AN ABSTRACT OF THE THESIS OF

Dirk Pflugmacher for the degree of Master of Science in Forest Science presented on November 30, 2007.

Title: Remote Sensing of Forest Aboveground Biomass using the Geoscience Laser Altimeter System.

Abstract approved:

_____

Warren B. Cohen

Accurate estimates of forest aboveground biomass are needed to reduce uncertainties in the terrestrial carbon flux. The Geoscience Laser Altimeter System (GLAS) onboard the Ice, Cloud and land Elevation Satellite is now the first spaceborne lidar sensor that will provide global estimates of vegetation height. This study investigated the utility of the GLAS sensor for large-scale biomass inventories by focusing on two important factors: the regional accuracy of GLAS-estimated forest height algorithms and the accuracy of general height-biomass allometric equations. Field data from the U.S Forest Service Inventory and Analysis (FIA) program was used to compare regional height estimates with GLAS predictions. GLAS algorithms provided generally accurate estimates of height and were on average 2-3 m lower than FIA estimates. The analysis of the regional variability of height-biomass relationships in the FIA data suggests that general non-species specific equations are applicable without a significant loss in prediction accuracy. Regional estimates of forest biomass from GLAS were about 20% lower than FIA estimates (difference between 39.7 – 58.2 Mg $ha^{-1}$).

# Remote Sensing of Forest Aboveground Biomass using the Geoscience Laser Altimeter System

by
Dirk Pflugmacher

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented November 30, 2007
Commencement June 2008

Master of Science thesis of Dirk Pflugmacher presented on November 30, 2007.

APPROVED:

_____

Major Professor, representing Forest Science

_____

Head of the Department of Forest Science

_____

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

_____

Dirk Pflugmacher, Author

# ACKNOWLEDGEMENTS

I would like to thank my advisor, Warren Cohen (U.S. Forest Service), for his guidance and the financial support for this research. His feedback and encouragement over the two years have been considerable. I would also like to thank my committee member Robert Kennedy (U.S. Forest Service) for his advice and inspiring discussions. He is a great mind; I have always enjoyed exchanging ideas with him. Thank you as well to Michael Lefsky (Colorado State University) for initiating the GLAS pilot project, thus making this work possible, and for inviting me to Colorado. Thanks to my fellow graduate student and office mate Al Kirschbaum for the many stimulating discussions and friendship. Thank you to everyone from the Laboratory for Applications of Remote Sensing in Ecology, the College of Forestry, and my family and friends for your support and guidance.

# CONTRIBUTION OF AUTHORS

Dr. Warren Cohen, Dr. Michael Lefsky and Dr. Robert Kennedy have been greatly involved in designing and guiding this research. Dr. Cohen, Dr. Lefsky, Dr. Kennedy and Dr. Turner have provided valuable comments to an earlier draft of chapter 2. Dr. Lefsky provided GLAS data and height estimates. Manuela Huso gave helpful advice with respect to the GLAS sampling.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDIX FIGURES

LIST OF APPENDIX TABLES

# Chapter 1 - Introduction

## Background

The concentration of atmospheric $CO_2$ has been increasing rapidly since the industrial revolution, which is likely to have implications for the global climate. The observed increase is attributed primarily to the burning of fossil fuels with contributions from cement manufacture, and to a lesser degree from deforestation and changing agricultural practices. At the same time the rate of carbon uptake by the ocean and the land has increased in the last decades and partly offset these trends (Denman et al. 2007). However, to what extent and for how long the terrestrial biosphere and the ocean can offset anthropogenic carbon emissions is highly uncertain.

The importance of monitoring anthropogenic and natural carbon sources and sinks is now widely recognized, and national and international research programs are making efforts to address these issues (Wofsy and Harriss 2002; Denman et al. 2007). One of the great uncertainties in the global carbon cycle is associated with the global land-atmosphere flux, which is currently most reliably estimated indirectly by deducting the residual between fossil fuel and cement emissions and the total uptake by the ocean and atmosphere (Denman et al. 2007). Direct observations of the carbon flux via flux measurements by eddy covariance technique (Law et al. 2001) or biomass inventories (Goodale et al. 2002) are too sparse, given the heterogeneity of

terrestrial ecosystems, to provide inferences with sufficient accuracy (Denman et al. 2007).

Global knowledge on distribution and rate of change of biomass stored in forest ecosystems is crucial to reduce the uncertainties in the terrestrial carbon flux (Houghton 2005). Although forested biomes cover only about a third of the ice-free land surface, they account for 80% of the global plant biomass; most (70-80%) is allocated aboveground (Chapin et al. 2002). National forest inventories provide valuable field data on forest attributes in many countries. However, their utility for global-scale studies is limited, because of methodological differences and the lack of observations in remote regions. Consequently, the potential value of space-based observations is high. Satellite remote sensing is an efficient means to monitor the Earth's surface in a spatially consistent way. However, mapping forest biomass has been challenging with existing satellite sensors (Lu 2006).

The most promising sensor technology for remote estimation of forest biomass today is lidar (Hese et al. 2005), which until recently has been limited to airborne systems. Studies have demonstrated that airborne lidar (light detection and ranging) is capable of measuring forest height with high accuracy and that forest aboveground biomass can be accurately estimated from lidar heights (Drake et al. 2003; Lefsky et al. 2002; Lefsky et al. 1999; Patenaude et al. 2004). The success of lidar in forest environments has ultimately led to efforts in the United States (VCL: Vegetation Canopy Lidar mission, Blair et al. 1999) and later in Germany (Carbon-3D mission, Hese et al. 2005) to implement a space-based lidar mission for vegetation studies. Airborne simulations with the Laser Vegetation Imaging Sensor (LVIS), designed for

the VCL mission, showed good accuracies for estimating aboveground biomass at test sites in Costa Rica (RMSE[1]=60.02-63.17 Mg $ha^{-1}$, Drake et al. 2002) and the U.S. (RMSE = 54.8-73.5 Mg $ha^{-1}$, Hyde et al. 2005). Nevertheless, the two proposed space missions did not receive funding from the respective space agencies.

In January, 2003, the Ice, Cloud and land Elevation Satellite (ICESat) was launched as part of NASA's Earth Observing System (EOS) of satellites to make global laser observations over the polar ice sheet, the land, the ocean and the atmosphere over a period of 3-5 years (Zwally et al. 2002). ICESat carries a single sensor: the Geoscience Laser Altimeter System (GLAS). GLAS utilizes three diode pumped Q-switched Nd:YAG lasers operating in the near-infrared wavelength (1064-nm) for measuring the elevation of surfaces and dense clouds and in the green wavelength (532-nm) for measuring the vertical distribution of clouds and aerosols. Each laser fires pulses over approximately month long periods three times a year (designated by sequential letters).

The terrestrial component of the mission provides measurements of land topography and vegetation canopy heights, which represents a unique opportunity for global biomass research. However, since the primary objective of the ICESat mission is to monitor changes in elevation of the Greenland and Antarctic ice sheets, the sensor design is not optimal for vegetation studies. For example, GLAS has a relatively large laser foot-print size (64 m diameter compared to 25 m diameter proposed for the VCL and Carbon-3D mission), which can decrease its accuracy. Nevertheless, in a pilot

[1] RMSE=root-mean-squared-error

study Lefsky et al. (2005) demonstrated that GLAS is able to predict forest height (RMSE= 4.85 m - 12.1 m) and biomass (RMSE=58.3 Mg $ha^{-1}$).

## GLAS

### *GLAS altimetry data*

GLAS uses a laser altimeter which determines the range distance between the satellite and the Earth's surface by measuring the round-trip travel time of a 1064-nm leaser pulse emitted from the sensor. The laser foot-print on the surface is elliptical with an equivalent circular area of 64 m diameter (Abshire et al. 2005). An onboard GPS receiver and an inertial reference system (IRS) is used to determine the precise position and attitude of the spacecraft and reference topographic measurements to a global datum. The predicted accuracy for the GLAS surface elevation measurements is 15 cm, (Zwally et al. 2002). The horizontal geolocation accuracy has varied between 4.6 and 17.4 m (National Snow and Ice Data Center 2008). However, there are factors that could diminish the actual geolocation accuracy such as instrumental errors, atmospheric scattering and surface conditions. To provide quality assurance, the ICESat data products are accompanied by flags indicating the quality of the orbit and attitude determination (predicted or precision), the loss of GPS data, the correction algorithms applied to the range and the waveform-detected presence of cloud layers.

For each laser pulse, GLAS records the 1064-nm wavelength energy of the echo pulse as a function of time. Thus, the instrument does not yield a discrete measurement of elevation but acquires a vertical profile of the illuminated surfaces within the laser foot-print. The GLAS onboard digitizer records the entire laser pulse in 1 ns intervals, but only the transmitted portion and the surface echo are extracted and transmitted to the ground station (Zwally et al. 2002). Return waveforms are transmitted in 544 bins for ice sheets and land, and 200 bins for oceans and sea ice. The 544 bins of 1 ns intervals represent a range distance of 81.5 m, where 1 ns is the two-way travel time of the pulse corresponding to a range distance of 15 cm ($range = \frac{time \times c}{2}$, where $c$ is the speed of light = 30 cm·ns$^{-1}$). To extend the range over land surfaces to 150 m, the upper 151 bins are averaged to 4 ns intervals starting with laser 2 and operation period b. Global altimetry data are distributed under the product name GLA01.

In addition to the "raw" altimetry data, the ICESat science team provides global elevation data for polar ice sheets (GLA12), sea ice (GLA13), land surfaces (GLA14) and oceans (GLA15). The "standard" algorithm derives mean surface elevation from the centroid of a single Gaussian distribution (or the larger of two) fitted to the waveform. Fitting a Gaussian distribution helps to diminish the effects from small-scale irregularities in the surface and forward scattering (Brenner et al. 2003). The shape of the return waveform is affected by both the shape of the transmitted pulse and the distribution of surface heights within the laser foot-print. Since the transmitted waveform is Gaussian (with a 4-ns width at half the amplitude)

the return waveform is also close to Gaussian assuming the surface is even (no relief or vegetation) or the surface height variations are random, and no atmospheric forward scattering occurs (Brenner et al. 2003). Surface slope and roughness both result in a broadening of the return pulse. Thus, the variance of the Gaussian model can be used to make inferences on these two parameters.

In comparison to ice sheets, sea ice and oceans, land surfaces often depict a complex conglomerate of slope, roughness (relief and vegetation) and varying reflectance properties. In particular, the presence of tall vegetation and cultural features can result in complex, multimodal surface height profiles. Thus, an "alternate" approach is employed to parameterize land waveforms (Zwally et al. 2003). This approach defines the elevation of the first and last detected surfaces by a low-amplitude threshold (signal start and signal end). In this study the threshold was set to the mean background noise value reported in the GLA01 product (GLAS product variable: d_4nsBgMean) plus 4.5 times the standard deviation of the background noise (GLAS product variable: d_4nsBgSDEV). The range between signal start and signal end is referred to as the waveform extent. Further, the extent of the leading and trailing edge measured from signal start and signal end, respectively, is defined where the waveform crosses the mean energy above the background noise (Lefsky et al. 2007). If the return waveform is unimodal, the centroid of the waveform between signal start and signal end serves as an alternate estimate for mean ground elevation (Harding and Carabajal 2005). An alternate representation of the waveform model is the sum of up to six Gaussian components

fitted by a least-squares algorithm (Figure 1.1 and Figure 1.2) (Brenner et al. 2003). The alternate waveform model is defined as:

$$w(t) = \varepsilon + \sum_{m=1}^{N_P} A_m \times e^{\frac{-(t-t_m)^2}{2\sigma_m^2}} \quad (1)$$

where

$w(t)$ = the amplitude of the waveform at time $t$

$N_p$ = number of peaks in the waveform

$A_m$ = amplitude of the $m^{th}$ peak

$\varepsilon$ = bias (noise level) of the waveform

$t_m$ = position of $m^{th}$ peak at time $t$

$\sigma_m$ = standard deviation of the $m^{th}$ peak

*GLAS waveforms over vegetation*

Where waveforms occur over vegetated areas, the surface height distribution produced by the ground reflection is mixed with the height distribution of the vegetation surfaces. Differentiation of the vegetation signal from the ground therefore becomes increasingly difficult with increasing topographic relief (slope and roughness). In cases of small topographic relief, separation of the vegetation from the ground is relatively straight forward. Small and moderate relief typically yields a bi- or multimodal return waveform, where the first mode or modes correspond to one

or multiple vegetation layers (Harding and Carabajal 2005). Assuming sufficient energy penetrates through canopy gaps and reaches the ground, the last mode corresponds to the mean ground elevation. The signal start of the waveform indicates the elevation of the uppermost plant surface detected. Thus, maximum canopy height is correlated to the distance between the signal start and the centroid of the ground peak (Harding and Carabajal 2005). The extent of the leading edge is related to the variability in height (i.e. ruggedness) of the uppermost canopy, while the extent of the trailing edge is related to terrain slope (Figure 1.1) (Lefsky et al. 2007). Where topographic relief is high, the vegetation signal is convoluted with the ground return and inferences on the vegetation structure are more complicated (Figure 1.2).

*Vegetation height algorithms*

As a consequence of the large laser footprint, the main challenge in estimating vegetation heights using GLAS waveforms is associated with areas of high topographic complexity in combination with high variability in vegetation heights and low to medium vegetation density. Figure 1.2 shows an example for a waveform over forest in steep terrain, where the return energy from vegetation and ground surfaces becomes convoluted. Inferences on the vertical distribution of vegetation heights alone are not possible without making assumptions on the underlying ground. The first algorithms for vegetation height from Lefsky et al. (2005) utilized digital elevation models (DEMs) from the Shuttle Radar Topography Mission (SRTM) to account for the pulse broadening associated with topographic

slope. Lefsky et al. (2005) used coincident field data to calibrate models of maximum tree height as a function of waveform extent and terrain slope. The study demonstrated that height estimation in complex terrain is possible with GLAS. However, the main disadvantage of this approach was that it relied on an external source for terrain information that itself is biased over forested areas (Sun et al. 2003). For forest covered areas, the C-band interferometric SAR measures a height within the tree canopy and therefore overestimates ground elevation. In addition to the positive bias introduced from using SRTM data, previous algorithms revealed a negative bias in some cases as the result of using maximum height as predicted variable. Maximum height tends to be underestimated when the topmost canopy surface area is not sufficiently large enough to be detected in the waveform (i.e. a single tall tree).

The work reported here uses a new generation of algorithms developed by Lefsky et al. (2007). The revised algorithms yield estimates of mean forest height and account for topographic variations by incorporating additional waveform parameters instead of relying on auxiliary data sources. In addition to waveform extent, the new models include the extent of the leading and trailing edge (see Figure 1.2). Coincident field data at three test sites (Cascades, Oregon; Appalachians, Tennessee; Amazon, Brazil) and high resolution airborne lidar data at two test sites (Bartlett Experimental Forest, New Hampshire; Tahoe National Forest, California) were used for model calibration. Height estimates correspond to the mean height of the dominant and co-dominant trees for field sites and mean canopy height for sites with airborne lidar data. The regression models reported in Lefsky et al. (2007)

estimate a "correction factor" (*cf*), which is subtracted from the waveform extent to obtain an estimate of mean height. The equation for the *cf* is as follows:

$$cf = 8.96 + 1.52\ lf + 1.14\ tf - offset \tag{2}$$

where *offset* is a constant to account for differences between test sites. In this study, two correction factors were calculated using the *offset* from the test site in the Cascades, Oregon (*offset*=4.83) and the test site in the Appalachians, Tennessee (*offset*=1.15). *Lf* and *tf* denote the trailing and leading edge correction factor, a transformation of the trailing and leading edge extent, respectively:

$$tf = 3.4\sqrt{trail} + 0.92 \cdot trail - 88.5\frac{trail}{extent} + 2049.5 + \frac{trail}{extent^2} - 14171.4\frac{trail}{extent^3} \tag{3}$$

$$lf = 0.72 \cdot lead - 22.8\frac{lead}{extent} \tag{4}$$

where

*lead* = leading edge extent

*trail* = trailing edge extent

*extent* = waveform extent

**Summary**

Despite the challenges associated with high topography, GLAS currently provides the most promising approach for globally consistent observations of forest height and biomass from space. Recent studies have achieved good accuracies for estimating forest height (RMSE = 4.85-12.66 m, Lefsky et al. 2005; RMSE= 5 m, Lefsky et al. 2007) and biomass (RMSE=58.3 Mg $ha^{-1}$, Lefsky et al. 2005) at test sites across North and South America. However, whether the developed concepts are valid at larger spatial scales has yet to be determined. The study presented here examined the utility of GLAS data for large-scale biomass inventories at two large forested regions in the United States. Using GLAS for remote sensing of biomass is a novel approach that may greatly advance our knowledge on terrestrial carbon stores. As part of NASA's Terrestrial Ecology Program, this study aimed to advance the knowledge on state-of-the-art remote sensing and its potential contribution to global terrestrial ecosystem science.

