem}</i>	<i>a_{bark}</i>	<i>b_{bark}</i>	<i>a_{branches}</i>	<i>b_{branches}</i>	<i>a_{foliage}</i>	<i>b_{foliage}</i>
3.003E-05	-5.411E-05	-3.263E-06	2.248E-05	1.033E-04	-6.949E-04	2.076E-04	-5.625E-04
-2.040E-04	3.758E-04	2.171E-05	-1.551E-04	-6.949E-04	4.831E-03	-1.373E-03	3.808E-03
1.717E-04	-2.868E-04	1.721E-05	-1.301E-04	2.076E-04	-1.373E-03	7.629E-04	-2.052E-03
-4.589E-04	7.762E-04	-4.562E-05	3.507E-04	-5.625E-04	3.808E-03	-2.052E-03	5.673E-03
Tamarack (<i>Larix laricina</i>)							
7.636E-06	-4.134E-05	1.513E-07	-2.764E-06	-1.096E-06	1.933E-05	-3.301E-06	1.289E-05
-4.134E-05	2.350E-04	-8.055E-07	1.824E-05	6.118E-06	-1.155E-04	1.670E-05	-6.453E-05
1.513E-07	-8.055E-07	2.508E-07	-4.881E-06	-1.008E-08	2.122E-07	-1.619E-07	5.055E-07
-2.764E-06	1.824E-05	-4.881E-06	1.011E-04	2.814E-07	-8.314E-06	2.362E-06	-6.226E-06
-1.096E-06	6.118E-06	-1.008E-08	2.814E-07	2.138E-06	-3.824E-05	6.647E-06	-2.927E-05
1.933E-05	-1.155E-04	2.122E-07	-8.314E-06	-3.824E-05	7.257E-04	-1.188E-04	5.532E-04
-3.301E-06	1.670E-05	-1.619E-07	2.362E-06	6.647E-06	-1.188E-04	3.944E-05	-1.761E-04
1.289E-05	-6.453E-05	5.055E-07	-6.226E-06	-2.927E-05	5.532E-04	-1.761E-04	8.489E-04
Hophornbeam (<i>Ostrya virginiana</i>)							
1.861E-03	-3.530E-03	4.648E-04	-2.742E-03	1.491E-04	-1.920E-03	-1.689E-04	2.145E-03
-3.530E-03	6.945E-03	-1.018E-03	6.424E-03	-2.669E-04	3.560E-03	3.326E-04	-4.382E-03
4.648E-04	-1.018E-03	4.039E-04	-2.581E-03	4.371E-05	-6.581E-04	-1.017E-05	2.066E-04
-2.742E-03	6.424E-03	-2.581E-03	1.735E-02	-2.315E-04	3.814E-03	9.128E-05	-1.790E-03
1.491E-04	-2.669E-04	4.371E-05	-2.315E-04	1.210E-04	-1.547E-03	3.314E-05	-4.067E-04
-1.920E-03	3.560E-03	-6.581E-04	3.814E-03	-1.547E-03	2.007E-02	-4.200E-04	5.176E-03
-1.689E-04	3.326E-04	-1.017E-05	9.128E-05	3.314E-05	-4.200E-04	6.487E-05	-8.111E-04
2.145E-03	-4.382E-03	2.066E-04	-1.790E-03	-4.067E-04	5.176E-03	-8.111E-04	1.037E-02
Red Pine (<i>Pinus resinosa</i>)							
1.183E-05	-6.486E-05	5.122E-07	-8.484E-06	-5.359E-07	5.049E-05	-1.214E-06	1.819E-05
-6.486E-05	3.678E-04	-2.888E-06	5.370E-05	2.971E-06	-2.890E-04	6.709E-06	-1.064E-04
5.122E-07	-2.888E-06	3.337E-07	-5.656E-06	1.410E-08	-1.175E-06	2.465E-08	-3.555E-07
-8.484E-06	5.370E-05	-5.656E-06	1.045E-04	-1.619E-07	1.006E-05	-3.794E-07	3.797E-06
-5.359E-07	2.971E-06	1.410E-08	-1.619E-07	2.351E-07	-2.182E-05	4.169E-07	-6.068E-06
5.049E-05	-2.890E-04	-1.175E-06	1.006E-05	-2.182E-05	2.069E-03	-3.885E-05	5.833E-04
-1.214E-06	6.709E-06	2.465E-08	-3.794E-07	4.169E-07	-3.885E-05	1.866E-06	-2.706E-05
1.819E-05	-1.064E-04	-3.555E-07	3.797E-06	-6.068E-06	5.833E-04	-2.706E-05	4.143E-04
Red Spruce (<i>Picea rubens</i>)							
4.583E-05	-1.421E-04	6.638E-06	-9.727E-05	6.650E-08	-4.196E-05	1.210E-06	-5.798E-05
-1.421E-04	4.513E-04	-2.148E-05	3.252E-04	-2.240E-07	1.464E-04	-3.731E-06	1.844E-04
6.638E-06	-2.148E-05	2.824E-06	-4.226E-05	2.400E-09	-3.020E-06	9.673E-08	-5.760E-06
-9.727E-05	3.252E-04	-4.226E-05	6.579E-04	-7.391E-08	7.357E-05	-1.611E-06	1.001E-04
6.650E-08	-2.240E-07	2.400E-09	-7.391E-08	1.096E-08	-6.145E-06	1.092E-07	-4.937E-06
-4.196E-05	1.464E-04	-3.020E-06	7.357E-05	-6.145E-06	3.497E-03	-6.143E-05	2.824E-03
1.210E-06	-3.731E-06	9.673E-08	-1.611E-06	1.092E-07	-6.143E-05	1.422E-06	-6.448E-05
-5.798E-05	1.844E-04	-5.760E-06	1.001E-04	-4.937E-06	2.824E-03	-6.448E-05	2.970E-03
White Pine (<i>Pinus strobus</i>)							
1.666E-04	-4.477E-04	1.646E-05	-2.360E-04	-3.211E-06	1.564E-04	-1.224E-07	-2.238E-06
-4.477E-04	1.224E-03	-4.485E-05	6.595E-04	8.703E-06	-4.357E-04	4.687E-07	-1.760E-06
1.646E-05	-4.485E-05	2.524E-06	-3.685E-05	-5.353E-07	2.598E-05	1.839E-07	-2.062E-06
-2.360E-04	6.595E-04	-3.685E-05	5.609E-04	7.514E-06	-3.779E-04	-2.504E-06	2.268E-05
-3.211E-06	8.703E-06	-5.353E-07	7.514E-06	6.180E-07	-2.955E-05	1.636E-07	-1.172E-06
1.564E-04	-4.357E-04	2.598E-05	-3.779E-04	-2.955E-05	1.452E-03	-8.000E-06	7.041E-05
-1.224E-07	4.687E-07	1.839E-07	-2.504E-06	1.636E-07	-8.000E-06	5.614E-06	-5.647E-05
-2.238E-06	-1.760E-06	-2.062E-06	2.268E-05	-1.172E-06	7.041E-05	-5.647E-05	6.143E-04
Bigtooth Aspen (<i>Populus grandidentata</i>)							
2.082E-04	-6.920E-04	3.728E-05	-3.983E-04	3.988E-07	-1.680E-05	-1.154E-06	4.918E-05
-6.920E-04	2.334E-03	-1.253E-04	1.360E-03	-1.078E-06	3.046E-05	3.737E-06	-1.692E-04
3.728E-05	-1.253E-04	1.244E-05	-1.350E-04	2.687E-07	-1.821E-05	2.574E-07	-9.669E-06
-3.983E-04	1.360E-03	-1.350E-04	1.508E-03	-2.967E-06	2.027E-04	-2.841E-06	1.009E-04