**References**

Abshire, J.B., X.L. Sun, H. Riris, J.M. Sirota, J.F. McGarry, S. Palm, D.H. Yi, and P. Liiva. 2005. Geoscience Laser Altimeter System (GLAS) on the ICESat mission: On-orbit measurement performance. *Geophys. Res. Lett.* 32(21):doi:10.1029/2005GL024028.

Blair, J.B., D.L. Rabine, and M.A. Hofton. 1999. The Laser Vegetation Imaging Sensor: a medium-altitude, digitization-only, airborne laser altimeter for mapping vegetation and topography. *ISPRS J. Photogramm.* 54(2-3):115-122.

Brenner, A., H.J. Zwally, C. Bentley, B. Csathó, D.J. Harding, M.A. Hofton, J.-B. Minster, L. Roberts, J.L. Saba, R.H. Thomas, and Y. Donghui. 2003. *Derivation of range and range distributions from laser pulse waveform analysis for surface elevations, roughness, slope, and vegetation heights*. Algorithm Theoretical Basis Document 4.1. Available online at http://www.csr.utexas.edu/glas/atbd.html; last accessed Sep. 15, 2007.

Chapin, F.S., P.A. Matson, and H.A. Mooney. 2002. *Principles of terrestrial ecosystem ecology*. Springer, New York. xiv, 436 p. p.

Denman, K.L., G. Brasseur, A. Chidthaisong, P. Ciais, P.M. Cox, R.E. Dickinson, D. Hauglustaine, C. Heinze, E. Holland, D. Jacob, U. Lohmann, S. Ramachandran, P.L. da Silva Dias, S.C. Wofsy, and X. Zhang. 2007. Couplings between changes in the climate system and biogeochemistry. P. 499-587 in *Climate change 2007: The physical science basis*. *Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change*, Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, and H.L. Miller (eds.). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Drake, J.B., R.O. Dubayah, D.B. Clark, R.G. Knox, J.B. Blair, M.A. Hofton, R.L. Chazdon, J.F. Weishampel, and S.D. Prince. 2002. Estimation of tropical forest structural characteristics using large-footprint lidar. *Remote Sens. Environ.* 79(2-3):305-319.

Drake, J.B., R.G. Knox, R.O. Dubayah, D.B. Clark, R. Condit, J.B. Blair, and M. Hofton. 2003. Above-ground biomass estimation in closed canopy Neotropical forests using lidar remote sensing: factors affecting the generality of relationships. *Global Ecol. Biogeogr.* 12(2):147-159.

Goodale, C.L., M.J. Apps, R.A. Birdsey, C.B. Field, L.S. Heath, R.A. Houghton, J.C. Jenkins, G.H. Kohlmaier, W. Kurz, S.R. Liu, G.J. Nabuurs, S. Nilsson, and A.Z. Shvidenko. 2002. Forest carbon sinks in the northern hemisphere. *Ecol. Appl.* 12(3):891-899.

Harding, D.J., and C.C. Carabajal. 2005. ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure. *Geophys. Res. Lett.* 32(21):doi:10.1029/2005GL023471.

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

Houghton, R.A. 2005. Aboveground forest biomass and the global carbon balance. *Glob. Change Biol.* 11(6):945-958.

Hyde, P., R. Dubayah, B. Peterson, J.B. Blair, M. Hofton, C. Hunsaker, R. Knox, and W. Walker. 2005. Mapping forest structure for wildlife habitat analysis using waveform lidar: Validation of montane ecosystems. *Remote Sens. Environ*. 96(3-4):427-437.

Jenkins, J.C., D.C. Chojnacky, L.S. Heath, and R.A. Birdsey. 2003. National-scale biomass estimators for United States tree species. *For. Sci.* 49(1):12-35.

Law, B.E., P.E. Thornton, J. Irvine, P.M. Anthoni, and S. Van Tuyl. 2001. Carbon storage and fluxes in ponderosa pine forests at different developmental stages. *Glob. Change Biol.* 7(7):755-777

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

Lefsky, M.A., D. Harding, W.B. Cohen, G. Parker, and H.H. Shugart. 1999. Surface lidar remote sensing of basal area and biomass in deciduous forests of eastern Maryland, USA. *Remote Sens. Environ.* 67(1):83-98.

Lefsky, M.A., D.J. Harding, M. Keller, W.B. Cohen, C.C. Carabajal, F.D. Espirito-Santo, M.O. Hunter, R. de Oliveira, and P.B. de Camargo. 2005. Estimates of forest canopy height and aboveground biomass using ICESat. *Geophys. Res. Lett.* 32(22):doi:10.1029/2005GL023971.

Lefsky, M.A., M. Keller, Y. Pang., P.B. de Camargo, M.O. Hunter. 2007. Revised method for forest canopy height estimation from Geoscience Laser Altimeter System waveforms. *J. Appl. Rem. Sens.* 1(1): 1(1):013537-18

Lu, D.S. 2006. The potential and challenge of remote sensing-based biomass estimation. *Int. J. Remote Sens.* 27(7):1297-1328.

National Snow and Ice Data Center, Laser Operational Periods - Attributes, Available at: http://nsidc.org/data/icesat/glas_laser_ops_attrib.pdf (access 2 January 2008).

Patenaude, G., R.A. Hill, R. Milne, D.L.A. Gaveau, B.B.J. Briggs, and T.P. Dawson. 2004. Quantifying forest above ground carbon content using LiDAR remote sensing. *Remote Sens. Environ.* 93(3):368-380.

Sun, G., K.J. Ranson, V.I. Khairuk, and K. Kovacs. 2003. Validation of surface height from shuttle radar topography mission using shuttle laser altimeter. *Remote Sens. Environ.* 88(4):401-411.

Wofsy, S.C., and R.C. Harriss. 2002. *The North American Carbon Program (NACP)*. Report of the NACP Committee of the U.S. Interagency Carbon Cycle Science Program US Global Change Research Program. 75.

Zwally, H.J., B. Schutz, W. Abdalati, J. Abshire, C. Bentley, A. Brenner, J. Bufton, J. Dezio, D. Hancock, D. Harding, T. Herring, B. Minster, K. Quinn, S. Palm, J. Spinhirne, and R. Thomas. 2002. ICESat's laser measurements of polar ice, atmosphere, ocean, and land. *J. Geodyn.* 34(3-4):405-445.

**Figures**

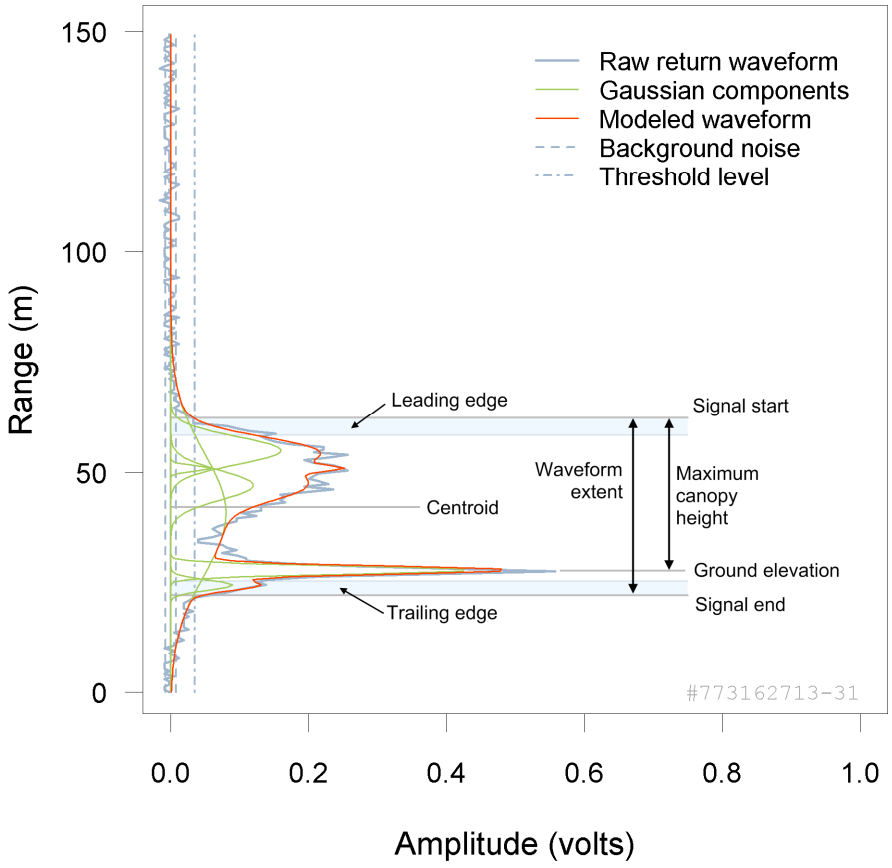


Figure 1.1. Bimodal GLAS waveform over forest land with small topographic slope. The first peak represents the canopy layers and the last peak represents the ground surface. Signal start and signal end are the first and last crossing of a low-amplitude threshold. The extent of the leading and trailing edge (light blue area) is defined where the waveform crosses the mean energy level within the waveform extent after the signal start and before the signal end, respectively. The "alternate" model form of the waveform (red line) is the sum of up to six Gaussian distributions (green lines) fitted to the six largest (by area) peaks. The centroid of the last Gaussian peak is used to estimate the mean ground elevation.

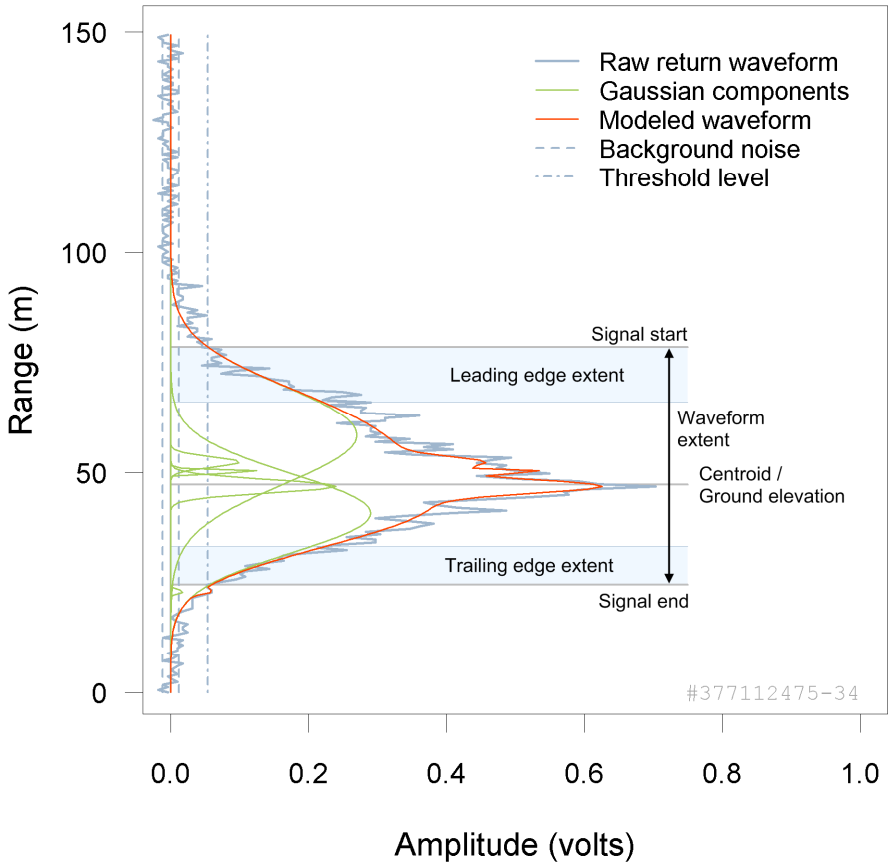


Figure 1.2. Unimodal GLAS waveform over forest land with steep terrain. Signal start and signal end represent the first and last crossing of a low-amplitude threshold. The extent of the leading and trailing edge (light blue area) is defined where the waveform crosses the mean energy level within the waveform extent after the signal start and before the signal end, respectively. The "alternate" model form of the waveform (red line) is the sum of up to six Gaussian distributions (green lines) fitted to the six largest (by area) peaks. The centroid of the waveform between signal start and signal end is an (alternate) estimate for the mean elevation of the ground surface.

# Chapter 2 - Regional applicability of forest height and aboveground biomass models for the Geoscience Laser Altimeter System

Dirk Pflugmacher[1*], Warren B. Cohen[2], Robert E. Kennedy[2], Michael A. Lefsky[3]

[1] Oregon State University, Department of Forest Science, 321 Richardson Hall, Corvallis, OR 97331, United States

[2] USDA Forest Service, Pacific Northwest Research Station, Forestry Sciences Laboratory, 3200 SW Jefferson Way, Corvallis, OR 97331, United States

[3] Center for Ecological Analysis of Lidar, Department of Natural Resources, Colorado State University, 131 Forestry Building, Fort Collins, CO 80523-1472, United States

Forest Science
5400 Grosvenor Lane
Bethesda, MD 20814-2198, USA
*Submitted*

**Abstract**

Accurate estimates of forest aboveground biomass are needed to reduce uncertainties in global and regional terrestrial carbon fluxes. In this study we investigated the utility of the Geoscience Laser Altimeter System (GLAS) onboard the Ice, Cloud and land Elevation Satellite for large-scale biomass inventories. GLAS is the first spaceborne lidar sensor that will provide global estimates of forest height. We compared accuracy and regional variability of GLAS height estimates with data from the U.S. Forest Service Inventory and Analysis (FIA) program and found that current GLAS algorithms provided generally accurate estimates of height. GLAS heights were on average 2.4 m lower than FIA estimates in a conifer dominated study region and 3.3 m lower than FIA estimates in a broadleaf tree dominated study region. To translate GLAS-estimated heights into forest biomass requires general allometric equations. Analyses of the regional variability of forest height–biomass relationships using FIA field data indicate that general non-species-specific equations can be applied without a significant loss of prediction accuracy. We developed biomass models from FIA data and applied them to the GLAS-estimated heights. Regional estimates of forest biomass from GLAS differed between 39.7 – 58.2 Mg $ha^{-1}$ (18% - 27%) when compared to FIA.

## Introduction

Accurate estimates of forest biomass, its spatial distribution and rate of change are required to quantify global and regional terrestrial carbon fluxes, and to formulate mitigation strategies for current and future greenhouse gas emissions. Current observations of terrestrial carbon flux are too sparse and the ecosystems are too heterogeneous to allow global assessment of the net biospheric flux with sufficient accuracy (Denman et al. 2007). National inventories provide the most extensive field observations and thus have been a key information source in many carbon studies (Goodale et al. 2002). However, the lack of field surveys in remote areas like the tropics and northern latitudes as well as methodological differences, raise uncertainties when carbon statistics are aggregated to the global scale (Houghton 2005). In addition most inventories are designed to provide inferences on the basis of administrative units or large regions. This is an important limitation, as information on carbon flux is needed on a spatial scale small enough to be linked to individual landscape units (i.e. forest stands) as they undergo natural disturbances, succession or land-use changes (Houghton 2005).

While spatially explicit and consistent earth observations are a primary strength of satellite remote sensing, remotely sensed estimation of forest biomass remains a challenging task (Lu 2006). Aboveground biomass is a three dimensional variable. Hence, the capability of satellite sensors to provide accurate estimates depends on their ability to discriminate vertical forest structure. Many studies have demonstrated that canopy reflectance (Dong et al. 2003; Labrecque et al. 2006) and

radar backscatter (Ranson et al. 1997) are correlated with aboveground biomass. However, both are two-dimensional measures and become asymptotic with canopy closure, limiting their ability to predict biomass in high biomass forests (Imhoff et al. 1998; Turner et al. 1999). Lidar and interferometric synthetic aperture radar (InSAR) are both promising technologies in that they provide a measure of the vertical structure. Since forest biomass is directly related to height (Drake et al. 2002; Lefsky et al. 1999; Lefsky et al. 2005a; Woodhouse 2006), lidar and InSAR represent more robust tools to determine biomass consistently across large areas. Nevertheless, InSAR has not yet achieved accuracies comparable to lidar (Treuhaft et al. 2004).

A spaceborne lidar mission is currently the most promising approach for accurate and globally consistent sampling of forest height and biomass. The Geoscience Laser Altimeter System (GLAS) onboard the Ice, Cloud and land Elevation Satellite (ICESat) is the first lidar mission to acquire global measurements of vegetation height (Zwally et al. 2002). Height estimates have been obtained by means of empirical algorithms developed at seven training sites in North and South America (Lefsky et al. 2005a; Lefsky et al. 2007). Training sites were selected to cover a range of biomes and topographic conditions. To be regionally applicable the generality of the estimation algorithms must be tested.

Another potential source of uncertainty in GLAS-based biomass estimation is determined by the degree to which environmental factors such as tree species composition, stand structure and site conditions affect the relationship between lidar-height and forest biomass. To convert GLAS-estimated vegetation heights into estimates of forest biomass on a large geographic scale will require general allometric

equations that are applicable over a range of forest types and conditions. Regional to global-level studies do not have the detailed information available as local-level studies. However, empirical knowledge on the relationships between lidar-derived canopy height metrics and total aboveground biomass of all trees within the laser foot-print currently exists only locally (e.g. Drake et al. 2002; Lefsky et al. 2005a). Consequently, there is a need for developing regionally to globally applicable allometric equations.

The goal of this study was to investigate the applicability of GLAS data for conducting regional and global forest biomass inventories. The applicability of GLAS data to broader areas primarily depends on two factors: the regional accuracy of GLAS height estimates and the accuracy of height-to-biomass allometry. In this study we investigated these two factors using GLAS data and extensive field data from the U.S. Forest Service, Forest Inventory and Analysis (FIA) program for two large regions. Our objectives were to 1) evaluate the regional applicability of GLAS height estimation algorithms, 2) develop and evaluate regional models of height and biomass relationships, and 3) evaluate regional estimates of forest biomass derived from GLAS-estimated heights.

**Methods**

Ideally, to evaluate the accuracy of height and biomass estimates, we would use reference data that has co-located GLAS waveforms and field-based height and biomass measurements. However, this approach is costly and therefore only feasible

for a small sample of waveforms (i.e. for algorithm training sites). To validate the accuracy of GLAS estimates over a large geographic region therefore requires a different approach. Alternatively, national forest inventories provide extensive field measurements of forest height and biomass that can be used to validate GLAS estimates on a regional level.

In this study we used an approach that compared regional estimates of forest height derived from GLAS waveforms with regional estimates from FIA (objective 1). FIA data was used to develop plot-level height-biomass models (objective 2), which we applied to the GLAS-estimated heights and compared the resulting biomass estimates with estimates from FIA (objective 3).The region of applicability of height and biomass models within the study regions was evaluated by grouping FIA and GLAS samples into ecologically meaningful strata.

*Study Regions*

Our research was focused on the Cascade Mountains in the Pacific Northwest region and the Appalachian Mountains in the eastern U.S. (Figure 2.1). These two regions play an important role in the terrestrial carbon cycle and span a wide range of topographic and floristic conditions. The forests in the Cascades study region are mainly comprised of needleleaf trees, with Douglas-fir (*Pseudotsuga menziesii*) being the dominant species, followed by Sitka spruce (*Picea sitchensis*), mountain hemlock (*Tsuga mertensiana*) and lodgepole pine (*Pinus contorta*). In the Appalachian study region, broadleaf deciduous species dominate (e.g. oak and hickory species).

The precise study region boundaries were determined based on inventory data availability and the geographic extent of ecological sections and subsections (Table 2.1) as defined by the Forest Service (McNab and Avers 1994). Due to limited availability of recent statewide forest inventory data at the time of this analysis, the Cascades study region was confined to the state of Oregon. Similarly, the northern boundary of the Appalachian study region is bounded by the state borders of Kentucky and Virginia.

*GLAS data*

The Geoscience Laser Altimeter System (GLAS) includes a waveform digitizing lidar sensor. The instrument emits a laser pulse with an approximate ground footprint of 64 m (Abshire et al. 2005). Within each footprint, laser energy is reflected back by all intercepting surfaces, resulting in a waveform that represents a vertical height profile of laser-illuminated surfaces.

In flat terrain and homogeneous forests, stand height is closely related to waveform extent, which is defined as the vertical distance between the first and last elevations at which the waveform energy exceeds a threshold level (Harding and Carabajal 2005). However, due to the relatively large GLAS footprint, the separation of the ground return from the vegetation surfaces is complicated in steep terrain and heterogeneous forest cover (Lefsky et al. 2005a). The first algorithm that accounted for terrain effects incorporated a digital elevation model (Lefsky et al. 2005a), and was

a proof-of-concept that vegetation heights and biomass can be predicted from GLAS waveforms under complex terrain conditions.

The study presented here utilizes a new generation of height algorithms that do not require auxiliary topographic information, but are based solely on the properties of the waveform (Lefsky et al. 2007). The height algorithms are calibrated to estimate the mean height of the dominant/co-dominant trees using field plots that coincide with GLAS samples. The coincident field plots that were used to train the height algorithms for the Cascades and Appalachian study regions are located in the Willamette National Forest and the Great Smoky Mountain National Park, respectively (see Lefsky et al. 2007).

*Forest inventory data*

Recent, annual field data from the U.S. Forest Service FIA program was used as reference data set, as it provides a consistent and extensive data source suitable for large scale studies. In the last decade, FIA has gradually moved from a periodic sampling scheme that was regionally specific to a nationally consistent, annual sampling scheme (McRoberts et al. 2005). The new annual forest inventory is based on a 5 to 10 year measurement cycle, but delivers a complete systematic sample for each state on a yearly basis.

Field data are collected by FIA on permanent plots. Each field plot consists of a set of four circular subplots, over which most tree measurements are taken. Trees with a diameter at breast height (dbh) of 12.5 cm and larger are measured within a 7.3-

m fixed radius, saplings (2.5 to 12.5 cm dbh) are measured on a 2-m microplot, and large trees (dbh > 102 cm) are measured on an 18-m fixed-radius macroplot. For each plot, FIA field crews assign one or multiple condition classes based on a series of predetermined discrete variables such as land use, forest type, stand size, tree density and ownership (Bechtold and Scott 2005). To avoid boundary plots (i.e. between forest/non-forest, forest types or distinctly different successional stages) we selected only plots where all four subplots were located completely on forest land and in the same condition class (Table 2.2).

Tree heights are measured in the field as the total length of a tree from the ground to the tip of the apical meristem. In some cases heights are visually estimated. Trees with estimated heights represented only a small proportion in this study (at average 3.6% per plot). For each plot, we calculated mean height of open grown, dominant and co-dominant trees (dcd-height), as this height metric was used to calibrate GLAS waveforms. Identification of dominant and co-dominant trees is based on the FIA crown class classification, which describes the position of a tree within the upper canopy layer (USDA Forest Service 2006). In addition to dcd-height, we computed maximum height, as the height of the tallest tree, and mean height of all trees with dbh greater than 2.5 cm.

We estimated total aboveground (oven-dry weight) biomass of all live trees (dbh 2.5 cm and larger) for each inventory plot from a set of ten allometric equations developed for large-scale studies (Jenkins et al. 2003). The “Jenkins equations” distinguish between four hardwood and six softwood species groups and use a simple log-linear regression model with dbh as predictor variable. Although FIA also reports

total aboveground tree mass, these estimates are based on regional volume tables or models. Methodological differences (e.g., model form, parameter attributes and nonlinear regression method) between FIA districts have been found to introduce regional biases (Hansen 2002). Conversely, the Jenkins equations are consistent in the way trees of defined dimension and species are treated across the U.S. On rare occasions, trees have been measured at root collar (instead of dbh); in these cases the FIA biomass estimate was used. To obtain plot-level estimates of live aboveground biomass (Mg $ha^{-1}$), we multiplied the tree biomass values with a trees-per-acre expansion factor reported by FIA and calculated the sum of the biomass of all live trees per plot.

*Estimation of regional parameters*

We derived regional estimates of mean forest height and biomass from FIA and GLAS samples using simple and stratified estimation. Both methods require a probability sampling design, e.g. random or systematic sample selection. FIA uses a systematic sample on a hexagonal grid with an approximate spacing of 5.3 km, which is expected to produce a random, equal probability sample (Scott et al. 2005). We calculated stratified estimates of means ( $\bar{y}$ ) and standard errors (*SE*) using the standard formulae from (Cochran 1977), ignoring finite population correction factors:

$$\bar{y} = \sum_{h=1}^{k} w_h \bar{y}_h \qquad (1)$$

$$SE(\bar{y}) = \sqrt{\sum_{h=1}^{H} w_h^2 \frac{s_h^2}{n_h}} \quad (2)$$

where $\bar{y}_h$ and $s_h^2$ depict the mean and variance, and $w_h$ and $n_h$ depict the weight and sample size of stratum *h*, respectively. Stratum weights were computed from digital maps showing forest/non-forest, forest type groups and ecological subsections (described in section 2.5). In case of simple estimation, *H* and $w_h$ are equal to 1.

Like FIA, GLAS features systematic sampling. In contrast to airborne lidar instruments, GLAS does not provide images of canopy height, but takes samples along transects (orbit tracks) every 172 m (center to center footprint spacing). The orbit tracks form a non-orthogonal grid spaced 14.5 km at the equator and 7 km at 60 degrees latitude (Zwally et al. 2002). GLAS waveforms tend to saturate under cloudy sky conditions. Hence, the actual sampling density of cloud-free waveforms can vary.

We applied the GLAS height algorithms from Lefsky et al. (2007) to 18,346 cloud-free waveforms in the Cascade region and to 24,050 waveforms in the Appalachian study region acquired between October 2003 and November 2006. However, GLAS samples are very dense along transects such that neighboring observations could be spatially autocorrelated. The occurrence of spatial autocorrelation would violate the assumption of independence among samples. As a result the sample mean computed for GLAS-derived forest height and biomass would not be an unbiased estimator and the variance would be underestimated. To avoid spatial autocorrelation we randomly selected waveforms with the requirement to be at least 2 km apart. We determined the minimum spacing based on the range of a

semivariogram from GLAS height estimates for each study region. We repeated sampling of the GLAS data 1000 times for simple and stratified random estimation and calculated the mean of each sample using the same formulas as we used for FIA. Thus, repeated sampling allowed us to construct a sampling distribution of the mean parameter. Mean and standard deviation of that sampling distribution were used as estimates of the population mean (equation 3) and its standard error (equation 4).

$$\bar{y}_G = \frac{1}{R}\sum_{r=1}^{R}\bar{y}_r \tag{3}$$

$$SE(\bar{y}_G) = \sqrt{\frac{1}{R-1}\sum_{r=1}^{R}(\bar{y}_r - \bar{y}_G)^2} \tag{4}$$

where $\bar{y}_r$ is a simple or stratified mean for sample *r*; and $\bar{y}_G$ and $SE(\bar{y}_G)$ is an estimate of the population mean and standard error derived from *R=1000* samples, respectively.

*Regional applicability of GLAS height algorithms*

We evaluated the regional performance of GLAS height algorithms by comparing the frequency distributions of GLAS-estimated heights and FIA-estimated dcd-heights for each study region. GLAS heights less than 2 m and over non-forest land (National Land Cover Data, NLCD 2001) were omitted from the analysis. The NLCD is a Landsat-based land cover map for the conterminous United States that

distinguishes these forest classes: deciduous, evergreen and mixed forests as well as woody wetlands. We further compared mean FIA and GLAS estimates by forest type group and ecological subsection. The two variables are reported in the FIA database and are also available as geospatial data layers. We used the latter to determine the respective membership for the GLAS samples.

Forest type groups are aggregations of forest types into ecological groupings (Eyre 1980). The Forest Service distinguishes 28 national forest type groups. Forest types are specified by FIA for each plot condition. Since we analyzed only plots with unique condition classes, each plot was associated with a single forest type group. The corresponding spatial layer is a thematic map produced with data from the Moderate Resolution Imaging Spectroradiometer (MODIS) acquired in 2001 at a spatial resolution of 250 m (http://svinetfc4.fs.fed.us/) (Ruefenacht et al. in press). We used this map to determine forest type groups for the GLAS samples.

While the comparison by forest type groups explores the effect of tree species and species groups on the performance of the GLAS algorithm, the meaning of ecological subsections is more complex. Ecological subsections are geographic regions of similar surficial geology, lithology, geomorphic process, soil groups, subregional climate, and potential natural communities (USDA Forest Service 2006). A vector layer is available for download at http://ncrs2.fs.fed.us/4801/fiadb/fiadb_documentation/FIADB_DOCUMENTATION.htm (last accessed August 3, 2007).

Assuming FIA provided an unbiased estimate of the mean height for each stratum, we calculated the bias of the GLAS estimate as the difference between the

GLAS and FIA mean. To obtain reliable stratified estimates, a minimum of five plots per stratum are recommended by FIA (McRoberts 2006). Therefore, forest type groups with fewer than five plots were omitted from this analysis (see Table 2.3 for sample sizes). No ecological subsections contained less than 5 plots.

For each ecological subsection we determined mean slope and elevation from a 30-m digital elevation model to explore potential biases introduced by differences in terrain conditions. Also, given the size of the study regions, there may be biases that are indirectly related to the distance of the GLAS samples from the source of the height calibration dataset. To test for this effect, the proximity of each ecological subsection to the training site for the height calibration dataset was calculated (as the Euclidean distance between the geographic center of each subsection and the geographic center of the subsection containing the training data). Proximity was only calculated for the Appalachian study region. Due to the spatial arrangement and low number of ecological subsections in the Cascades region, a proximity measure was not meaningful.

*Regional applicability of height-biomass allometry*

The regional applicability of height-based allometric equations to predict forest biomass largely depends on the site-specific variability in the relationship between plot-level height and biomass. For each study region, we developed height-based biomass regression models using FIA plot data, and then tested the region of model

applicability (as for the height algorithms) using information on forest type groups and ecological subsections.

To develop a basic height-biomass model, we performed a logarithmic transformation of the predictor and response variables and applied a linear regression model of the form:

$$f(AGBM) = \exp(\beta_0 + \beta_1 \ln height) \quad (5)$$

where

*AGBM* = total aboveground (oven-dry weight) biomass (Mg $ha^{-1}$) including all live trees (dbh 2.5 cm and larger)

*height* = maximum/mean/dcd-height (m)

exp = exponential function

ln = log base e (2.718282)

We then included the factors forest type group and ecological subsection using indicator variables and multiple linear regression models with and without interaction terms. Adding these variables permitted the mean response (biomass) in the models to vary with different levels of the factors. We evaluated all models based on several statistics: the model's coefficient of determination, and the root-mean-squared-error (RMSE) and bias of the predicted versus observed values. Whereas the coefficient of determination is a measure of the strength of the model fit, RMSE and bias quantify the model's prediction accuracy. These were calculated as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{i.predicted} - y_{i.observed})^2} \quad (6)$$

$$Bias = \bar{y}_{predicted} - \bar{y}_{observed} \quad (7)$$

When factors are included in regression models, the sample is partitioned into groups according to the number of factor levels. As a result some of the groups may contain only few observations. To assure an adequate sample size for each factor level, we excluded forest type groups and ecological subsections with fewer than ten observations. As a result, four out of ten and seven out of ten forest type groups remained for analysis in the Cascade and Appalachian study region, respectively (see Table 2.3 for sample sizes). In addition, one of the thirty three ecological subsections located in the Appalachian region was removed.

We developed and tested regression models based on mean dcd-height, maximum and mean height, respectively. Since GLAS provides estimates of dcd-height, empirical relationships between biomass and this height metric were of primary interest. However, we also developed models using maximum and mean height as predictor variables and compared their performance.