<i>a_{stem}</i>	<i>b_{stem}</i>	<i>a_{bark}</i>	<i>b_{bark}</i>	<i>a_{branches}</i>	<i>b_{branches}</i>	<i>a_{foliage}</i>	<i>b_{foliage}</i>
3.988E-07	-1.078E-06	2.687E-07	-2.967E-06	3.603E-07	-2.550E-05	2.392E-07	-1.015E-05
-1.680E-05	3.046E-05	-1.821E-05	2.027E-04	-2.550E-05	1.890E-03	-1.729E-05	7.776E-04
-1.154E-06	3.737E-06	2.574E-07	-2.841E-06	2.392E-07	-1.729E-05	2.438E-06	-9.718E-05
4.918E-05	-1.692E-04	-9.669E-06	1.009E-04	-1.015E-05	7.776E-04	-9.718E-05	3.961E-03
Pin Cherry (<i>Prunus pensylvanica</i>)							
2.135E-06	-7.349E-06	8.980E-08	-1.051E-06	1.784E-07	-3.364E-06	9.006E-08	-8.134E-07
-7.349E-06	2.660E-05	-2.913E-07	3.795E-06	-6.489E-07	1.318E-05	-3.478E-07	3.583E-06
8.980E-08	-2.913E-07	8.311E-07	-1.047E-05	-8.198E-08	1.482E-06	-4.864E-08	3.256E-07
-1.051E-06	3.795E-06	-1.047E-05	1.356E-04	1.064E-06	-2.055E-05	6.218E-07	-4.617E-06
1.784E-07	-6.489E-07	-8.198E-08	1.064E-06	4.479E-07	-8.481E-06	2.968E-07	-2.527E-06
-3.364E-06	1.318E-05	1.482E-06	-2.055E-05	-8.481E-06	1.719E-04	-5.754E-06	5.548E-05
9.006E-08	-3.478E-07	-4.864E-08	6.218E-07	2.968E-07	-5.754E-06	8.504E-07	-7.060E-06
-8.134E-07	3.583E-06	3.256E-07	-4.617E-06	-2.527E-06	5.548E-05	-7.060E-06	6.882E-05
Black Cherry (<i>Prunus serotina</i>)							
3.281E-03	-2.666E-03	1.678E-04	-8.289E-04	1.879E-04	-6.528E-04	4.688E-06	-1.439E-05
-2.666E-03	2.227E-03	-1.558E-04	7.885E-04	-1.542E-04	5.830E-04	-3.854E-06	2.597E-05
1.678E-04	-1.558E-04	4.158E-04	-1.789E-03	2.122E-05	-6.085E-05	-1.201E-06	8.786E-06
-8.289E-04	7.885E-04	-1.789E-03	7.793E-03	-9.469E-05	2.865E-04	5.401E-06	-5.102E-05
1.879E-04	-1.542E-04	2.122E-05	-9.469E-05	1.176E-04	-4.816E-04	1.129E-05	-4.642E-05
-6.528E-04	5.830E-04	-6.085E-05	2.865E-04	-4.816E-04	2.330E-03	-4.432E-05	2.944E-04
4.688E-06	-3.854E-06	-1.201E-06	5.401E-06	1.129E-05	-4.432E-05	2.842E-05	-1.014E-04
-1.439E-05	2.597E-05	8.786E-06	-5.102E-05	-4.642E-05	2.944E-04	-1.014E-04	6.129E-04
White Oak (<i>Quercus alba</i>)							
6.651E-05	-2.898E-04	-1.676E-06	1.223E-05	9.274E-06	-2.508E-04	1.512E-05	-2.516E-04
-2.898E-04	1.395E-03	6.913E-06	-3.447E-05	-3.804E-05	1.114E-03	-6.254E-05	1.128E-03
-1.676E-06	6.913E-06	3.598E-06	-3.192E-05	-1.327E-06	3.211E-05	-8.859E-07	1.349E-05
1.223E-05	-3.447E-05	-3.192E-05	3.266E-04	1.079E-05	-2.475E-04	8.813E-06	-1.285E-04
9.274E-06	-3.804E-05	-1.327E-06	1.079E-05	5.288E-06	-1.313E-04	5.527E-06	-8.322E-05
-2.508E-04	1.114E-03	3.211E-05	-2.475E-04	-1.313E-04	3.403E-03	-1.380E-04	2.184E-03
1.512E-05	-6.254E-05	-8.859E-07	8.813E-06	5.527E-06	-1.380E-04	1.717E-05	-2.708E-04
-2.516E-04	1.128E-03	1.349E-05	-1.285E-04	-8.322E-05	2.184E-03	-2.708E-04	4.501E-03
Red Oak (<i>Quercus rubra</i>)							
2.221E-04	-3.871E-04	3.899E-05	-3.055E-04	2.459E-06	-9.731E-05	-5.319E-06	3.581E-05
-3.871E-04	7.113E-04	-6.636E-05	5.411E-04	-5.084E-06	2.087E-04	7.969E-06	-3.776E-05
3.899E-05	-6.636E-05	2.582E-05	-2.012E-04	2.637E-06	-9.387E-05	-3.782E-06	3.165E-05
-3.055E-04	5.411E-04	-2.012E-04	1.626E-03	-2.236E-05	7.978E-04	2.847E-05	-2.309E-04
2.459E-06	-5.084E-06	2.637E-06	-2.236E-05	7.959E-06	-2.743E-04	1.760E-06	-1.211E-05
-9.731E-05	2.087E-04	-9.387E-05	7.978E-04	-2.743E-04	9.593E-03	-6.146E-05	4.672E-04
-5.319E-06	7.969E-06	-3.782E-06	2.847E-05	1.760E-06	-6.146E-05	8.382E-06	-7.140E-05
3.581E-05	-3.776E-05	3.165E-05	-2.309E-04	-1.211E-05	4.672E-04	-7.140E-05	7.201E-04
Black Oak (<i>Quercus velutina</i>)							
2.135E-06	-7.349E-06	8.980E-08	-1.051E-06	1.784E-07	-3.364E-06	9.006E-08	-8.134E-07
-7.349E-06	2.660E-05	-2.913E-07	3.795E-06	-6.489E-07	1.318E-05	-3.478E-07	3.583E-06
8.980E-08	-2.913E-07	8.311E-07	-1.047E-05	-8.198E-08	1.482E-06	-4.864E-08	3.256E-07
-1.051E-06	3.795E-06	-1.047E-05	1.356E-04	1.064E-06	-2.055E-05	6.218E-07	-4.617E-06
1.784E-07	-6.489E-07	-8.198E-08	1.064E-06	4.479E-07	-8.481E-06	2.968E-07	-2.527E-06
-3.364E-06	1.318E-05	1.482E-06	-2.055E-05	-8.481E-06	1.719E-04	-5.754E-06	5.548E-05
9.006E-08	-3.478E-07	-4.864E-08	6.218E-07	2.968E-07	-5.754E-06	8.504E-07	-7.060E-06
-8.134E-07	3.583E-06	3.256E-07	-4.617E-06	-2.527E-06	5.548E-05	-7.060E-06	6.882E-05
American Basswood (<i>Tilia americana</i>)							
2.010E-05	-1.002E-04	9.015E-06	-8.615E-05	6.643E-06	-8.049E-05	1.416E-05	-1.312E-04
-1.002E-04	5.139E-04	-4.752E-05	4.687E-04	-3.233E-05	4.001E-04	-6.675E-05	6.151E-04
9.015E-06	-4.752E-05	4.688E-05	-4.423E-04	9.820E-06	-1.228E-04	1.482E-05	-1.423E-04
-8.615E-05	4.687E-04	-4.423E-04	4.255E-03	-9.304E-05	1.192E-03	-1.379E-04	1.353E-03