*Regional estimates of forest biomass*

To estimate forest biomass for each GLAS sample, we calibrated plot-level height-biomass models using FIA data and applied the "best" regional model to the GLAS estimates of height. Then we calculated and compared mean estimates of forest biomass from GLAS and FIA data using simple and stratified estimation by forest type groups and ecological subsections. Forest area for the ecological subsections was determined from the NLCD 2001.

Since a minimum of five plots per stratum is considered necessary for reliable stratified estimates (McRoberts 2006), we combined strata with fewer than five plots with ecologically similar strata. For FIA estimates in the Cascade region, we combined the western white pine and the other western softwoods group with the ponderosa pine group, and the alder/maple and tanoak/laurel group with the other western hardwoods group. In the Appalachian region it was necessary to pool together the longleaf/slash pine group with the loblolly/shortleaf pine group. For stratified GLAS estimates we combined the California mixed conifer group and the ponderosa pine group with the lodgepole pine group, the western oak group with the alder/maple group and the longleaf/slash pine group with the loblolly/shortleaf pine group.

## Results

### *Regional applicability of GLAS height algorithms*

Repeated random sampling from the GLAS data resulted for each study region in 1000 subsamples with an average sample size of 863 and 1948 shots for the Cascade and Appalachian region, respectively. The FIA sample consisted of 362 forest plots in the Cascades and 3054 forest plots in the Appalachian study region. Comparison of the frequency distributions of GLAS heights, which show the mean frequency of 1000 sample distributions, with frequency distributions from FIA suggests that the GLAS algorithms are generally accurate predictors of mean dcd-height (Figure 2.2).

In the Cascade region, the GLAS estimate of mean forest height was 25.3 m (± 0.5 m 95% confidence interval), which is 2.4 m lower than the FIA estimate of the mean (27.7 m ± 1.4 m). The agreement is higher between the medians of the two distributions (difference of 0.8 m), and also very good between the maximum heights (FIA: 65.7 m, GLAS mean of all subsample maxima: 64.7 m, GLAS maximum of all waveforms: 69.7 m). However, the frequency of FIA heights was higher in the height range between 30 m and 55 m compared to GLAS.

In the Appalachian region the GLAS estimate of mean forest height was 3.3 m lower than the FIA estimate (GLAS: 17.3 m ± 0.2 m; FIA: 20.6 m ± 0.2 m). Both frequency distributions are symmetric and exhibit similar standard deviations (FIA: 5.8 m, GLAS: 6.5 m). The maximum observed height of the 1000 GLAS subsamples

was 51.6 m (or 58.5 m when all GLAS waveforms are considered). Compared to FIA (39.8 m), this is a difference of 11.8 m. However, the majority of height estimates were within a similar range. 98% of the height estimates were between 3.4 m - 34.0 m (GLAS) and between 4.6 m – 33.4 m (FIA), respectively.

Examination of GLAS height distributions sampled by forest type groups revealed no substantive bias associated with tree species composition (Table 2.3). In the Cascade region, differences between GLAS and FIA heights varied from -5.4 m to -1.7 m for forest types with more than 5 plots. The bias of the Douglas-fir group, which was associated with the GLAS training plots, is nearly equivalent to the average bias across forest types (-4.5 m). Largest differences were observed in forest type groups with small sample sizes from FIA and GLAS (hemlock/Sitka spruce group and lodgepole pine group). The results are similar in the Appalachian region with biases varying between -6.4 m and 0.1. In comparison, biases of the forest type groups associated with the training plots were -1.6 m and -4.0 m (oak/pine group and oak/hickory forest group, respectively). Again, largest differences occurred in groups with small sample sizes (oak/gum/cypress group and elm/ash/cottonwood group).

Analyzing the mean differences between FIA and GLAS-estimated heights by ecological subsection revealed no association between biases and median topographic slope, median elevation or proximity to training data set (Figure 2.3 a-c). In the ecological subsection containing the training plots of the Appalachian region for example, GLAS estimated heights were on average -3.4 m lower than FIA heights. In comparison, we determined a difference of -4.0 m for the most distant subsection. Again, the largest biases occurred in subsections with limited FIA plot representation

(as in the forest type groups). These subsections also exhibited the largest standard errors in the GLAS and FIA estimates (Figure 2.3 d-f).

*The regional applicability of height-biomass models*

Empirical analysis of the height-biomass relationships revealed that the general region-specific allometic equations predicted biomass with accuracies comparable to forest-type-specific and ecological-subsection-specific equations. All regression models performed similarly across the study regions, in that there were only slight differences in model fit and prediction accuracy by factor and factor level (Table 2.4). The coefficient of determination of models using dcd-height (mean height of dominant/co-dominant trees), for example, varied between 0.74 and 0.77 in the Cascades study region, and between 0.60 and 0.64 in the Appalachian study region. Model fit and prediction accuracy improved in both regions when variations in forest type and ecological subsection were taken into account; however, these improvements were minimal. For example, RMSE's decreased between 6.6 - 12.2 $Mg \cdot ha^{-1}$ in the Cascades region and between 2.9 - 3.8 $Mg \cdot ha^{-1}$ in the Appalachian region, depending on the height metric used. These amounts correspond to less than 1% of the maximum observed aboveground biomass in each study area (1455.2 $Mg \cdot ha^{-1}$ and 541.1 $Mg \cdot ha^{-1}$ in the Cascades and Appalachians, respectively).

The performance of models using dcd-height was similar to models based on mean or maximum height. However, differences were greater in the Cascade region with higher biomass forests than in the Appalachian region. In the Cascade region,

prediction accuracies of maximum-height-models were between 0.8 and 6.0 $Mg \cdot ha^{-1}$ better than the accuracies of dcd-height-models. Conversely, maximum height produced biases two to three times as high as the biases observed with dcd-height. Finally, models using mean height were least accurate with RMSE values between 32.5 and 41.4 $Mg\ ha^{-1}$ greater than the RMSE values for dcd-height models. In the Appalachian region, differences between all height metrics were minimal, and did not exceed 1.7 $Mg \cdot ha^{-1}$.

*Regional biomass estimates*

To obtain regional estimates of mean aboveground biomass, we applied the regional biomass models based on FIA data to the GLAS heights. Since accounting for variations in ecological subsections and forest type groups did not improve the prediction accuracy of the models significantly, we proceeded with the simplest model that used mean dcd-height as a single predictor variable. This time, however, we included plots from all forest type groups and ecological subsections. The resulting models explained 74% and 60% of the variation in the FIA biomass data in the Cascade and Appalachian region, respectively. The equations are as follows (standard errors in square brackets):

$$AGBM_C = \exp(0.156\ [0.168] + 1.665\ [0.052] \cdot \ln \textit{dcd-height}) \quad (8)$$

$$AGBM_A = \exp(-0.484\ [0.079] + 1.777\ [0.026] \cdot \ln \textit{dcd-height}) \quad (9)$$

where:

| | | |
|---|---|---|
| $AGBM_C$ and $AGBM_A$ | = | total aboveground (oven-dry weight) biomass (Mg $ha^{-1}$) for the Cascade and Appalachian region, respectively |
| *dcd-height* | = | mean height of the dominant/co-dominant trees (m) |
| exp | = | exponential function |
| ln | = | logarithm base e (2.718282) |

The Cascade model had a bias of -2.66 Mg $ha^{-1}$ and yielded an RMSE of 174.7 Mg $ha^{-1}$, which corresponds to 12.0% of the maximum observed biomass value. In comparison, the bias of the Appalachian model was -7.52 Mg $ha^{-1}$ with an RMSE of 60.6 Mg $ha^{-1}$ or 11.2% of the maximum biomass of this region (Figure 2.4).

Table 2.5 shows the regional estimates of biomass from GLAS and FIA data based on simple and stratified estimation. Simple estimates of biomass from GLAS were 40.9 and 58.2 Mg $ha^{-1}$ lower than estimated by FIA in the Appalachian and Cascade region, respectively. Stratification by forest type or ecological subsection did not decrease standard errors of the estimates or the discrepancy between FIA and GLAS significantly. The results are in agreement with our previous findings (section 3.1), that differences between FIA- and GLAS-estimated heights are not explained by this type of stratification.

**Discussion**

In this study, we validated the regional applicability of height algorithms for the GLAS sensor. The height algorithms used here correspond to two training sites located in the Cascade and Appalachian region in the U.S. According to Lefsky et al. (2007), GLAS algorithms explained 90% (Cascades) and 40% (Appalachians) of the variance in the training data with an RMSE of 5.91 m (Cascades) and 4.86 m (Appalachians), respectively. The results of our study show good agreement between GLAS and FIA dcd-heights, despite the relatively few plots available for direct calibration of GLAS height algorithms. GLAS heights are on average 2 - 3 m lower than FIA heights. We were not able to detect any patterns of disagreement associated with forest type, mean slope, elevation or proximity of training data. Further research will need to focus more specific on the GLAS waveforms and their response to varying forest structure and topographic conditions. This will require additional training sites (potentially with high resolution lidar data) and simulation exercises (Yong et al. in prep).

To explore the regional accuracy of GLAS-estimated forest heights, we built on an approach that compared regional distributions of forest heights with data from the U.S. Forest Service FIA program. We choose this approach because it had a large geographic scope, and it provided inferences on an application-oriented level. However, the results of this study were not only affected by the accuracy of the height algorithms, but also by other factors inherent to the GLAS data (e.g. the temporal and spatial consistency) and the sampling scheme. An additional weakness of this region-

level analysis was its limited flexibility in linking the uncertainties of the height estimates to potentially important environmental factors (e.g. stand structure, terrain conditions). To obtain site-specific inferences will require reference measurements coincident with GLAS samples. However, since neither field nor high resolution lidar data are currently available for extensive areas, such a study would be confined to small geographic regions.

We explored the regional applicability of height-biomass allometic equations and found that regional models based on height as a single predictor variable performed equally well compared to models that accounted for variations in forest types and ecological subsections. This suggests that generalized, non-site and non-species-specific allometric equations can be useful for large-scale estimation of forest biomass. Our results are consistent with former studies from Lefsky et al. (2002, 2005b) who reported that height-biomass relationships were robust across the investigated forest types and even major biomes. Similarly, Mette et al. (2003) compared height-based equations for two hardwood and two coniferous tree species using German forest yield tables and found that the overall variability in the equations between species and stand ages was negligible (less than 15%). In this study we developed separate equations for study regions dominated by coniferous and broadleaf forests. Thus, we did not attempt to draw inferences across these two types of forests. Nevertheless, the results are promising as it is unlikely that accurate and global fine-scale maps for tree species composition will be available in the near future. We do anticipate that stand density, canopy cover and perhaps a classification by leaf-type might reduce the residual variance observed in our biomass models. However, more

research is required to quantify these effects. Conventional remote sensing has been successful in retrieving such two-dimensional variables. Future research should focus on possible synergies of GLAS with other sensor technologies.

This study showed that current GLAS algorithms for forest height can be extended to the regional scale with good accuracy. However, differences between GLAS and FIA estimates of forest biomass were still significant. Since the error associated with the allometric equations was relatively small more accurate height estimates are needed for regional biomass inventories. Other sources of error are related to the measurements and allometric equations at the tree-level which may propagate to the plot and landscape level; but they are not specific to GLAS and thus were beyond the scope of this study.

GLAS is the first spaceborne lidar sensor relevant for vegetation studies. As algorithms for vegetation height are continued to be refined, GLAS will provide a unique opportunity for large-scale observations of forest structure and biomass. In addition, forest research with GLAS will inform policy makers and earth scientists to guide future space missions. The Committee on Earth Science and Applications from Space: A Community Assessment and Strategy for the Future and the National Research Council (2007) have included in their recommendation for the next decade of space missions two operations that might support lidar observations of forests. Existing inventories programs like the FIA will be an important tool to provide the field observations necessary for validation.

**References**

Abshire, J.B., X.L. Sun, H. Riris, J.M. Sirota, J.F. McGarry, S. Palm, D.H. Yi, and P. Liiva. 2005. Geoscience Laser Altimeter System (GLAS) on the ICESat mission: On-orbit measurement performance. *Geophys. Res. Lett.* 32(21):doi:10.1029/2005GL024028.

Bechtold, W.A., and C.T. Scott. 2005. The Forest Inventory and Analysis plot design. P. 27-42 in *The Enhanced Forest Inventory and Analysis program - National sampling design and estimation procedures*, Bechtold, W.A., and P.L. Patterson (eds.). USDA For. Serv. Gen. Tech. Rep. SRS-80. 98 p.

Cochran, W.G. 1977. *Sampling techniques*. Third edition.Wiley, New York. 428 p.

Committee on Earth Science and Applications from Space: A Community Assessment and Strategy for the Future, and National Research Council. 2007. *Earth science and applications from space: National imperatives for the next decade and beyond*. The National Academies Press, Washington, D.C. 456 p.

Denman, K.L., G. Brasseur, A. Chidthaisong, P. Ciais, P.M. Cox, R.E. Dickinson, D. Hauglustaine, C. Heinze, E. Holland, D. Jacob, U. Lohmann, S. Ramachandran, P.L. da Silva Dias, S.C. Wofsy, and X. Zhang. 2007. Couplings between changes in the climate system and biogeochemistry. P. 499-587 in *Climate change 2007: The physical science basis*. *Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change*, Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, and H.L. Miller (eds.). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Dong, J.R., R.K. Kaufmann, R.B. Myneni, C.J. Tucker, P.E. Kauppi, J. Liski, W. Buermann, V. Alexeyev, and M.K. Hughes. 2003. Remote sensing estimates of boreal and temperate forest woody biomass: carbon pools, sources, and sinks. *Remote Sens. Environ.* 84(3):393-410.

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

Eyre, F.H. 1980. *Forest cover types of the United States and Canada*. Society of American Foresters. 148.

Goodale, C.L., M.J. Apps, R.A. Birdsey, C.B. Field, L.S. Heath, R.A. Houghton, J.C. Jenkins, G.H. Kohlmaier, W. Kurz, S.R. Liu, G.J. Nabuurs, S. Nilsson, and A.Z.

Shvidenko. 2002. Forest carbon sinks in the northern hemisphere. *Ecol. Appl.* 12(3):891-899.

Hansen, M. 2002. Volume and biomass estimation in FIA: national consistency vs. regional accuracy. P. 109-120 in *Proc. of the third annual forest inventory and analysis symposium*, McRoberts, R.E., G.A. Reams, P.C. Van Deusen, J.W. Moser (eds.). USDA For. Serv. 216 p.

Harding, D.J., and C.C. Carabajal. 2005. ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure. *Geophys. Res. Lett.* 32(21):doi:10.1029/2005GL023471.

Houghton, R.A. 2005. Aboveground forest biomass and the global carbon balance. *Glob. Change Biol.* 11(6):945-958.

Imhoff, M.L., S. Carson, and P. Johnson. 1998. A low-frequency radar experiment for measuring vegetation biomass. *IEEE T. Geosci. Remote.* 36(6):1988-1991.

Jenkins, J.C., D.C. Chojnacky, L.S. Heath, and R.A. Birdsey. 2003. National-scale biomass estimators for United States tree species. *For. Sci.* 49(1):12-35.

Labrecque, S., R.A. Fournier, J.E. Luther, and D. Piercey. 2006. A comparison of four methods to map biomass from Landsat-TM and inventory data in western Newfoundland. *Forest Ecol. Manag.* 226(1-3):129-144.

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

Lefsky, M.A., D. Harding, W.B. Cohen, G. Parker, and H.H. Shugart. 1999. Surface lidar remote sensing of basal area and biomass in deciduous forests of eastern Maryland, USA. *Remote Sens. Environ.* 67(1):83-98.

Lefsky, M.A., D.J. Harding, M. Keller, W.B. Cohen, C.C. Carabajal, F.D. Espirito-Santo, M.O. Hunter, R. de Oliveira, and P.B. de Camargo. 2005a. Estimates of forest canopy height and aboveground biomass using ICESat. *Geophys. Res. Lett.* 32(22):doi:10.1029/2005GL023971.

Lefsky, M.A., A.T. Hudak, W.B. Cohen, and S.A. Acker. 2005b. Geographic variability in lidar predictions of forest stand structure in the Pacific Northwest. *Remote Sens. Environ.* 95(4):532-548.