<i>a_{stem}</i>	<i>b_{stem}</i>	<i>a_{bark}</i>	<i>b_{bark}</i>	<i>a_{branches}</i>	<i>b_{branches}</i>	<i>a_{foliage}</i>	<i>b_{foliage}</i>
6.643E-06	-3.233E-05	9.820E-06	-9.304E-05	4.790E-05	-6.067E-04	4.352E-05	-4.374E-04
-8.049E-05	4.001E-04	-1.228E-04	1.192E-03	-6.067E-04	7.889E-03	-5.481E-04	5.634E-03
1.416E-05	-6.675E-05	1.482E-05	-1.379E-04	4.352E-05	-5.481E-04	8.970E-05	-8.990E-04
-1.312E-04	6.151E-04	-1.423E-04	1.353E-03	-4.374E-04	5.634E-03	-8.990E-04	9.333E-03
Eastern Hemlock (<i>Tsuga canadensis</i>)							
9.201E-06	-4.413E-05	1.323E-06	-2.697E-05	1.157E-06	-1.643E-05	3.031E-07	-7.408E-07
-4.413E-05	2.251E-04	-5.954E-06	1.260E-04	-6.121E-06	8.989E-05	-9.222E-07	2.419E-06
1.323E-06	-5.954E-06	9.866E-07	-2.045E-05	-6.512E-07	8.271E-06	1.080E-06	-3.897E-06
-2.697E-05	1.260E-04	-2.045E-05	4.421E-04	1.226E-05	-1.539E-04	-2.118E-05	7.889E-05
1.157E-06	-6.121E-06	-6.512E-07	1.226E-05	9.355E-06	-1.265E-04	-3.231E-06	1.240E-05
-1.643E-05	8.989E-05	8.271E-06	-1.539E-04	-1.265E-04	1.768E-03	3.941E-05	-1.307E-04
3.031E-07	-9.222E-07	1.080E-06	-2.118E-05	-3.231E-06	3.941E-05	4.764E-05	-1.932E-04
-7.408E-07	2.419E-06	-3.897E-06	7.889E-05	1.240E-05	-1.307E-04	-1.932E-04	8.551E-04
Balsam Fir (<i>Abies balsamea</i>)							
2.770E-06	-1.676E-05	4.138E-07	-1.165E-05	-5.923E-08	2.384E-06	-1.264E-06	4.521E-06
-1.676E-05	1.064E-04	-2.520E-06	7.490E-05	3.197E-07	-1.530E-05	7.659E-06	-3.034E-05
4.138E-07	-2.520E-06	1.823E-07	-5.269E-06	-1.275E-08	7.027E-07	-4.478E-07	1.771E-06
-1.165E-05	7.490E-05	-5.269E-06	1.614E-04	5.010E-07	-3.085E-05	1.333E-05	-5.782E-05
-5.923E-08	3.197E-07	-1.275E-08	5.010E-07	5.350E-07	-2.502E-05	1.450E-06	-6.109E-06
2.384E-06	-1.530E-05	7.027E-07	-3.085E-05	-2.502E-05	1.207E-03	-7.092E-05	3.203E-04
-1.264E-06	7.659E-06	-4.478E-07	1.333E-05	1.450E-06	-7.092E-05	2.109E-05	-8.706E-05
4.521E-06	-3.034E-05	1.771E-06	-5.782E-05	-6.109E-06	3.203E-04	-8.706E-05	3.965E-04
Mountain Maple (<i>Acer spicatum</i>)							
2.135E-06	-7.349E-06	8.980E-08	-1.051E-06	1.784E-07	-3.364E-06	9.006E-08	-8.134E-07
-7.349E-06	2.660E-05	-2.913E-07	3.795E-06	-6.489E-07	1.318E-05	-3.478E-07	3.583E-06
8.980E-08	-2.913E-07	8.311E-07	-1.047E-05	-8.198E-08	1.482E-06	-4.864E-08	3.256E-07
-1.051E-06	3.795E-06	-1.047E-05	1.356E-04	1.064E-06	-2.055E-05	6.218E-07	-4.617E-06
1.784E-07	-6.489E-07	-8.198E-08	1.064E-06	4.479E-07	-8.481E-06	2.968E-07	-2.527E-06
-3.364E-06	1.318E-05	1.482E-06	-2.055E-05	-8.481E-06	1.719E-04	-5.754E-06	5.548E-05
9.006E-08	-3.478E-07	-4.864E-08	6.218E-07	2.968E-07	-5.754E-06	8.504E-07	-7.060E-06
-8.134E-07	3.583E-06	3.256E-07	-4.617E-06	-2.527E-06	5.548E-05	-7.060E-06	6.882E-05
Black Ash (<i>Fraxinus nigra</i>)							
5.260E-05	-1.730E-04	1.971E-05	-1.874E-04	8.975E-06	-7.307E-05	-1.376E-05	6.662E-05
-1.730E-04	6.014E-04	-6.428E-05	6.429E-04	-2.261E-05	1.608E-04	4.374E-05	-2.205E-04
1.971E-05	-6.428E-05	2.494E-05	-2.332E-04	-5.868E-06	3.165E-05	-1.891E-05	9.922E-05
-1.874E-04	6.429E-04	-2.332E-04	2.251E-03	6.212E-05	-3.962E-04	1.743E-04	-9.311E-04
8.975E-06	-2.261E-05	-5.868E-06	6.212E-05	7.393E-05	-6.308E-04	4.834E-05	-3.402E-04
-7.307E-05	1.608E-04	3.165E-05	-3.962E-04	-6.308E-04	6.094E-03	-3.706E-04	3.089E-03
-1.376E-05	4.374E-05	-1.891E-05	1.743E-04	4.834E-05	-3.706E-04	7.685E-05	-4.617E-04
6.662E-05	-2.205E-04	9.922E-05	-9.311E-04	-3.402E-04	3.089E-03	-4.617E-04	3.160E-03
Green Ash (<i>Fraxinus pennsylvanica</i>)							
2.135E-06	-7.349E-06	8.980E-08	-1.051E-06	1.784E-07	-3.364E-06	9.006E-08	-8.134E-07
-7.349E-06	2.660E-05	-2.913E-07	3.795E-06	-6.489E-07	1.318E-05	-3.478E-07	3.583E-06
8.980E-08	-2.913E-07	8.311E-07	-1.047E-05	-8.198E-08	1.482E-06	-4.864E-08	3.256E-07
-1.051E-06	3.795E-06	-1.047E-05	1.356E-04	1.064E-06	-2.055E-05	6.218E-07	-4.617E-06
1.784E-07	-6.489E-07	-8.198E-08	1.064E-06	4.479E-07	-8.481E-06	2.968E-07	-2.527E-06
-3.364E-06	1.318E-05	1.482E-06	-2.055E-05	-8.481E-06	1.719E-04	-5.754E-06	5.548E-05
9.006E-08	-3.478E-07	-4.864E-08	6.218E-07	2.968E-07	-5.754E-06	8.504E-07	-7.060E-06
-8.134E-07	3.583E-06	3.256E-07	-4.617E-06	-2.527E-06	5.548E-05	-7.060E-06	6.882E-05
Norway Spruce (<i>Picea abies</i>)							
1.469E-06	-7.244E-06	1.979E-07	-4.146E-06	-2.037E-07	4.118E-06	-8.240E-07	3.154E-06
-7.244E-06	3.712E-05	-1.001E-06	2.189E-05	9.941E-07	-2.090E-05	4.163E-06	-1.681E-05
1.979E-07	-1.001E-06	5.902E-08	-1.277E-06	-1.357E-08	3.081E-07	-8.983E-08	3.716E-07
-4.146E-06	2.189E-05	-1.277E-06	2.978E-05	3.285E-07	-7.916E-06	2.085E-06	-9.204E-06

<i>a_{stem}</i>	<i>b_{stem}</i>	<i>a_{bark}</i>	<i>b_{bark}</i>	<i>a_{branches}</i>	<i>b_{branches}</i>	<i>a_{foliage}</i>	<i>b_{foliage}</i>
-2.037E-07	9.941E-07	-1.357E-08	3.285E-07	2.601E-07	-5.373E-06	8.240E-07	-3.220E-06
4.118E-06	-2.090E-05	3.081E-07	-7.916E-06	-5.373E-06	1.162E-04	-1.757E-05	7.250E-05
-8.240E-07	4.163E-06	-8.983E-08	2.085E-06	8.240E-07	-1.757E-05	6.042E-06	-2.382E-05
3.154E-06	-1.681E-05	3.716E-07	-9.204E-06	-3.220E-06	7.250E-05	-2.382E-05	1.004E-04
<i>Black Spruce (Picea mariana)</i>							
1.058E-06	-7.598E-06	3.650E-07	-8.240E-06	-5.623E-07	6.855E-06	-1.396E-06	2.916E-06
-7.598E-06	5.665E-05	-2.638E-06	6.105E-05	4.016E-06	-5.082E-05	1.055E-05	-2.369E-05
3.650E-07	-2.638E-06	3.661E-07	-8.324E-06	-2.377E-07	2.970E-06	-1.121E-06	2.432E-06
-8.240E-06	6.105E-05	-8.324E-06	1.932E-04	5.450E-06	-7.024E-05	2.658E-05	-6.065E-05
-5.623E-07	4.016E-06	-2.377E-07	5.450E-06	3.748E-06	-4.806E-05	1.116E-05	-2.511E-05
6.855E-06	-5.082E-05	2.970E-06	-7.024E-05	-4.806E-05	6.394E-04	-1.478E-04	3.494E-04
-1.396E-06	1.055E-05	-1.121E-06	2.658E-05	1.116E-05	-1.478E-04	7.748E-05	-1.778E-04
2.916E-06	-2.369E-05	2.432E-06	-6.065E-05	-2.511E-05	3.494E-04	-1.778E-04	4.336E-04
<i>Balsam Poplar (Populus balsamea)</i>							
1.065E-05	-6.500E-05	2.924E-06	-2.988E-05	1.987E-06	-5.176E-05	7.035E-07	-8.801E-06
-6.500E-05	4.221E-04	-1.691E-05	1.794E-04	-1.200E-05	3.286E-04	-4.574E-06	6.389E-05
2.924E-06	-1.691E-05	1.195E-05	-1.270E-04	1.018E-06	-2.531E-05	2.995E-07	-1.457E-06
-2.988E-05	1.794E-04	-1.270E-04	1.403E-03	-1.049E-05	2.528E-04	-2.293E-06	-4.896E-06
1.987E-06	-1.200E-05	1.018E-06	-1.049E-05	1.869E-06	-4.956E-05	7.482E-08	-1.873E-06
-5.176E-05	3.286E-04	-2.531E-05	2.528E-04	-4.956E-05	1.417E-03	-4.335E-06	1.254E-04
7.035E-07	-4.574E-06	2.995E-07	-2.293E-06	7.482E-08	-4.335E-06	3.242E-06	-3.997E-05
-8.801E-06	6.389E-05	-1.457E-06	-4.896E-06	-1.873E-06	1.254E-04	-3.997E-05	5.782E-04
<i>Trembling Aspen (Populus tremuloides)</i>							
8.529E-06	-4.453E-05	1.417E-06	-2.698E-05	-8.178E-08	2.689E-06	3.877E-07	-5.444E-06
-4.453E-05	2.388E-04	-7.485E-06	1.472E-04	3.612E-07	-1.105E-05	-2.200E-06	3.418E-05
1.417E-06	-7.485E-06	5.378E-07	-1.024E-05	7.365E-08	-2.981E-06	-1.144E-07	1.003E-06
-2.698E-05	1.472E-04	-1.024E-05	2.036E-04	-1.414E-06	5.876E-05	1.770E-06	-1.120E-05
-8.178E-08	3.612E-07	7.365E-08	-1.414E-06	6.781E-07	-2.697E-05	2.168E-06	-2.644E-05
2.689E-06	-1.105E-05	-2.981E-06	5.876E-05	-2.697E-05	1.110E-03	-8.630E-05	1.088E-03
3.877E-07	-2.200E-06	-1.144E-07	1.770E-06	2.168E-06	-8.630E-05	1.279E-05	-1.542E-04
-5.444E-06	3.418E-05	1.003E-06	-1.120E-05	-2.644E-05	1.088E-03	-1.542E-04	1.926E-03
<i>Northern White Cedar (Thuja occidentalis)</i>							
1.362E-05	-6.174E-05	2.673E-06	-7.094E-05	2.269E-07	-2.420E-06	-5.231E-07	1.205E-06
-6.174E-05	2.907E-04	-1.235E-05	3.382E-04	-9.409E-07	1.077E-05	1.220E-06	6.506E-06
2.673E-06	-1.235E-05	1.657E-06	-4.424E-05	-1.399E-07	1.333E-06	1.860E-07	-1.281E-06
-7.094E-05	3.382E-04	-4.424E-05	1.214E-03	3.601E-06	-3.263E-05	-6.063E-06	4.955E-05
2.269E-07	-9.409E-07	-1.399E-07	3.601E-06	1.132E-05	-1.075E-04	9.641E-06	-6.325E-05
-2.420E-06	1.077E-05	1.333E-06	-3.263E-05	-1.075E-04	1.095E-03	-9.399E-05	6.739E-04
-5.231E-07	1.220E-06	1.860E-07	-6.063E-06	9.641E-06	-9.399E-05	1.434E-05	-9.554E-05
1.205E-06	6.506E-06	-1.281E-06	4.955E-05	-6.325E-05	6.739E-04	-9.554E-05	7.038E-04
<i>American Elm (Ulmus americana)</i>							
1.628E-05	-1.147E-04	2.379E-06	-9.407E-05	4.998E-06	-4.313E-05	-2.500E-06	1.234E-05
-1.147E-04	8.200E-04	-1.686E-05	6.751E-04	-3.528E-05	2.988E-04	1.775E-05	-9.640E-05
2.379E-06	-1.686E-05	5.519E-06	-2.273E-04	3.328E-07	-5.459E-06	-4.004E-07	1.864E-06
-9.407E-05	6.751E-04	-2.273E-04	9.495E-03	-1.284E-05	2.151E-04	1.580E-05	-6.110E-05
4.998E-06	-3.528E-05	3.328E-07	-1.284E-05	1.778E-05	-1.432E-04	-2.239E-07	9.186E-06
-4.313E-05	2.988E-04	-5.459E-06	2.151E-04	-1.432E-04	1.265E-03	4.644E-06	-8.636E-05
-2.500E-06	1.775E-05	-4.004E-07	1.580E-05	-2.239E-07	4.644E-06	9.309E-06	-3.864E-05
1.234E-05	-9.640E-05	1.864E-06	-6.110E-05	9.186E-06	-8.636E-05	-3.864E-05	2.681E-04
<i>Unidentified</i>							
2.135E-06	-7.349E-06	8.980E-08	-1.051E-06	1.784E-07	-3.364E-06	9.006E-08	-8.134E-07
-7.349E-06	2.660E-05	-2.913E-07	3.795E-06	-6.489E-07	1.318E-05	-3.478E-07	3.583E-06
8.980E-08	-2.913E-07	8.311E-07	-1.047E-05	-8.198E-08	1.482E-06	-4.864E-08	3.256E-07
-1.051E-06	3.795E-06	-1.047E-05	1.356E-04	1.064E-06	-2.055E-05	6.218E-07	-4.617E-06

<i>a_{stem}</i>	<i>b_{stem}</i>	<i>a_{bark}</i>	<i>b_{bark}</i>	<i>a_{branches}</i>	<i>b_{branches}</i>	<i>a_{foliage}</i>	<i>b_{foliage}</i>
1.784E-07	-6.489E-07	-8.198E-08	1.064E-06	4.479E-07	-8.481E-06	2.968E-07	-2.527E-06
-3.364E-06	1.318E-05	1.482E-06	-2.055E-05	-8.481E-06	1.719E-04	-5.754E-06	5.548E-05
9.006E-08	-3.478E-07	-4.864E-08	6.218E-07	2.968E-07	-5.754E-06	8.504E-07	-7.060E-06
-8.134E-07	3.583E-06	3.256E-07	-4.617E-06	-2.527E-06	5.548E-05	-7.060E-06	6.882E-05