Lefsky, M.A., M. Keller, Y. Pang., P.B. de Camargo, M.O. Hunter. 2007. Revised method for forest canopy height estimation from Geoscience Laser Altimeter System waveforms. *J. Appl. Rem. Sens.* 1(1): 1(1):013537-18

Lu, D.S. 2006. The potential and challenge of remote sensing-based biomass estimation. *Int. J. Remote Sens.* 27(7):1297-1328.

McNab, W.H. and P.E. Avers. 1994. *Ecological subregions of the United States: section descriptions*. USDA For. Serv. WO-WSA-5. Available online at http://www.fs.fed.us/land/pubs/ecoregions/; last accessed Jan. 26, 2007.

McRoberts, R.E., W.A. Bechtold, P.L. Patterson, C.T. Scott, and G.A. Reams. 2005. The enhanced Forest Inventory and Analysis program of the USDA Forest Service: Historical perspective and announcement of statistical documentation. *J. Forest.* 103(6):304-308.

McRoberts, R.E. 2006. A model-based approach to inventory stratification. P. 126 in *Proc. of the sixth annual forest inventory and analysis symposium*, McRoberts, R.E., G.A. Reams, P.C. Van Duesen, and W.H. McWilliams (eds.). USDA For. Serv. 132 p.

Mette, T., K.P. Papathanassiou, I. Hajnsek, and R. Zimmermann. 2003. Forest biomass estimation using polarimetric SAR interferometry. in *Proc. of the workshop on POLinSAR – Applications of SAR polarimetry and polarimetric interferometry (ESA SP-529)*. 14-16 January 2003, Frascati, Italy. Published on CDROM., p.23.1

Ranson, K.J., G. Sun, R.H. Lang, N.S. Chauhan, R.J. Cacciola, and O. Kilic. 1997. Mapping of boreal forest biomass from spaceborne synthetic aperture radar. *J. Geophys. Res.* 102(D24):29599-29610.

Ruefenacht, B., M.V. Finco, M.D. Nelson, R. Czaplewski, E.H. Helmer, J.A. Blackard, G.R. Holden, A.J. Lister, D. Salajanu, and D. Weyermann., K. Winterberger. in press. Conterminous US and Alaska forest type mapping using Forest Inventory and Analysis data. *Photogramm. Eng. Rem. S.*

Scott, C.T., W.A. Bechtold, G.A. Reams, W.D. Smith, J.A. Westfall, M.H. Hansen, and G.G. Moisen. 2005. Sample-based estimators used by the Forest Inventory and Analysis National Information Management System. P. 43-67 in *The Enhanced Forest Inventory and Analysis program - National sampling design and estimation procedures*, Bechtold, W.A., and P.L. Patterson (eds.). USDA For. Serv. Gen. Tech. Rep. SRS-80. 98 p.

Treuhaft, R.N., B.E. Law, and G.P. Asner. 2004. Forest attributes from radar interferometric structure and its fusion with optical remote sensing. *Bioscience* 54(6):561-571.

Turner, D.P., W.B. Cohen, R.E. Kennedy, K.S. Fassnacht, and J.M. Briggs. 1999. Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites. *Remote Sens. Environ.* 70(1):52-68.

USDA Forest Service. 2006. *The Forest Inventory and Analysis database: Database description and users guide version 2.1*. P. 211. Available online at http://fia.fs.fed.us/; last accessed Jan. 26, 2007.

Woodhouse, I.H. 2006. Predicting backscatter-biomass and height-biomass trends using a macroecology model. *IEEE T. Geosci. Remote.* 44(4):871-877.

Zwally, H.J., B. Schutz, W. Abdalati, J. Abshire, C. Bentley, A. Brenner, J. Bufton, J. Dezio, D. Hancock, D. Harding, T. Herring, B. Minster, K. Quinn, S. Palm, J. Spinhirne, and R. Thomas. 2002. ICESat's laser measurements of polar ice, atmosphere, ocean, and land. *J. Geodyn.* 34(3-4):405-445.

Zwally, H.J., R. Schutz, C. Bentley, J. Bufton, T. Herring, J. Minster, J. Spinhirne, and R. Thomas. 2003. *GLAS/ICESat L2 Antarctic and Greenland Ice Sheet Altimetry Data V001*. Boulder, CO: National Snow and Ice Data Center. Digital media.

**Figures**

Figure 2.1. Cascade (Western U.S.) and Appalachian study region (Eastern U.S.)

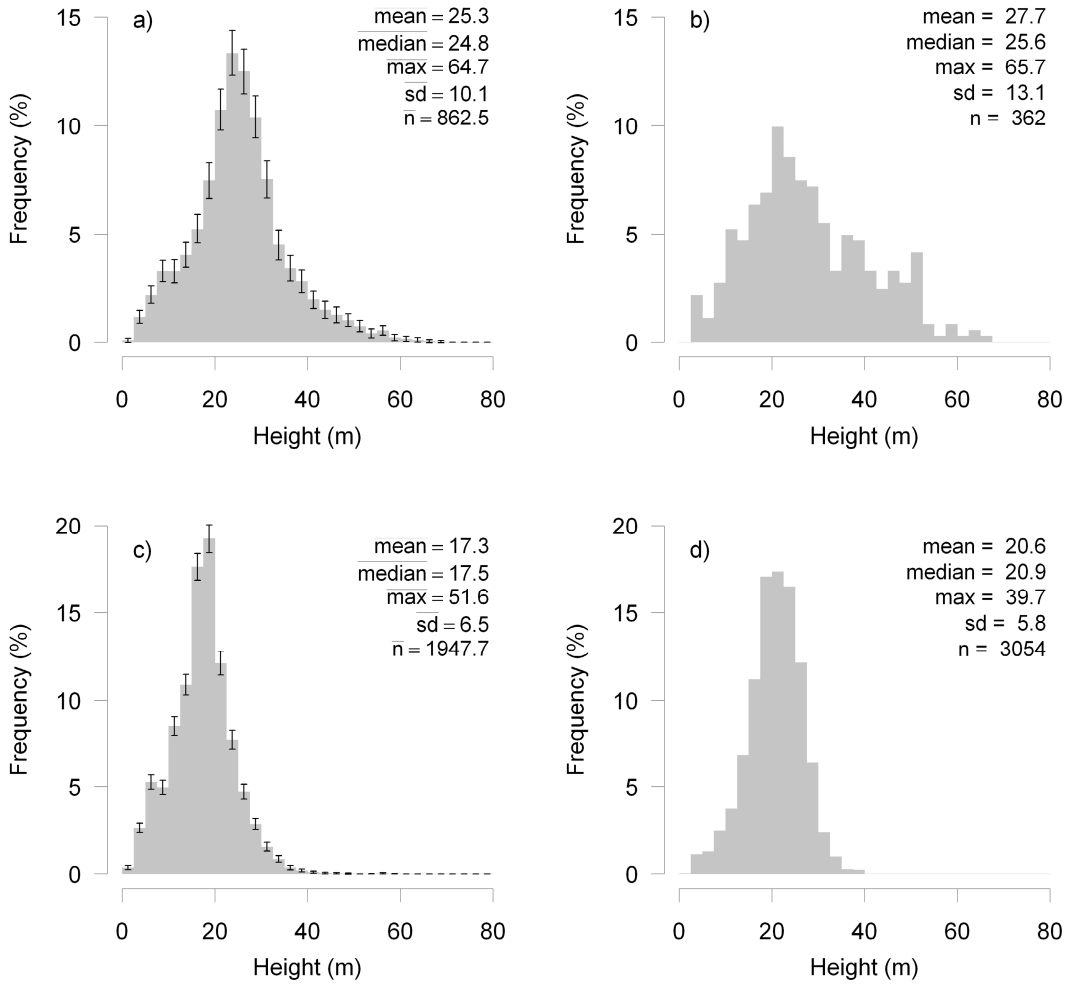


Figure 2.2. Frequency distributions for forest height in the Cascade (a, b) and Appalachian study region (c, d). Figures a) and c) show the mean frequency and the standard error for 1000 repeated, random GLAS samples. Estimates of the mean, median, maximum (max) and standard deviation (sd) depict the mean of these estimates across the 1000 samples which have a mean sample size of size = $\overline{n}$ . Figures b) and d) show the frequency distributions of FIA samples based on the mean height of the dominant/co-dominant trees (dcd-height).

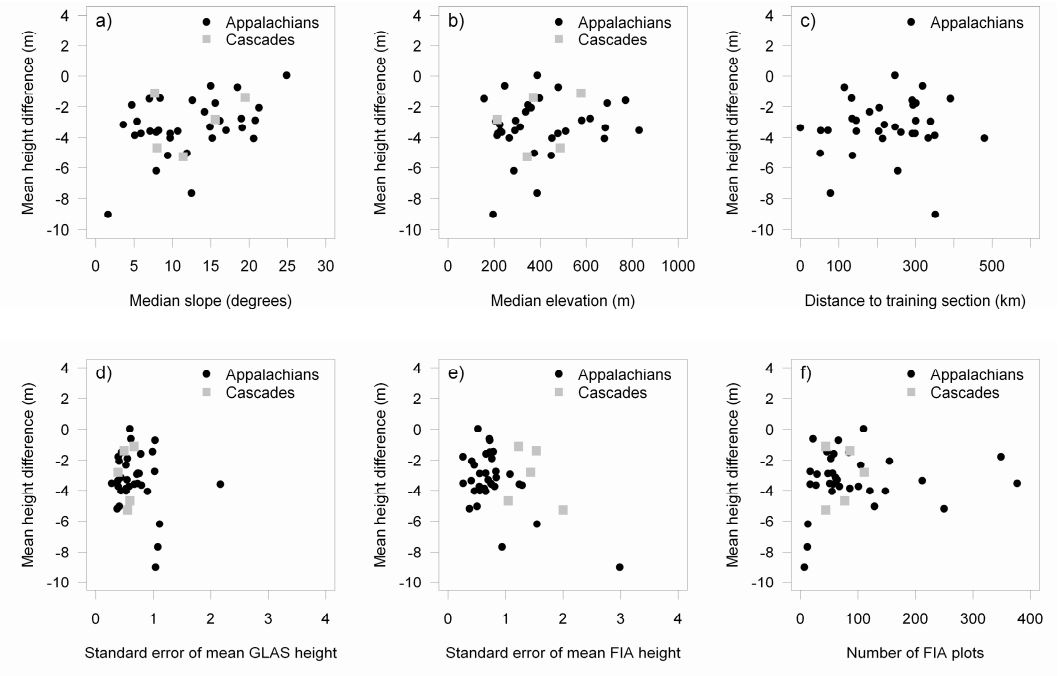


Figure 2.3. Scatterplots showing the bias (difference) between mean GLAS and mean FIA height for ecological subsections versus a) median slope and b) elevation of forested area, c) distance to the geographic center of the ecological subsection with the training data, d) standard error of mean GLAS height, e) standard error of mean FIA dcd-height, f) number of FIA plots.

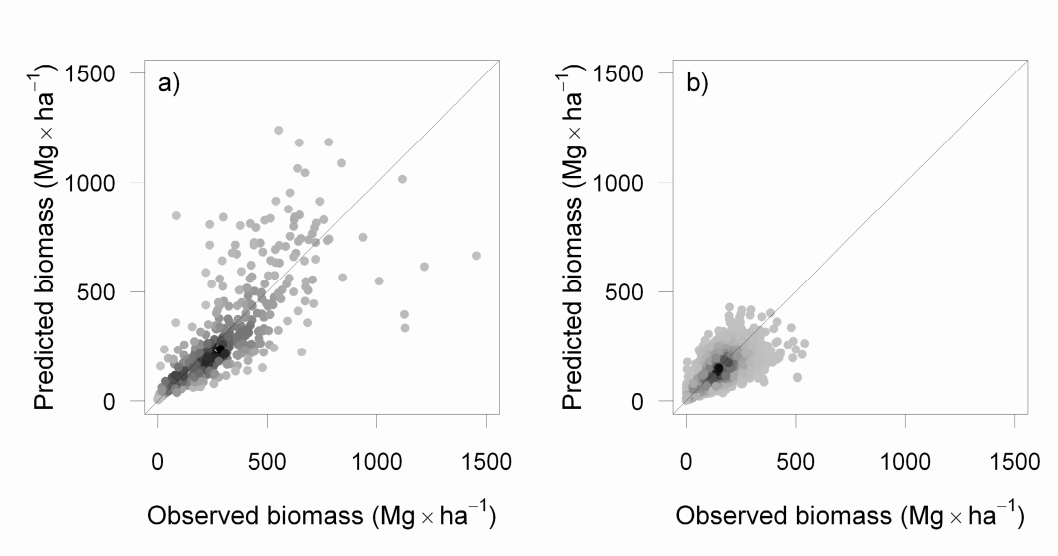


Figure 2.4. Observed versus predicted aboveground biomass (Mg ha$^{-1}$) for the Cascade (left, RMSE=174.7 Mg ha$^{-1}$, n = 362) and Appalachian study region (right, RMSE=60.6 Mg ha$^{-1}$, n=3054)

## Tables

Table 2.1. U.S. Forest Service ecological sections and subsections of the study regions

| Study region | Ecological section | Ecological subsections |
|---|---|---|
| Cascades | Western Cascades | West Cascade Slope Forest, Western Cascades Highland Forest, Cascade Crest Forest and Volcanic Peaks, Southern Oregon Cascades, Southern Oregon Cascade Highlands |
| Appalachians | Southern Unglaciated Allegheny Plateau | Teays Plateau, Kinniconick and Licking Knobs |
| | Northern Cumberland Plateau | Rugged Eastern Hills, Kinniconick and Licking Knobs, Southwestern Escarpment, Sequatchie Valley, Miami-Scioto Plain-Tipton Till Plain |
| | Central Ridge and Valley | Rolling Limestone Hills |
| | Southern Appalachian Piedmont | Schist Hills |
| | Southern Cumberland Plateau | Shale Hills and Mountain, Sandstone Plateau, Table Plateau, Sandstone Mountain, Moulton Valley, Southern Cumberland Valleys |
| | Southern Ridge and Valley | Chert Valley, Sandstone-Shale and Chert Ridge, Sandstone Ridge, Shaley Limestone Valley |
| | Northern Ridge and Valley | Ridge and Valley, Great Valley of Virginia |
| | Northern Cumberland Mountains | Western Coal Fields, Eastern Coal Fields, Black Mountains, Southern Cumberland Mountains, Pine and Cumberland Mountains |
| | Blue Ridge Mountains | Northern Blue Ridge Mountains, Central Blue Ridge Mountains, Southern Blue Ridge Mountains, Metasedimentary Mountains |

Table 2.2. Source and number of selected FIA plots

| State | Report year | Cycle | Subcycle | #Plots |
|---|---|---|---|---|
| Alabama | 2004 | 8 | 5/5 | 456 |
| Georgia | 2004 | 8 | 7/7 | 320 |
| Kentucky | 2004 | 5 | 5/5 | 603 |
| North Carolina | 2005 | 8 | 3/5 | 236 |
| Oregon | 2005 | 5 | 5/10 | 362 |
| South Carolina | 2001 | 3 | 5/5 | 30 |
| Tennessee | 2004 | 7 | 5/5 | 653 |
| Virginia | 2001 | 7 | 5/5 | 756 |

Table 2.3. Mean and standard error (SE) of FIA and GLAS heights by forest type group in the Cascade and the Appalachian study region. GLAS statistics are based on 1000 random samples with a mean sample size = #Shots. Bias is the difference between the mean GLAS and mean FIA estimate.

| Forest type group | Area ($10^3$ ha) | FIA dcd-height (m) | | | GLAS height (m) | | | Bias (m) |
|---|---|---|---|---|---|---|---|---|
| | | Mean | SE | #Plots | Mean | SE | #Shots | |
| *Cascades* | | | | | | | | |
| Douglas-fir | 2021.08 | 32.11 | 1.07 | 181 | 26.62 | 0.27 | 725.3 | -5.5 |
| Ponderosa pine | 17.53 | 22.60 | 3.90 | 7 | 13.60 | 1.54 | 3.5 | -9.0 |
| Western white pine | 0.10 | 30.08 | - | 1 | - | - | - | - |
| Fir/spruce/mountain hemlock | 657.86 | 23.48 | 0.78 | 122 | 21.80 | 0.41 | 238.9 | -1.7 |
| Lodgepole pine | 62.93 | 17.07 | 1.12 | 16 | 11.62 | 0.88 | 23.5 | -5.4 |
| Hemlock/Sitka spruce | 25.01 | 30.02 | 2.76 | 26 | 24.66 | 1.43 | 26.3 | -5.4 |
| Other western softwoods | 0.54 | 16.60 | - | 1 | - | - | - | - |
| California mixed conifer | 1.69 | - | - | - | 7.34 | 0.86 | 1.0 | - |
| Elm/Ash/Cottonwood | 0.04 | - | - | - | - | - | - | - |
| Aspen/Birch | 0.08 | - | - | - | - | - | - | - |
| Alder/maple | 0.69 | 17.81 | 5.24 | 3 | 20.89 | 0.48 | 4.0 | 3.1 |
| Western oak | 0.76 | - | - | - | 22.36 | 0.50 | 1.0 | - |
| Tanoak/laurel | 0.04 | 12.05 | 4.26 | 2 | - | - | - | - |
| Other western hardwoods | 0.07 | 10.12 | 4.51 | 3 | - | - | - | - |
| *Appalachians* | | | | | | | | |
| White/red/jack pine | 95.93 | 21.59 | 0.83 | 74 | 18.21 | 0.56 | 30.4 | -3.4 |
| Spruce/fir | 0.01 | 19.18 | 1.97 | 7 | - | - | - | - |
| Longleaf/slash pine | 9.96 | 15.65 | 3.71 | 4 | 9.99 | 1.16 | 2.0 | -5.7 |
| Loblolly/shortleafpine | 758.29 | 14.13 | 0.40 | 229 | 14.20 | 0.22 | 272.5 | 0.1 |
| Pinyon/juniper | 2.23 | 12.03 | 1.00 | 9 | - | - | - | - |
| Oak/pine | 427.99 | 17.86 | 0.34 | 277 | 16.22 | 0.22 | 211.9 | -1.6 |
| Oak/hickory | 12128.11 | 21.49 | 0.11 | 2348 | 17.52 | 0.10 | 1863.9 | -4.0 |
| Oak/gum/cypress group | 34.64 | 21.02 | 1.85 | 13 | 14.61 | 0.73 | 19.0 | -6.4 |
| Elm/ash/cottonwood | 21.24 | 21.02 | 1.06 | 13 | 15.75 | 0.87 | 9.0 | -5.3 |
| Maple/beech/birch | 100.28 | 22.96 | 0.49 | 80 | 16.68 | 0.54 | 16.7 | -6.3 |

Table 2.4. Comparison of biomass models that account for forest type group (fortypgrp) or ecological subsection (ecosubcd) or both with (*) and without (+) interaction. Plot heights are defined as height of the tallest tree (max-ht), mean height (mean-ht), and mean height of the dominant/co-dominant trees (dcd-ht). Model evaluation is based on the root-mean-squared-error (RMSE) and BIAS of the predicted (fitted) versus observed biomass values, and the $R^2$ of the regression model.

| Model terms | Cascades | | | Appalachians | | |
|---|---|---|---|---|---|---|
| | RMSE | BIAS | $R^2$ | RMSE | BIAS | $R^2$ |
| log(dcd-ht) | 173.13 | -3.43 | 0.74 | 60.58 | -7.53 | 0.60 |
| log(dcd-ht)*ecosubcd | 168.71 | -6.06 | 0.76 | 57.31 | -7.09 | 0.64 |
| log(dcd-ht)*fortypgrp | 165.52 | -6.04 | 0.76 | 60.84 | -7.54 | 0.60 |
| log(dcd-ht)*fortypgrp+log(dcd-ht)*ecosubcd | 164.60 | -7.16 | 0.77 | 58.02 | -6.79 | 0.64 |
| log(dcd-ht)+ecosubcd | 166.37 | -5.52 | 0.75 | 58.61 | -6.53 | 0.62 |
| log(dcd-ht)+fortypgrp | 164.96 | -6.08 | 0.76 | 61.13 | -7.38 | 0.60 |
| log(dcd-ht)+fortypgrp+ecosubcd | 162.69 | -6.34 | 0.76 | 59.49 | -6.30 | 0.63 |
| log(max-ht) | 167.13 | -15.57 | 0.73 | 61.49 | -8.94 | 0.59 |
| log(max-ht)*ecosubcd | 164.71 | -16.58 | 0.74 | 58.69 | -8.47 | 0.63 |
| log(max-ht)*fortypgrp | 164.75 | -18.01 | 0.74 | 61.95 | -8.72 | 0.59 |
| log(max-ht)*fortypgrp+log(max-ht)*ecosubcd | 162.72 | -17.52 | 0.74 | 59.69 | -7.86 | 0.64 |
| log(max-ht)+ecosubcd | 161.50 | -17.47 | 0.74 | 60.02 | -7.84 | 0.61 |
| log(max-ht)+fortypgrp | 164.15 | -17.49 | 0.74 | 62.11 | -8.77 | 0.59 |
| log(max-ht)+fortypgrp+ecosubcd | 160.52 | -17.76 | 0.74 | 61.10 | -7.53 | 0.62 |
| log(mean-ht) | 205.59 | -5.71 | 0.67 | 61.54 | -8.02 | 0.57 |
| log(mean-ht)*ecosubcd | 210.09 | -5.91 | 0.71 | 58.61 | -7.65 | 0.61 |
| log(mean-ht)*fortypgrp | 200.67 | -6.04 | 0.71 | 61.05 | -8.27 | 0.57 |
| log(mean-ht)*fortypgrp+log(mean-ht)*ecosubcd | 200.41 | -7.45 | 0.73 | 58.94 | -7.60 | 0.61 |
| log(mean-ht)+ecosubcd | 204.16 | -5.55 | 0.70 | 59.57 | -7.12 | 0.59 |
| log(mean-ht)+fortypgrp | 201.07 | -6.17 | 0.71 | 61.53 | -7.96 | 0.57 |
| log(mean-ht)+fortypgrp+ecosubcd | 197.87 | -6.21 | 0.71 | 59.98 | -7.02 | 0.59 |

Table 2.5. Means and standard errors (SE) of the regional biomass estimates from GLAS and FIA in Mg $ha^{-1}$ based on simple and stratified estimation

| **Region** | **Estimation method** | **Biomass FIA** | | **Biomass GLAS** | |
|---|---|---|---|---|---|
| | | Mean | SE | Mean | SE |
| Appalachians | simple, no strata | 148.06 | 1.33 | 107.20 | 1.06 |
| | stratified by eco-subsection | 148.92 | 1.29 | 109.19 | 1.14 |
| | stratified by forest type group | 150.98 | 1.35 | 106.80 | 1.02 |
| Cascades | simple, no strata | 334.03 | 11.75 | 275.87 | 4.61 |
| | stratified by eco-subsection | 336.50 | 12.06 | 270.91 | 4.07 |
| | stratified by forest type group | 352.85 | 13.85 | 272.90 | 4.11 |

# Chapter 3 - Conclusion

There are great uncertainties in the current and future ability of forest ecosystems to offset anthropogenic carbon emissions. New approaches are required to improve our knowledge on the distribution of plant biomass across the globe. While satellite remote sensing is perhaps the most appropriate tool for consistent and frequent observations of the Earth's surface, most existing sensors have failed to produce accurate estimates of forest biomass under closed-canopy conditions. Light detection and ranging (lidar) is capable of predicting forest biomass with high accuracies, but until recently sensor technologies have been limited to airborne systems.

With the Geoscience Laser Altimeter System (GLAS) onboard NASA's ICESat satellite, now global lidar observations have become available that could potentially be used to inventory global biomass stocks. With over 900 million laser observations globally (by June of 2005, Abshire et al. 2005), GLAS represents a unique opportunity for global carbon research. The objective of this study was to determine the utility of GLAS for large-scale biomass inventories. The utility of GLAS data for biomass inventories primarily depends on two factors: the accuracy of GLAS-estimated forest heights and the accuracy of allometric equations that are needed to transform estimates of forest height into estimates of biomass.

By comparing GLAS estimates with field data from the U.S Forest Service Inventory and Analysis (FIA) Program, this study found that locally trained height algorithms were applicable to larger regions without introducing regional bias. GLAS-estimated heights were generally accurate, but consistently lower than FIA estimates by 2-3 m. It can be expected that GLAS is not as accurate as high resolution airborne lidar (due to the larger footprint). However, more research on the effects of complex terrain and stand heterogeneity might improve future algorithms. Future research could include additional test sites and theoretical models based on radiative transfer theory (e.g. Koetz et al. 2006). Since the study regions in this research were selected to include complex terrain and high biomass forests, the results obtained here might represent the lower limit of accuracy that can be expected with GLAS.

This study investigated the uncertainty associated with general, non-species specific allometric equations to convert GLAS-estimates of vegetation height into estimates of biomass. Species-specific allometric equations are not feasible for large-scale studies. However, it is important to evaluate the trade-offs and potential biases associated with general equations. The research presented here showed that two general equations based on a single measure of stand height (for a conifer and broadleaf dominated study region, respectively) achieved comparable accuracies on the regional scale relative to equations that accounted for variations in forest type and other ecological strata. The accuracies of the allometric equations were similar to accuracies achieved with airborne simulations of the Vegetation Canopy Lidar mission, where biomass was directly estimated from waveform parameters (Drake et

al. 2002; Hyde et al. 2005). However, allometric error does not take into account the error associated with the GLAS height estimates.

This research demonstrated the potential of a space-borne lidar mission for large-scale biomass inventories. GLAS provides a first attempt to provide global observations of forest height and biomass. The results are promising, but further research is needed to improve current algorithms. Given the current lack of knowledge on the spatial distribution of forest biomass, it is not certain what the minimum accuracy requirements are to effectively reduce the error in current climate models. Since GLAS is a sampling instrument, its utility for biomass inventories does not only depend on the accuracy of individual observations, but also the sampling design and its ability to capture the spatial variability of forest biomass. Future research should focus on methods to combine GLAS samples with spatially continuous data from optical or radar remote sensing.

## Bibliography

Abshire, J.B., X.L. Sun, H. Riris, J.M. Sirota, J.F. McGarry, S. Palm, D.H. Yi, and P. Liiva. 2005. Geoscience Laser Altimeter System (GLAS) on the ICESat mission: On-orbit measurement performance. *Geophys. Res. Lett.* 32(21):doi:10.1029/2005GL024028.

Blair, J.B., D.L. Rabine, and M.A. Hofton. 1999. The Laser Vegetation Imaging Sensor: a medium-altitude, digitization-only, airborne laser altimeter for mapping vegetation and topography. *ISPRS J. Photogramm.* 54(2-3):115-122.

Bechtold, W.A., and C.T. Scott. 2005. The Forest Inventory and Analysis plot design. P. 27-42 in *The Enhanced Forest Inventory and Analysis program - National sampling design and estimation procedures*, Bechtold, W.A., and P.L. Patterson (eds.). USDA For. Serv. Gen. Tech. Rep. SRS-80. 98 p.

Blair, J.B., D.L. Rabine, and M.A. Hofton. 1999. The Laser Vegetation Imaging Sensor: a medium-altitude, digitization-only, airborne laser altimeter for mapping vegetation and topography. *ISPRS J. Photogramm.* 54(2-3):115-122.

Brenner, A., H.J. Zwally, C. Bentley, B. Csathó, D.J. Harding, M.A. Hofton, J.-B. Minster, L. Roberts, J.L. Saba, R.H. Thomas, and Y. Donghui. 2003. *Derivation of range and range distributions from laser pulse waveform analysis for surface elevations, roughness, slope, and vegetation heights*. Algorithm Theoretical Basis Document 4.1. Available online at http://www.csr.utexas.edu/glas/atbd.html; last accessed Sep. 15, 2007.

Chapin, F.S., P.A. Matson, and H.A. Mooney. 2002. *Principles of terrestrial ecosystem ecology*. Springer, New York. xiv, 436 p. p.

Cochran, W.G. 1977. *Sampling techniques*. Third edition.Wiley, New York. 428 p.

Committee on Earth Science and Applications from Space: A Community Assessment and Strategy for the Future, and National Research Council. 2007. *Earth science and applications from space: National imperatives for the next decade and beyond*. The National Academies Press, Washington, D.C. 456 p.

Denman, K.L., G. Brasseur, A. Chidthaisong, P. Ciais, P.M. Cox, R.E. Dickinson, D. Hauglustaine, C. Heinze, E. Holland, D. Jacob, U. Lohmann, S. Ramachandran, P.L. da Silva Dias, S.C. Wofsy, and X. Zhang. 2007. Couplings between changes in the climate system and biogeochemistry. P. 499-587 in *Climate change 2007: The physical science basis. Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change*,

Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, and H.L. Miller (eds.). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Dong, J.R., R.K. Kaufmann, R.B. Myneni, C.J. Tucker, P.E. Kauppi, J. Liski, W. Buermann, V. Alexeyev, and M.K. Hughes. 2003. Remote sensing estimates of boreal and temperate forest woody biomass: carbon pools, sources, and sinks. *Remote Sens. Environ.* 84(3):393-410.

Drake, J.B., R.O. Dubayah, D.B. Clark, R.G. Knox, J.B. Blair, M.A. Hofton, R.L. Chazdon, J.F. Weishampel, and S.D. Prince. 2002. Estimation of tropical forest structural characteristics using large-footprint lidar. *Remote Sens. Environ.* 79(2-3):305-319.

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

Drake, J.B., R.G. Knox, R.O. Dubayah, D.B. Clark, R. Condit, J.B. Blair, and M. Hofton. 2003. Above-ground biomass estimation in closed canopy Neotropical forests using lidar remote sensing: factors affecting the generality of relationships. *Global Ecol. Biogeogr.* 12(2):147-159.

Eyre, F.H. 1980. *Forest cover types of the United States and Canada*. Society of American Foresters. 148.

Goodale, C.L., M.J. Apps, R.A. Birdsey, C.B. Field, L.S. Heath, R.A. Houghton, J.C. Jenkins, G.H. Kohlmaier, W. Kurz, S.R. Liu, G.J. Nabuurs, S. Nilsson, and A.Z. Shvidenko. 2002. Forest carbon sinks in the northern hemisphere. *Ecol. Appl.* 12(3):891-899.

Hansen, M. 2002. Volume and biomass estimation in FIA: national consistency vs. regional accuracy. P. 109-120 in *Proc. of the third annual forest inventory and analysis symposium*, McRoberts, R.E., G.A. Reams, P.C. Van Deusen, J.W. Moser (eds.). USDA For. Serv. 216 p.

Harding, D.J., and C.C. Carabajal. 2005. ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure. *Geophys. Res. Lett.* 32(21):doi:10.1029/2005GL023471.

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

Houghton, R.A. 2005. Aboveground forest biomass and the global carbon balance. *Glob. Change Biol.* 11(6):945-958.

Hyde, P., R. Dubayah, B. Peterson, J.B. Blair, M. Hofton, C. Hunsaker, R. Knox, and W. Walker. 2005. Mapping forest structure for wildlife habitat analysis using waveform lidar: Validation of montane ecosystems. *Remote Sens. Environ*. 96(3-4):427-437.

Imhoff, M.L., S. Carson, and P. Johnson. 1998. A low-frequency radar experiment for measuring vegetation biomass. *IEEE T. Geosci. Remote.* 36(6):1988-1991.

Jenkins, J.C., D.C. Chojnacky, L.S. Heath, and R.A. Birdsey. 2003. National-scale biomass estimators for United States tree species. *For. Sci.* 49(1):12-35.

Koetz, B., F. Morsdorf, G. Sun, K.J. Ranson, K. Itten, and B. Allgöwer. 2006. Inversion of a lidar waveform model for forest biophysical parameter estimation. *IEEE T. Geosci. Remote. Lett.* 3(1):49-53.

Labrecque, S., R.A. Fournier, J.E. Luther, and D. Piercey. 2006. A comparison of four methods to map biomass from Landsat-TM and inventory data in western Newfoundland. *Forest Ecol. Manag.* 226(1-3):129-144.

Law, B.E., P.E. Thornton, J. Irvine, P.M. Anthoni, and S. Van Tuyl. 2001. Carbon storage and fluxes in ponderosa pine forests at different developmental stages. *Glob. Change Biol.* 7(7):755-777

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

Lefsky, M.A., D. Harding, W.B. Cohen, G. Parker, and H.H. Shugart. 1999. Surface lidar remote sensing of basal area and biomass in deciduous forests of eastern Maryland, USA. *Remote Sens. Environ.* 67(1):83-98.