BIBLIOGRAPHY

- [1] Ahmed, R., Siqueira, P., Hensley, S., Chapman, B., and Bergen, K. A Survey of Temporal Decorrelation from Spaceborne L-band Repeat-pass InSAR. *Remote Sensing of Environment* (in print).
- [2] Anderson, J., Hardy, E., Roach, J., and Witmer, R. A land use and land cover classification system for use with remote sensor data. *U.S. Geological Survey Prof. Paper*, 964 (1976), 28.
- [3] Anthes, R. A. *Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond*, National Research Council Committee on Earth Science and Applications from Space. National Academic Press, 2007.
- [4] Askne, J., Dammert, P., Ulander, L., and Smith, G. C-band repeat-pass interferometric SAR observations of the forest. *Geoscience and Remote Sensing, IEEE Transactions on* 35, 1 (Jan 1997), 25–35.
- [5] Askne, J., Santoro, M., Smith, G., and Fransson, J. E. S. Multitemporal repeat-pass SAR interferometry of boreal forests. *IEEE Transactions on Geoscience and Remote Sensing* 41 (July 2003), 1540–1550.
- [6] Attema, E. P. W., and Ulaby, F. T. Vegetation modeled as a water cloud. *Radio Science* 13 (1978), 357–364.
- [7] Avery, T E., and Burkhardt, H. E. *Forest Measurements*, fifth ed. McGraw Hill, 2002.
- [8] Baskerville, G.L. Use of logarithmic regression in the estimation of plant biomass. *Canadian Journal of Forest Research* 2, 49 (1972).
- [9] Blair, J. B., Rabine, D. L., and Hofton, M. A. The laser vegetation imaging sensor: a medium-altitude, digitisation-only, airborne laser altimeter for mapping vegetation and topography. *ISPRS Journal of Photogrammetry and Remote Sensing* 54 (July 1999), 115–122(8).
- [10] Brenneman, B.B., Fredrick, D.J., Gardner, W.E., Schoenhoefen, L.H., and Marsh, P.L. Biomass fo species and stand of West Virginia hardwoods. In *Central Hardwood Forest Conference II, West Lafayette* (1978), Purdue University, pp. 159–178.
- [11] Buonaccorsi, J. P. *Measurement error, models, methods and applications*. Chapman and Hall/CRC interdisciplinary statistics series, 2010.

- [12] Castel, T., Beaudoin, A., Stach, N., Stussi, N., Toan, T. Le, and Durand, P. Sensitivity of space-borne SAR data to forest parameters over sloping terrain. Theory and experiment. *International Journal of Remote Sensing* 22, 12 (2001), 2351–2376.
- [13] Chave, Jerome, Condit, Richard, Aguilar, Salomon, Hernandez, Andres, Lao, Suzanne, and Perez, Rolando. Error propagation and scaling for tropical forest biomass estimates. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 359, 1443 (2004), 409–420.
- [14] Cloude, S. R. Polarization coherence tomography. *Radio Science* 41 (Aug. 2006), 4017–+.
- [15] Cloude, S. R., and Papathanassiou, K. P. Polarimetric optimization in radar interferometry. *Electronics Letters* 33, 13 (1997), 1176–1178.
- [16] Cloude, S. R., and Papathanassiou, K. P. Polarimetric SAR interferometry. *IEEE Trans. Geosci. Remote Sensing* 36, 5 (1998), 1551–1565.
- [17] Cloude, S. R., and Papathanassiou, K. P. Three-stage inversion process for polarimetric SAR interferometry. In *IEE Proc. Radar Sonar Navigation* (2003), vol. 150, pp. 125–134.
- [18] Cloude, S. R., and Pottier, E. A review of target decomposition theorems in radar polarimetry. *IEEE Transactions on Geoscience and Remote Sensing* 34 (Mar. 1996), 498–518.
- [19] Cloude, Shane. *Polarisation: Applications in Remote Sensing*. Oxford, 2009.
- [20] Cloude, S.R. Dual-baseline coherence tomography. *Geoscience and Remote Sensing Letters, IEEE* 4, 1 (Jan. 2007), 127–131.
- [21] Colin, E., Titin-Schnaider, C., and Tabbara, W. Investigation on Different Interferometric Coherence Optimization Methods. In *Applications of SAR Polarimetry and Polarimetric Interferometry* (Apr. 2003), vol. 529 of *ESA Special Publication*.
- [22] Colin, E., Titin-Schnaider, C., and Tabbara, W. An interferometric coherence optimization method in radar polarimetry for high-resolution imagery. *IEEE Trans. Geosci. Remote Sensing* 44, 1 (2006), 167–175.
- [23] Cook, J. R., and Stefanski, L. A. Simulation-extrapolation estimation in parametric measurement error models. *Journal of the American Statistical Association* 89, 428 (December 1994), 1314–1328.
- [24] Crow, T.R., and Erdmann, G.G. Weight and volume equations and tables for red maple in the lake states. *Research Paper, USDA, Forest Service, North Central Forest Experiment Station NC-242* (1983).

- [25] Crow, T.R., and Schlaegel, B.E. A guide to using regression equations for estimating tree biomass. *North. J. Appl. For.* 5 (1988), 15–22.
- [26] Curlander, J.C., and McDonough, R.N. *Synthetic aperture radar: systems and signal processing*. Wiley series in remote sensing. Wiley, 1991.
- [27] Dobson, M. C., Ulaby, F. T., adn T. L. Sharik, L. E. Pierce, Bergen, K. M., Kellendorfer, J., Kendra, J. R., Li, E., Lin, Y. C., Nashashibi, A., Sarabandi, K., and Siqueira, P. Estimation of forest biophysical characteristics in northern Michigan with SIR-C/X-SAR. *IEEE Trans. Geosci. Remote Sensing* 33, 4 (1995), 877–895.
- [28] Donnellan, A., Rosen, P., Ranson, J., and Zebker, H. Deformation, ecosystem strcutre and dynamics of ice (DESDynI). In *Int. Geoscience and Remote Sensing Symposium* (July 2008), vol. III, IEEE, pp. 3–8.
- [29] Drake, J., Dubayah, R., Knox, R., Clark, D., and Blair, J. B. Sensitivity of large-footprint lidar to canopy structure and biomass in a neotropical rainforest. *Remote Sensing of Environment* 81 (2002), 378–392.
- [30] Drake, Jason B., Dubayah, Ralph O., Clark, David B., Knox, Robert G., Blair, J. Bryan, Hofton, Michelle A., Chazdon, Robin L., Weishampel, John F., and Prince, Steve. Estimation of tropical forest structural characteristics using large-footprint lidar. *Remote Sensing of Environment* 79, 2-3 (2002), 305 – 319.
- [31] Drake, Jason B., Knox, Robert G., Dubayah, Ralph O., Clark, David B., Condit, Richard, Blair, J. Bryan, and Hofton, Michelle. Above-ground biomass estimation in closed canopy neotropical forests using lidar remote sensing: factors affecting the generality of relationships. *Global Ecology and Biogeography* 12 (2003).
- [32] Draper, Norman R., and Smith, Harry. *Applied Regression Analysis (Wiley Series in Probability and Statistics)*, third ed. Wiley-Interscience, Apr. 1998.
- [33] Durden, S. L., Zyl, J. J. Van, and Zebker, H. A. Modeling and observation of radar polarization signature of forested areas. *IEEE Trans. Geosci. Remote Sensing* 27, 3 (1989), 290–301.
- [34] Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Osokin, M., Burbank, D., and Alsdorf, D. The shuttle radar topography mission. *Review of Geophysics* 45 (2007).
- [35] Fisher, R. T. The management of the Harvard Forest. *Harvard Forest Bulletin* 1 (1921), 1–55.