Lefsky, M.A., D.J. Harding, M. Keller, W.B. Cohen, C.C. Carabajal, F.D. Espirito-Santo, M.O. Hunter, R. de Oliveira, and P.B. de Camargo. 2005a. Estimates of forest canopy height and aboveground biomass using ICESat. *Geophys. Res. Lett.* 32(22):doi:10.1029/2005GL023971.

Lefsky, M.A., A.T. Hudak, W.B. Cohen, and S.A. Acker. 2005b. Geographic variability in lidar predictions of forest stand structure in the Pacific Northwest. *Remote Sens. Environ.* 95(4):532-548.

Lefsky, M.A., M. Keller, Y. Pang., P.B. de Camargo, M.O. Hunter. 2007. Revised method for forest canopy height estimation from Geoscience Laser Altimeter System waveforms. *J. Appl. Rem. Sens.* 1(1): 1(1):013537-18

Lu, D.S. 2006. The potential and challenge of remote sensing-based biomass estimation. *Int. J. Remote Sens.* 27(7):1297-1328.

McNab, W.H. and P.E. Avers. 1994. *Ecological subregions of the United States: section descriptions*. USDA For. Serv. WO-WSA-5. Available online at http://www.fs.fed.us/land/pubs/ecoregions/; last accessed Jan. 26, 2007.

McRoberts, R.E., W.A. Bechtold, P.L. Patterson, C.T. Scott, and G.A. Reams. 2005. The enhanced Forest Inventory and Analysis program of the USDA Forest Service: Historical perspective and announcement of statistical documentation. *J. Forest.* 103(6):304-308.

McRoberts, R.E. 2006. A model-based approach to inventory stratification. P. 126 in *Proc. of the sixth annual forest inventory and analysis symposium*, McRoberts, R.E., G.A. Reams, P.C. Van Duesen, and W.H. McWilliams (eds.). USDA For. Serv. 132 p.

Mette, T., K.P. Papathanassiou, I. Hajnsek, and R. Zimmermann. 2003. Forest biomass estimation using polarimetric SAR interferometry. in *Proc. of the workshop on POLinSAR – Applications of SAR polarimetry and polarimetric interferometry (ESA SP-529)*. 14-16 January 2003, Frascati, Italy. Published on CDROM., p.23.1

Patenaude, G., R.A. Hill, R. Milne, D.L.A. Gaveau, B.B.J. Briggs, and T.P. Dawson. 2004. Quantifying forest above ground carbon content using LiDAR remote sensing. *Remote Sens. Environ.* 93(3):368-380.

Ranson, K.J., G. Sun, R.H. Lang, N.S. Chauhan, R.J. Cacciola, and O. Kilic. 1997. Mapping of boreal forest biomass from spaceborne synthetic aperture radar. *J. Geophys. Res.* 102(D24):29599-29610.

Ruefenacht, B., M.V. Finco, M.D. Nelson, R. Czaplewski, E.H. Helmer, J.A. Blackard, G.R. Holden, A.J. Lister, D. Salajanu, and D. Weyermann., K. Winterberger. in press. Conterminous US and Alaska forest type mapping using Forest Inventory and Analysis data. *Photogramm. Eng. Rem. S.*

Scott, C.T., W.A. Bechtold, G.A. Reams, W.D. Smith, J.A. Westfall, M.H. Hansen, and G.G. Moisen. 2005. Sample-based estimators used by the Forest Inventory and Analysis National Information Management System. P. 43-67 in *The Enhanced Forest Inventory and Analysis program - National sampling design and estimation procedures*, Bechtold, W.A., and P.L. Patterson (eds.). USDA For. Serv. Gen. Tech. Rep. SRS-80. 98 p.

Sun, G., K.J. Ranson, V.I. Khairuk, and K. Kovacs. 2003. Validation of surface height from shuttle radar topography mission using shuttle laser altimeter. *Remote Sens. Environ.* 88(4):401-411.

Treuhaft, R.N., B.E. Law, and G.P. Asner. 2004. Forest attributes from radar interferometric structure and its fusion with optical remote sensing. *Bioscience* 54(6):561-571.

Turner, D.P., W.B. Cohen, R.E. Kennedy, K.S. Fassnacht, and J.M. Briggs. 1999. Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites. *Remote Sens. Environ.* 70(1):52-68.

USDA Forest Service. 2006. *The Forest Inventory and Analysis database: Database description and users guide version 2.1*. P. 211. Available online at http://fia.fs.fed.us/; last accessed Jan. 26, 2007.

Wofsy, S.C., and R.C. Harriss. 2002. *The North American Carbon Program (NACP)*. Report of the NACP Committee of the U.S. Interagency Carbon Cycle Science Program US Global Change Research Program. 75.

Woodhouse, I.H. 2006. Predicting backscatter-biomass and height-biomass trends using a macroecology model. *IEEE T. Geosci. Remote.* 44(4):871-877.

Zwally, H.J., B. Schutz, W. Abdalati, J. Abshire, C. Bentley, A. Brenner, J. Bufton, J. Dezio, D. Hancock, D. Harding, T. Herring, B. Minster, K. Quinn, S. Palm, J. Spinhirne, and R. Thomas. 2002. ICESat's laser measurements of polar ice, atmosphere, ocean, and land. *J. Geodyn.* 34(3-4):405-445.

Zwally, H.J., R. Schutz, C. Bentley, J. Bufton, T. Herring, J. Minster, J. Spinhirne, and R. Thomas. 2003. *GLAS/ICESat L2 Antarctic and Greenland Ice Sheet Altimetry Data V001*. Boulder, CO: National Snow and Ice Data Center. Digital media.

# Appendix

Table A.1. Parameters for estimating aboveground biomass (source: Jenkins et al., 2003)

| | Species group | Parameters* | |
|---|---|---|---|
| | | $\beta_0$ | $\beta_1$ |
| Hardwood | Aspen/alder/cottonwood/willow | –2.2094 | 2.3867 |
| | Soft maple/birch | –1.9123 | 2.3651 |
| | Mixed hardwood | –2.4800 | 2.4835 |
| | Hard maple/oak/hickory/beech | –2.0127 | 2.4342 |
| Softwood | Cedar/larch | –2.0336 | 2.2592 |
| | Douglas-fir | –2.2304 | 2.4435 |
| | True fir/hemlock | –2.5384 | 2.4814 |
| | Pine | –2.5356 | 2.4349 |
| | Spruce | –2.0773 | 2.3323 |
| Woodland | Juniper/oak/mesquite | –0.7152 | 1.7029 |

* Biomass equation:

$$bm = \exp(\beta_{0+} \beta_1 \ln dbh)$$

where

*bm* = total aboveground biomass (kg) for trees 2.5 cm dhh and larger
*dbh* = diameter at breast height (cm)
exp = exponential function
ln = natural logarithm

Table A.2 Ecological subsections in the study regions

| Study region | Ecological subsections | |
|---|---|---|
| | **Code** | **Name** |
| Appalachians | 221Eb | Teays Plateau |
| | 221En | Kinniconick and Licking Knobs |
| | 221Ha | Rugged Eastern Hills |
| | 221Hb | Kinniconick and Licking Knobs |
| | 221Hc | Southwestern Escarpment |
| | 221Hd | Sequatchie Valley |
| | 221He | Miami-Scioto Plain-Tipton Till Plain |
| | 221Ja | Rolling Limestone Hills |
| | 221Jb | Sandstone Hills |
| | 221Jc | Holston Valley |
| | 231Ag | Schist Hills |
| | 231Ca | Shale Hills and Mountain |
| | 231Cb | Sandstone Plateau |
| | 231Cc | Table Plateau |
| | 231Cd | Sandstone Mountain |
| | 231Ce | Moulton Valley |
| | 231Cf | Southern Cumberland Valleys |
| | 231Cg | Sequatchie Valley |
| | 231Da | Chert Valley |
| | 231Db | Sandstone-Shale and Chert Ridge |
| | 231Dc | Sandstone Ridge |
| | 231De | Shaley Limestone Valley |
| | M221Aa | Ridge and Valley |
| | M221Ab | Great Valley of Virginia |
| | M221Ca | Western Coal Fields |
| | M221Cb | Eastern Coal Fields |
| | M221Cc | Black Mountains |
| | M221Cd | Southern Cumberland Mountains |
| | M221Ce | Pine and Cumberland Mountain |
| | M221Da | Northern Blue Ridge Mountains |
| | M221Db | Central Blue Ridge Mountains |
| | M221Dc | Southern Blue Ridge Mountains |
| | M221Dd | Metasedimentary Mountains |
| Cascades | M242Ba | West Cascade Slope Forest |
| | M242Bb | Western Cascades Highland Forest |
| | M242Bc | Cascade Crest Forest and Volcanic Peaks |
| | M242Be | Southern Oregon Cascade Highlands |
| | M242Bg | Southern Oregon Cascades |

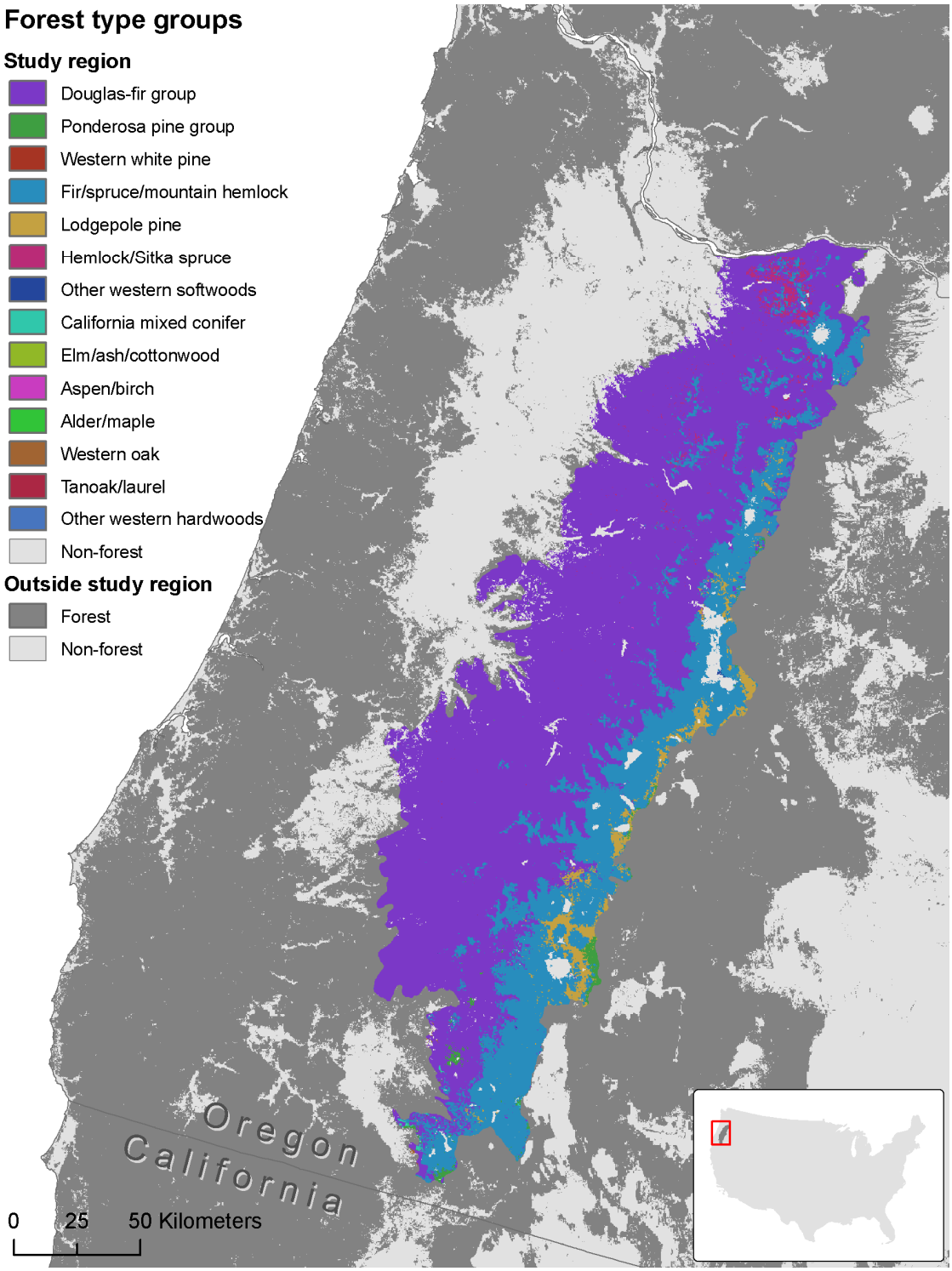


Figure A.1. U.S. Forest Type Groups of the Cascade study region (after Ruefenacht et al. in press)