- [36] Flynn, T., Tabb, M., and Carande, R. Coherence region shape extraction for vegetation parameter estimation in polarimetric sar interferometry. In *Geoscience and Remote Sensing Symposium, 2002. IGARSS '02. 2002 IEEE International* (2002), vol. 5, pp. 2596 – 2598 vol.5.
- [37] Fornaro, G., Lombardini, F., and Serafino, F. Three-dimensional multipass SAR focusing: experiments with long-term spaceborne data. *Geoscience and Remote Sensing, IEEE Transactions on* 43, 4 (April 2005), 702–714.
- [38] Fornaro, G., Serafino, F., and Soldovieri, F. Three-dimensional focusing with multipass SAR data. *Geoscience and Remote Sensing, IEEE Transactions on* 41, 3 (March 2003), 507–517.
- [39] Freedman, B., Duinker, P.N., Barclay, H., Morash, R., and Prager, U. Forest biomass and nutrient studies in central nova scotia. *Information Report, Maritimes Forest Research Centre, Canada M-X-134* (1982).
- [40] Freeman, A., Alves, M., Chapman, B., Cruz, J., Kim, Y., Shaffer, S., Sun, J., Turner, E., , and Sarabandi, K. SIR-C data quality and calibration results. *IEEE Transactions on Geoscience and Remote Sensing* 33, 4 (1995), 848–857.
- [41] Freeman, A., and Durden, S.L. A three-component scattering model for polarimetric sar data. *Geoscience and Remote Sensing, IEEE Transactions on* 36, 3 (may 1998), 963 –973.
- [42] Freeman, A., Rosen, P., Jordan, R., Johnson, W.T.K., Hensley, S., Sweestser, T., Loverro, A., Smith, J., Sprague, G., and Shen, Y. DESDynI - a NASA mission for ecosystems, solid earth and cryosphere science. In *Int. Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry* (January 2009), ESA, pp. 26–30.
- [43] Freeman, Anthony. SAR Calibration: An Overview. *IEEE Trans. Geosci. Remote Sensing* 30, 6 (1992).
- [44] Frieswyk, T. S., and DiGiovanni, D. M. Biomass statistics for Maryland 1986. *USDA Forest Service Northeastern Forest Experiment Station, Resource Bulletin NE-113* (1990).
- [45] Fuller, Wayne A. *Measurement Error Models*. John Wiley and Sons, Inc., 2008.
- [46] Gatelli, F., Guamieri, A.M., Parizzi, F., Pasquali, P., Prati, C., and Rocca, F. The wavenumber shift in sar interferometry. *Geoscience and Remote Sensing, IEEE Transactions on* 32, 4 (jul 1994), 855 –865.
- [47] Gertner, G., and Kohl, M. An assessment of some nonsampling errors in a national survey using an error budget. *Forest Science* 38 (1 August 1992), 525–538(14).

- [48] Gesch, D., Oimoen, M., Greenlee, S., Nelson, C., Steuck, M., and Tyler, D. The national elevation dataset. *Photogrammetric Engineering and Remote Sensing* 68 (2002), 5–11.
- [49] Gierull, C., and Sikaneta, I.C. Estimating the effective number of looks in interferometric SAR data. *IEEE Transactions on Geoscience and Remote Sensing* 40, 8 (2002), 1733–1742.
- [50] Graham, L. C. Synthetic interferometer radar for topographic mapping. *Proceedings of the IEEE* 62, 6 (June 1974), 763–768.
- [51] Guttman, N., and Quayle, R. Historical perspective of the U.S. climate divisions. *Bulletin of the American Meteorological Society* 77 (1996), 293–303.
- [52] Hagberg, J. O., Ulander, L. M. H., and Askne, J. Repeat-pass SAR interferometry over forested terrain. *IEEE Trans. Geosci. Remote Sensing* 33, 2 (1995).
- [53] Hajnsek, I., Lee, S. K., and Papathanassiou, K. P. Tropical-forest-parameter estimation by means fo PolInSAR: The INDREX-II campaign. *IEEE Trans. Geosci. Remote Sensing* 47, 2 (2009), 481–493.
- [54] Hensley, S., Chapin, E., and Michel, T. R. Improved processing of airsar data based on the geosar processor. In *AIRsar Earth Science and Applications Workshop* (March 2002).
- [55] Hensley, S., Wheeler, K., Sadowy, G., Jones, C., Shaffer, S., Zebker, H., Miller, T., Heavey, B., Chuang, E., Chao, R., Vines, K., Nishimoto, K., Prater, J., Carrico, B., Chamberlain, N., Shimada, J., Simard, M., Chapman, B., Muellerschoen, R., Le, C., Michel, T., Hamilton, G., Robison, D., Neumann, G., Meyer, R., Smith, P., Granger, J., Rosen, P., Flower, D., and Smith, R. The uavasar instrument: Description and first results. In *Radar Conference, 2008. RADAR '08. IEEE* (May 2008), pp. 1–6.
- [56] Hensley, Scott. Desdyni biomass error model and assosiated spreadsheet. *Technical Report, JPL, Interoffice Memorandum* (2010).
- [57] Hocker, H.W., and Early, D.J. Biomass and leaf area equations for northern forest species. *University of New Hampshire Research Report* 102, 27 (1983).
- [58] Homer, J., Longstaff, I.D., and Callaghan, G. High resolution 3-D SAR via multi-baseline interferometry. In *Geoscience and Remote Sensing Symposium, 1996. IGARSS '96. 'Remote Sensing for a Sustainable Future.', International* (May 1996), vol. 1, pp. 796–798 vol.1.
- [59] Honer, T.G. Weight relationships in open- and forest-grown balsam fir trees. *University of Maine at Orono, Misc. publication. IUFRO section 25*, 132 (1971), 65–78.

- [60] Hyde, P., Dubayah, R., Peterson, B., Blair, J. B., Hofton, M., Hunsaker, C., Knox, C., and Walker, W. Mapping forest structure for wildlife habitat analysis using waveform lidar: Validation of montane ecosystems. *Remote Sensing of Environment* 96 (2005), 427–437.
- [61] Hyde, P., Dubayah, R., Walker, W., Blair, B., Hofton, M., and Hunsaker, C. Mapping forest structure fro wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy. *Remote Sensing of Environment* 102 (Jan 2006), 63–73.
- [62] Hyde, P., Nelson, R., Kimes, D., and Levine, E. Exploring lidar-radar synergy—predicting aboveground biomass in a southwestern ponderosa pine forest using LiDAR, SAR and InSAR. *Remote Sensing of Environment* 106, 1 (2007), 28 – 38.
- [63] Imhoff, M. L. Radar backscatter and biomass saturation: ramifications for global biomass inventory. *IEEE Trans. Geosci. Remote Sensing* 33, 2 (1995).
- [64] Jenkins, J.C., Chojnacky, D.C., Heath, L.S., and Birdsey, R.A. Comprehensive database of diameter-based biomass regressions for north american tree species. *US Forest Service Northeastern experiment station research paper NE-319* (2004).
- [65] Jenkins, J.C., Chojnacky, D.C., Heath, L.S., and Birdsey, R.A. National-scale biomass estimators for united states tree species. *Forest Science* 49 (February 2003), 12–35(24).
- [66] Jokela, E. J., P., Van Gorp K., D., Briggs R., and H., White E. Biomass estimation equations for norway spruce in new york. *Canadian journal of forest research* 16, 2 (1986).
- [67] Jones, C., Hensley, S., and Michel, T. Topography-dependent motion compensation: Application to uavssar data. In *Radar Conference, 2009 IEEE* (may 2009), pp. 1 –6.
- [68] Kellendorfer, J. M., Walker, W. S., Pierce, L. E., Dobson, C., Fites, J. A., Hunsaker, C., Vona, J., and Clutter, M. Vegetation height estimation from shuttle radar topography mission and national elevation datasets. *Remote Sensing of the Environment* 93 (2004), 339–358.
- [69] Ker, M.F. Biomass equations for ten major species in Cumberland County, Nova Scotia. *Canadian Forest Service Maritime Forest Research Center, Report M-X-148* (1980).
- [70] Ker, M.F. Tree biomass equations for seven species in southwestern New Brunswick. *Canadian Forest Service Maritime Forest Research Center, Report M-X-114* (1980).

- [71] Lambert, M.C., Ung, C. H., and Raulier, F. Canadian national tree above-ground biomass equations. *Canadian Journal of Forestry Research* 35 (2005), 1996–2018.
- [72] Le Toan, T., Beaudoin, A., Riou, J., and Guyon, D. Relating forest biomass to SAR data. *IEEE Trans. Geosci. Remote Sensing* 30, 2 (1992), 403–411.
- [73] Lee, J.S., and Pottier, E. *Polarimetric Radar Imaging : From basics to applications*. Optical Science and Engineering. CRC Press, Taylor & Francis Editor,, Jan. 2009.
- [74] Lefsky, M. A., Cohen, W. B., Acker, S. A., Parker, G. G., and Harding, T. A. Spies D. Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. *Remote Sensing of Environment* 70 (1999), 339–361.
- [75] Lefsky, Michael A, Cohen, Warren B., Harding, David J., Parker, Geoffrey G., Acker, Steven A., and Gower, S. Thomas. Lidar remote sensing of above-ground biomass in three biomes. *Global Ecology and Biogeography* 11 (2002).
- [76] Lefsky, Michael A, Harding, D, Cohen, W.B, Parker, G, and Shugart, H.H. Surface lidar remote sensing of basal area and biomass in deciduous forests of eastern maryland, usa. *Remote Sensing of Environment* 67, 1 (1999), 83 – 98.
- [77] Lefsky, Michael A., Hudak, Andrew T., Cohen, Warren B., and Acker, S.A. Geographic variability in lidar predictions of forest stand structure in the pacific northwest. *Remote Sensing of Environment* 95, 4 (2005), 532 – 548.
- [78] Lefsky, Michael A., Hudak, Andrew T., Cohen, Warren B., and Acker, S.A. Patterns of covariance between forest stand and canopy structure in the pacific northwest. *Remote Sensing of Environment* 95, 4 (2005), 517 – 531.
- [79] Li, F. K., and Goldstein, R. M. Studies of multibaseline spaceborne interferometric synthetic aperture radars. *IEEE Transactions on Geoscience and Remote Sensing* 28 (1990), 88–98.
- [80] Lucas, R., Armston, J., Fairfax, R., Fensham, R., Accad, A., Carreiras, J., Kelley, J., Bunting, P., Clewley, D., Bray, S., Metcalfe, D., Dwyer, J., Bowen, M., Eyre, T., Laidlaw, M., and Shimada, M. An evaluation of the alos palsar l-band backscatter - above ground biomass relationship queensland, australia: Impacts of surface moisture condition and vegetation structure. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of* 3, 4 (dec. 2010), 576 –593.
- [81] Luckman, Adrian, Baker, John, Honzk, Miroslav, and Lucas, Richard. Tropical forest biomass density estimation using jers-1 sar: Seasonal variation, confidence limits, and application to image mosaics. *Remote Sensing of Environment* 63, 2 (1998), 126 – 139.

- [82] Luckman, Adrian, Baker, John, Kuplich, Tatiana Mora, da Costa Freitas Yanasse, Corina, and Frery, Alejandro C. A study of the relationship between radar backscatter and regenerating tropical forest biomass for spaceborne sar instruments. *Remote Sensing of Environment* 60, 1 (1997), 1 – 13.
- [83] MacArthur, Robert H., and Horn, Henry S. Foliage profile by vertical measurements. *Ecology* 50, 5 (1969), 802–804.
- [84] Mandelbrot, B. B. *The fractal geometry of nature*. W.H.Freeman, San Francisco, 1977.
- [85] Mann, H., and Whitney, D. On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Science* 18, 1 (1947), 50–60.
- [86] Means, Joseph E., Acker, Steven A., Harding, David J., Blair, J. Bryan, Lefsky, Michael A., Cohen, Warren B., Harmon, Mark E., and McKee, W. Arthur. Use of large-footprint scanning airborne lidar to estimate forest stand characteristics in the western cascades of oregon. *Remote Sensing of Environment* 67, 3 (1999), 298 – 308.
- [87] Mette, T., Papathanassiou, K. P., and Hajnesk, I. Forest biomass estimation using polarimetric SAR interferometry. In *Proceedings of IEEE Geoscience and Remote Sensing Symposium (IGARSS)* (June 2002), pp. 22–26.
- [88] Nelson, R., Krabill, W., and Mclean, G. Determining forest canpopy characteristics using airborne laser data. *Remote Sensing of Environment* 15 (1984), 201–212.
- [89] Neumann, M., Ferro-Famil, L., and Reigber, A. Multibaseline polarimetric SAR interferometry coherence optimization. *Geoscience and Remote Sensing Letters, IEEE* 5, 1 (Jan. 2008), 93–97.
- [90] Neumann, M., Ferro-Famil, L., and Reigber, A. Multibaseline Polarimetric SAR Interferometry Coherence Optimization. *IEEE Geoscience and Remote Sensing Letters* 5 (Jan. 2008), 93–97.
- [91] Ni-Meister, Wenge, Lee, Shihyan, Strahler, Alan H., Woodcock, Curtis E., Schaaf, Crystal, Yao, Tian, Ranson, K. Jon, Sun, Guoqing, and Blair, J. Bryan. Assessing general relationships between aboveground biomass and vegetation structure parameters for improved carbon estimate from lidar remote sensing. *Journal of Geophysical Research* 115 (2010).
- [92] Niklas, K. J. *Plant allometry: The scaling of form and process*. University of Chicago Press, 1994.
- [93] Nilsson, M. Estimation of tree heights and stand volume using an airborne lidar system. *Remote Sensing of Environment* 56 (1986), 1–7.

- [94] Osawa, A. Inverse relationship of crown fractal dimension to self-thinning exponent of tree populations: a hypothesis. *Can. J. For. Res.* 25, 10 (1995), 16081617.
- [95] Papathanassiou, K. P., and Cloude, S. R. Single-baseline polarimetric SAR interferometry. *IEEE Trans. Geosci. Remote Sensing* 39, 11 (2001), 2352–2363.
- [96] Papathanassiou, K. P., and Cloude, S. R. The effect of temporal decorrelation on the inversion of forest parameters from PolInSAR data. In *Proceedings of IEEE Geoscience and Remote Sensing Symposium (IGARSS)* (July 2003), pp. 1429–1431.
- [97] Parresol, B.R. Additivity of nonlinear biomass equations. *Canadian Journal of Forest Research* 31 (May 2001), 865–878(14).
- [98] Pastor, J., Aber, J.D., and Melillo, J.M. Biomass prediction using generalized allometric regression for some northeast tree species. *Forest ecology and management* 7 (1984), 265–274.
- [99] Pastor, John, and Bockheim, Jim. Biomass and production of an aspen - mixed hardwood - spodosol ecosystem in northern wisconsin. *Canadian Journal of Forest Research* 11, No. 1, 1 (1981), 132–138.
- [100] Perala, D.A., and Alban, D.H. Allometric biomass estimators for aspen-dominated ecosystems in the upper great lakes. *US Forest Service research paper NC-314* (1994).
- [101] Praks, J., and M. Hallikainen, F. Kugler, K. P. Papathanassiou. Coherence tomography for boreal forest: Comparison with HUTSCAT scatterometer measurements. In *Proceedings of the EUSAR Conference* (June 2008).
- [102] Prati, C., and Rocca, F. Improving slant-range resolution with multiple sar surveys. *Aerospace and Electronic Systems, IEEE Transactions on* 29, 1 (jan 1993), 135 –143.
- [103] Pulliainen, J., Engdahl, M., and Hallikainen, M. Feasibility of multi-temporal interferometric SAR data for stand-level estimation of boreal forest stem volume. *Remote Sensing of Environment* 85, 4 (2003), 397 – 409.
- [104] Reigber, A., and Moreira, A. First demonstration of airborne SAR tomography using multibaseline L-band data. *Geoscience and Remote Sensing, IEEE Transactions on* 38, 5 (Sep 2000), 2142–2152.
- [105] Reigber, A., and Scheiber, R. Airborne differential SAR interferometry: first results at L-band. *Geoscience and Remote Sensing, IEEE Transactions on* 41, 6 (June 2003), 1516–1520.

- [106] Richards, J. A., Sun, G. Q., and Simonet, D. S. L-band radar backscatter modeling of forest stands. *IEEE Trans. Geosci. Remote Sensing* **GE-25**, 4 (1987).
- [107] Rodriguez, E., and Martin, J. M. Theory and design of interferometric synthetic aperture radars. *Proc. of Inst. Elect. Eng.* **139**, 2 (April 1992), 147–159.
- [108] Rosen, P. A., Hensley, S., Joughin, I. R., Li, F. K., Madsen, S. N., Rodriguez, E., and Goldstein, R. M. Synthetic aperture radar interferometry. *Proc. of Inst. Elect. Eng.* **88**, 3 (March 2000), 333–382.
- [109] Saatchi, S., Halligan, K., Despain, D.G., and Crabtree, R.L. Estimation of forest fuel load from radar remote sensing. *Geoscience and Remote Sensing, IEEE Transactions on* **45**, 6 (june 2007), 1726 –1740.
- [110] Santoro, M., Askne, J., Smith, G., and Fransson, J. Stem volume retrieval in boreal forests from ERS-1/2 interferometry. *Remote Sensing of Environment* **81** (2002), 19–35.
- [111] Santoro, M., Shvidenko, A., McCallum, I., Askne, J., and Schmullius, C. Properties of ERS-1/2 coherence in the Siberian Boreal forest and implications for stem volume retrieval. *Remote Sensing of the Environment* **106** (2007), 154–172.
- [112] Schmitt, M.D.C, and Grigal, D.F. Generalized biomass estimation equations for Betula Papyrifera Marsh. *Canadian Journal of Forestry Research* **11** (1981), 837–840.
- [113] Simard, M., Zhang, K., Rivera-Monroy, V.H., Ross, M.S., Ruiz, P.L., Castaneda-Moya, E., Twilley, R.R., and Rodriguez, E. Mapping height and biomass of mangrove forests in everglades national park with SRTM elevation data. *Photogrammetric Engineering and Remote Sensing* **72**, 3 (2006), 299–311.
- [114] Slatton, K. C., Crawford, M. M., and Evans, B. L. Fusing interferometric radar and laser altimeter data to estimate surface topography and vegetation heights. *Geoscience and Remote Sensing, IEEE Transactions on* **39**, 11 (Nov 2001), 2470–2482.
- [115] Stehman, S., Wickham, J., Smith, J., and Yang, L. Thematic accuracy of the 1992 national land-cover data for the eastern United States: Statistical methodology and regional results. *Remote Sensing of Environment* **86** (2003), 500–516.
- [116] Stout, B. Species distribution and soils in the Harvard Forest. *Harvard Forest Bulletin* **24** (1952), 1–29.
- [117] Tebaldini, S. Algebraic synthesis of forest scenarios from multibaseline PolInSAR data. *Geoscience and Remote Sensing, IEEE Transactions on* **47**, 12 (Dec. 2009), 4132–4142.

- [118] Ter-Mikaelian, M. T., and Korzukin, M. D. Biomass equations for sixty-five North American tree species. *Forest Ecology and Management* 97 (1997), 1–24.
- [119] Touzi, R., Lopez, A., Bruniquel, J., and Vachon, P.W. Coherence estimation for SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing* 37, 1 (1999), 135–149.
- [120] Treuhaft, R. N., Madsen, S., Moghaddam, S. N., and Zyl, J. J. Van. Vegetation characteristics and underlying topography from interferometric radar. *Radio Science* 31 (1996), 1449–1485.
- [121] Treuhaft, R. N., and Siqueira, P. R. Vertical structure of vegetated land surfaces from interferometric and polarimetric radar. *Radio Science* 35, 1 (2000), 141–177.
- [122] Treuhaft, R.N., Chapman, B.D., dos Santos, J.R., Dutra, L.V., Goncalves, F.G., J.C., de Graca, P.M.A., and Drake, J.B. Vegetation profiles in tropical forests from multibaseline interferometric synthetic aperture radar, field, and lidar measurements. *Journal Of Geophysical Research* 114, D23110 (2009-Dec. 15 2009).
- [123] Tritton, M.L, and Hornbeck, J.W. Biomass equations for majore trees species in the northeast. *US Forest Service Northeastern experiment station research paper NE-619* (1982).
- [124] Ulaby, F. T., Sarabandi, K., McDonald, K., Whitt, M., and Dobson, M. C. Michigan microwave canopy scattering model. *International Journal of Remote Sensing* 11 (1990), 1223–1253.
- [125] Ulaby, Fawaz T., Moore, Richard K., and Fung, Adrian K. *Microwave Remote Sensing, Active and Passive*, vol. 2. Artech House Inc, Norwood, MA, 1986.
- [126] Ulander, Lars. Radiometric slope correction of synthetic-aperture radar images. *IEEE Trans. Geosci. Remote Sensing* 34, 5 (1996).
- [127] van Zyl, J.J. Unsupervised classification of scattering behavior using radar polarimetry data. *Geoscience and Remote Sensing, IEEE Transactions on* 27, 1 (jan 1989), 36 –45.
- [128] Vogelmann, J., Sohl, T., Campbell, P., and Shaw, D. Regional land cover characterization using LANDSAT thematic mapper data and ancillary data sources. *Environmental Monitoring and Assessment* 51 (1998), 415–428.
- [129] Walker, W. S., Kellendorfer, J. M., Lapoint, E., Hoppus, M., and Westfall, J. An empirical InSAR-optical fusion approach to mapping vegetation canopy height. *Remote Sensing of the Environment* 109 (2007), 482–499.

- [130] Wegmuller, U., and Werner, C. Retrieval of vegetation parameters with SAR interferometry. *Geoscience and Remote Sensing, IEEE Transactions on* 35, 1 (Jan 1997), 18–24.
- [131] West, Geoffrey B., Brown, James H., and Enquist, Brian J. A general model for the origin of allometric scaling laws in biology. *Science* 276, 5309 (1997), 122–126.
- [132] Whittaker, R. H., Bormann, F. H., Likens, G. E., and Siccamo, T. G. The hubbard brook ecosystem study: Forest biomass and production. *Ecological Monographs* 44, 2 (1974), 233–254.
- [133] Wiant, H.V., Sheetz, C.E., Colaninno, A., DeMoss, J.C., and Castadena, F. Tables and procedures for estimating weights of some appalachian hardwoods. *West Virginia Agriculture Experiment Station Bulletin* 659-T (1977).
- [134] Wu, S. T. Potential application of multi-polarization SAR for plantation pine biomass estimation. *IEEE Trans. Geosci. Remote Sensing GRS-25* (1987), 403–409.
- [135] Yamada, H., Yamaguchi, Y., Kim, Y., and Rodriguez, E. Polarimetric SAR interferometry for forest analysis based on the ESPRIT algorithm. *IEICE Transactions on Electronics E84C* (2001), 1917–1924.
- [136] Young, H. E., Ribe, J. H., and Wainwright, K. Weight tables for tree and shrub species in maine. *Life Sciences and Agriculture Experiment Station, University of Maine at Orono, Miscellaneous Report* 230 (1980).
- [137] Zebker, H. A., and Goldstein, R. M. Topographic mapping from interferometer synthetic aperture radar observations. *Journal of Geophysical Research* 91 (Apr. 1986), 4993–5000.
- [138] Zebker, H. A., and Villasenor, J. Decorrelation in interferometric radar echoes. *IEEE Transactions on Geoscience and Remote Sensing* 30 (September 1992), 950–959.
- [139] Zeide, B., and Pfeifer, P. A method for estimation of fractal dimension of tree crowns. *Forest Science* 37 (1 November 1991), 1253–1265(13).