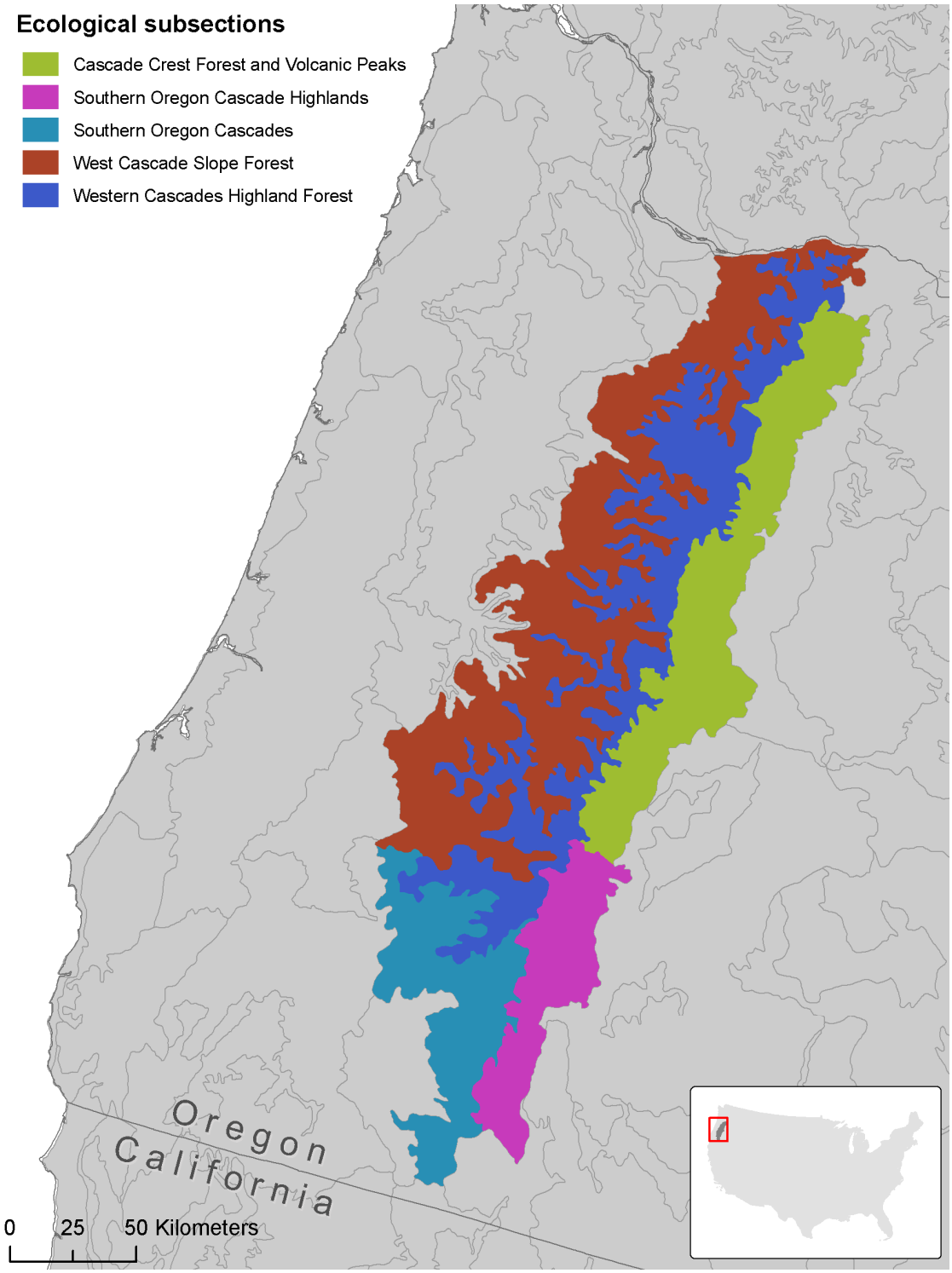


Figure A.2. Ecological subsections of the Cascade study region

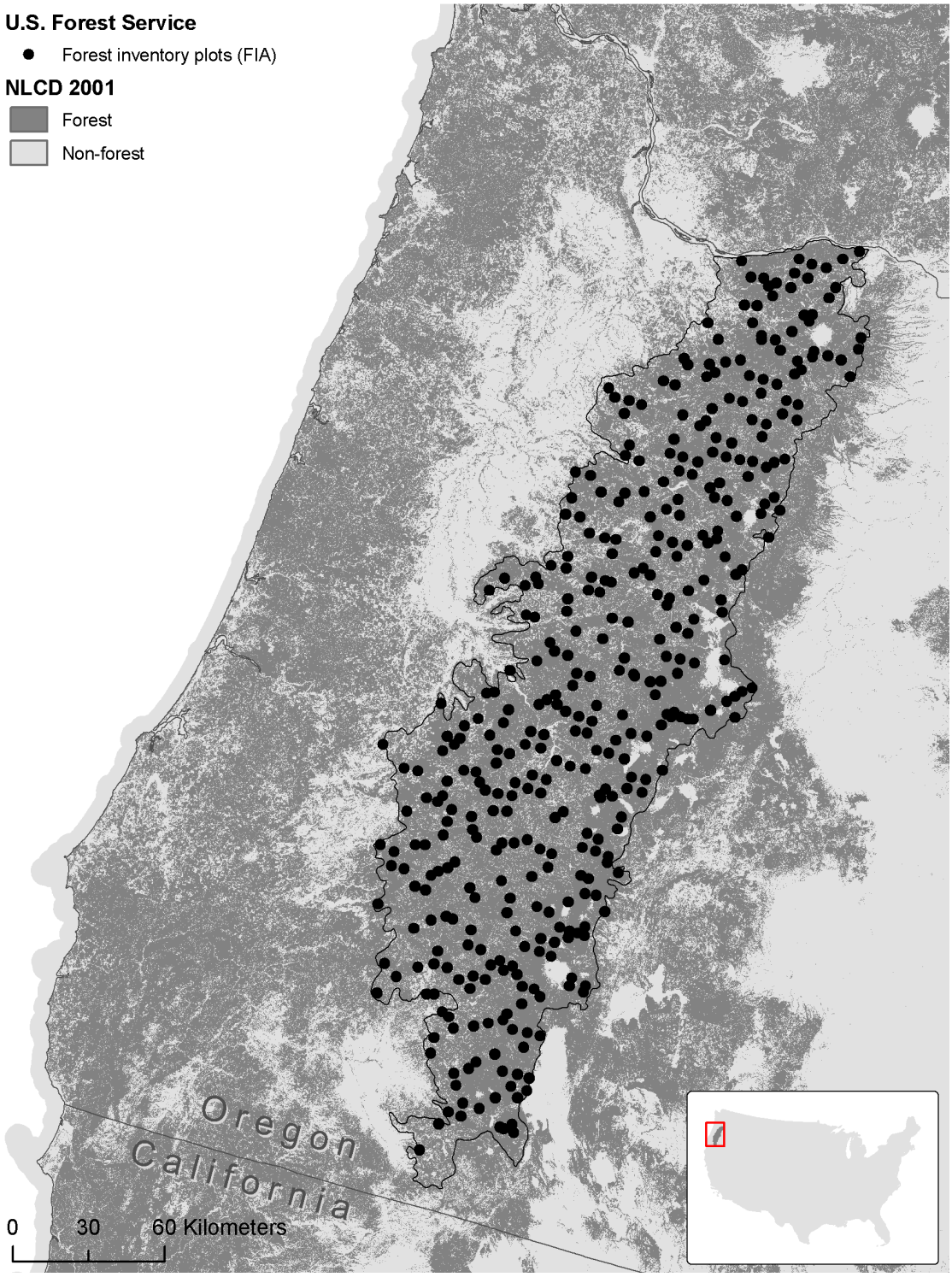


Figure A.3. Field plots from the U.S. Forest Service Annual FIA program used in this study for the Cascade region


Figure A.4 Cloud-free GLAS laser observations over forested areas in the Cascade region

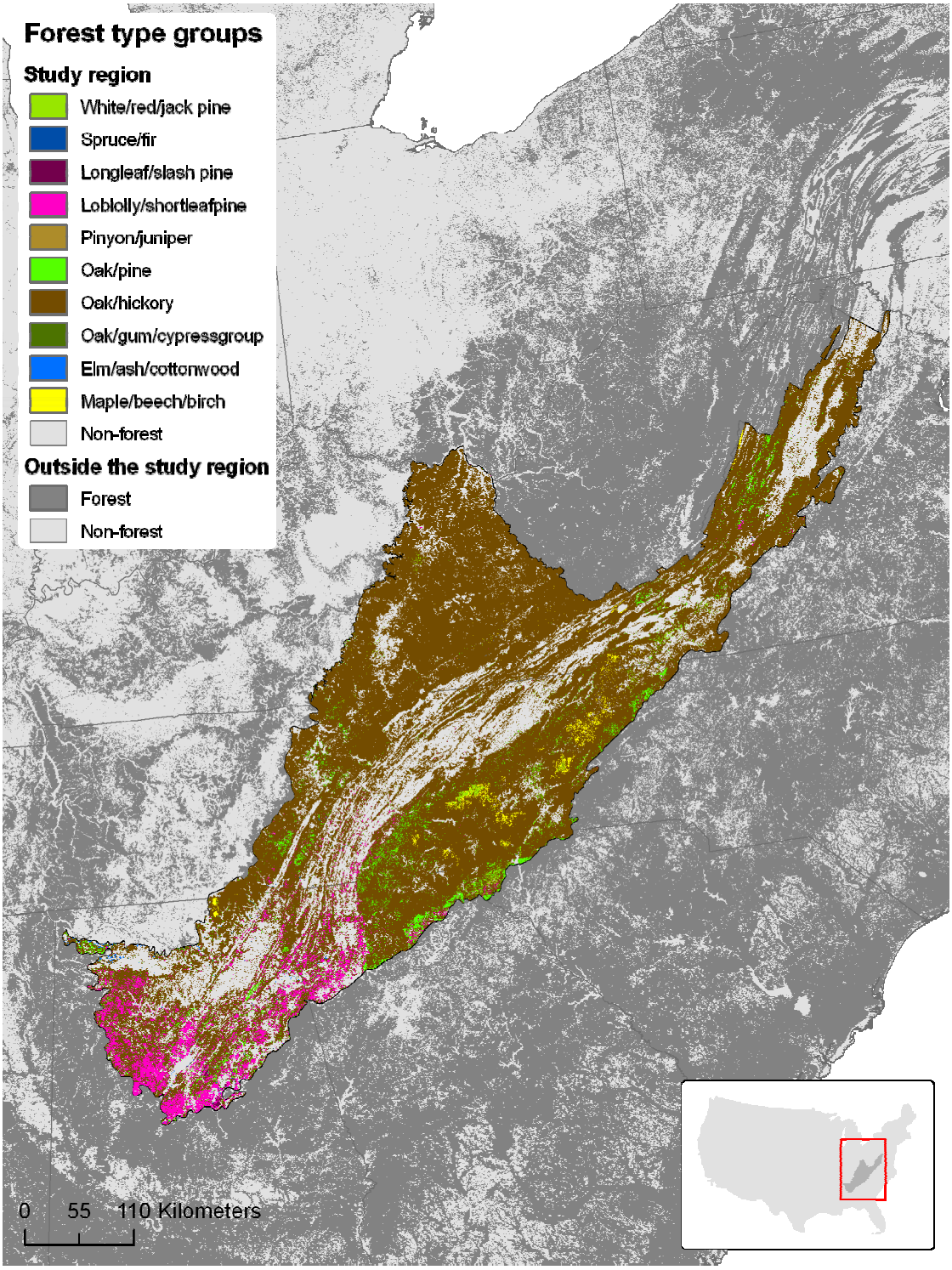


Figure A.5. U.S. Forest Type Groups of the Appalachian study region (after Ruefenacht et al. in press)

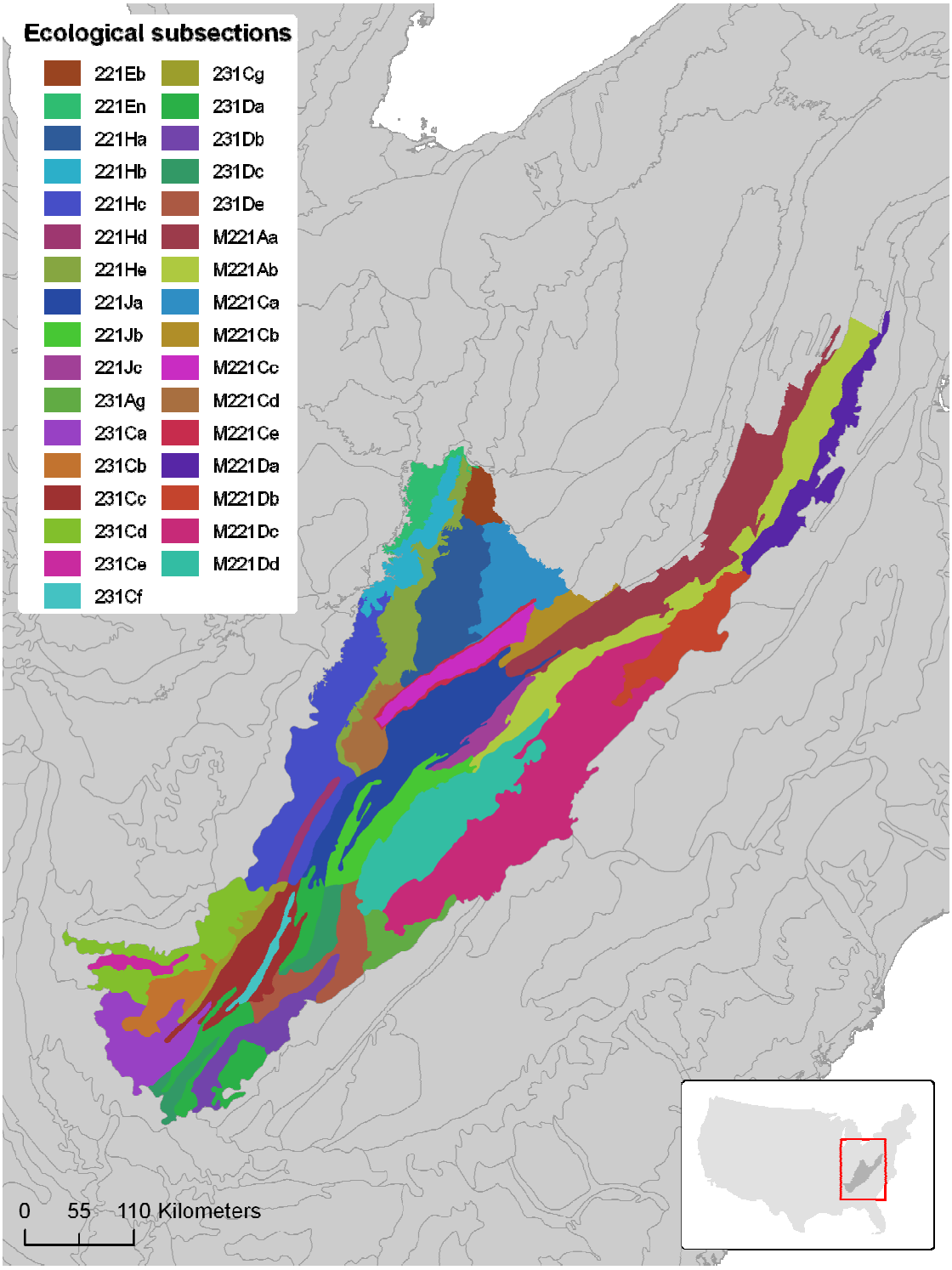


Figure A.6. Ecological subsections of the Appalachian study region. See table Table A.2 for subsection names.

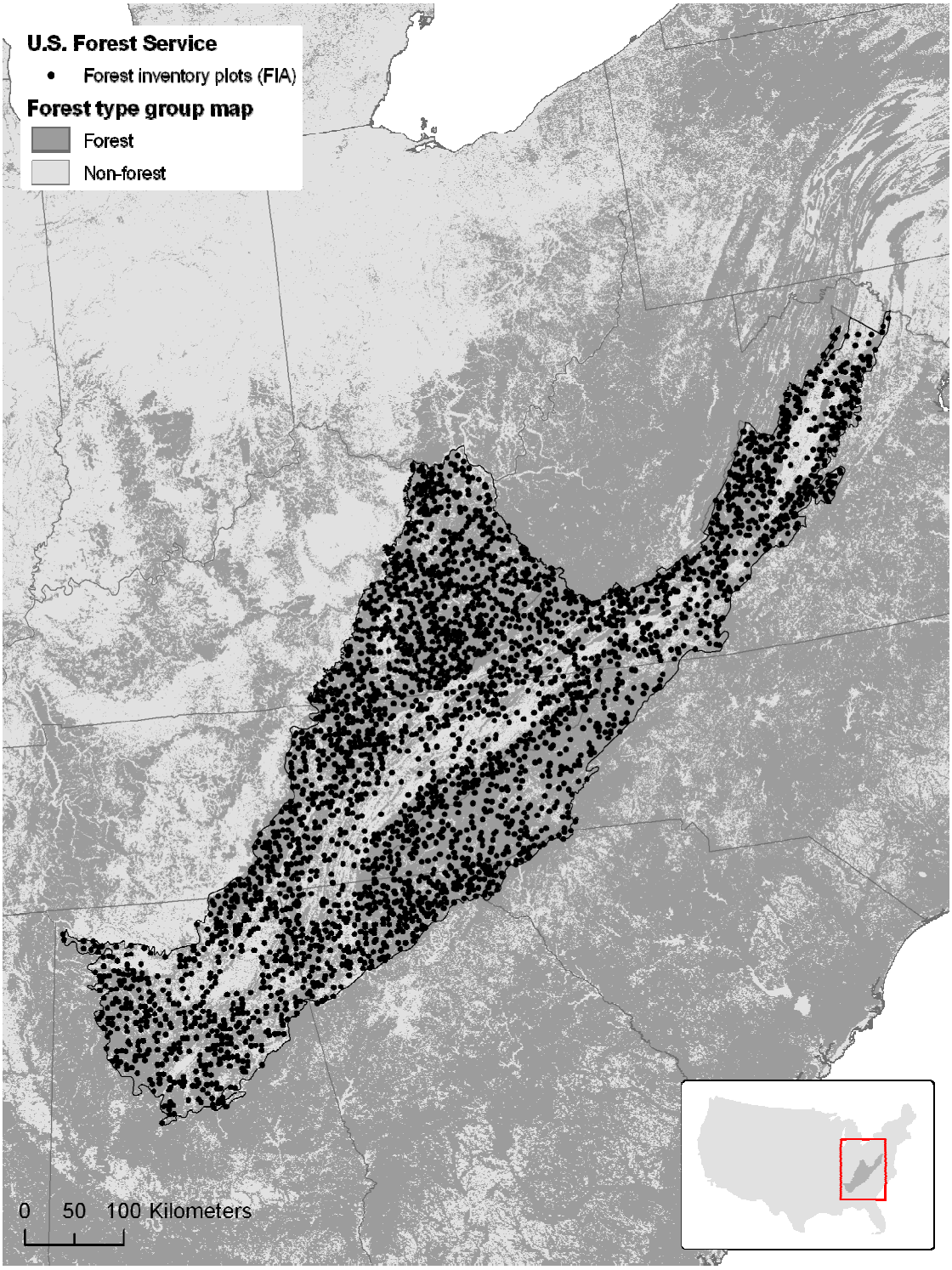


Figure A.7. Field plots from the U.S. Forest Service Annual FIA program used in this study for the Appalachian region

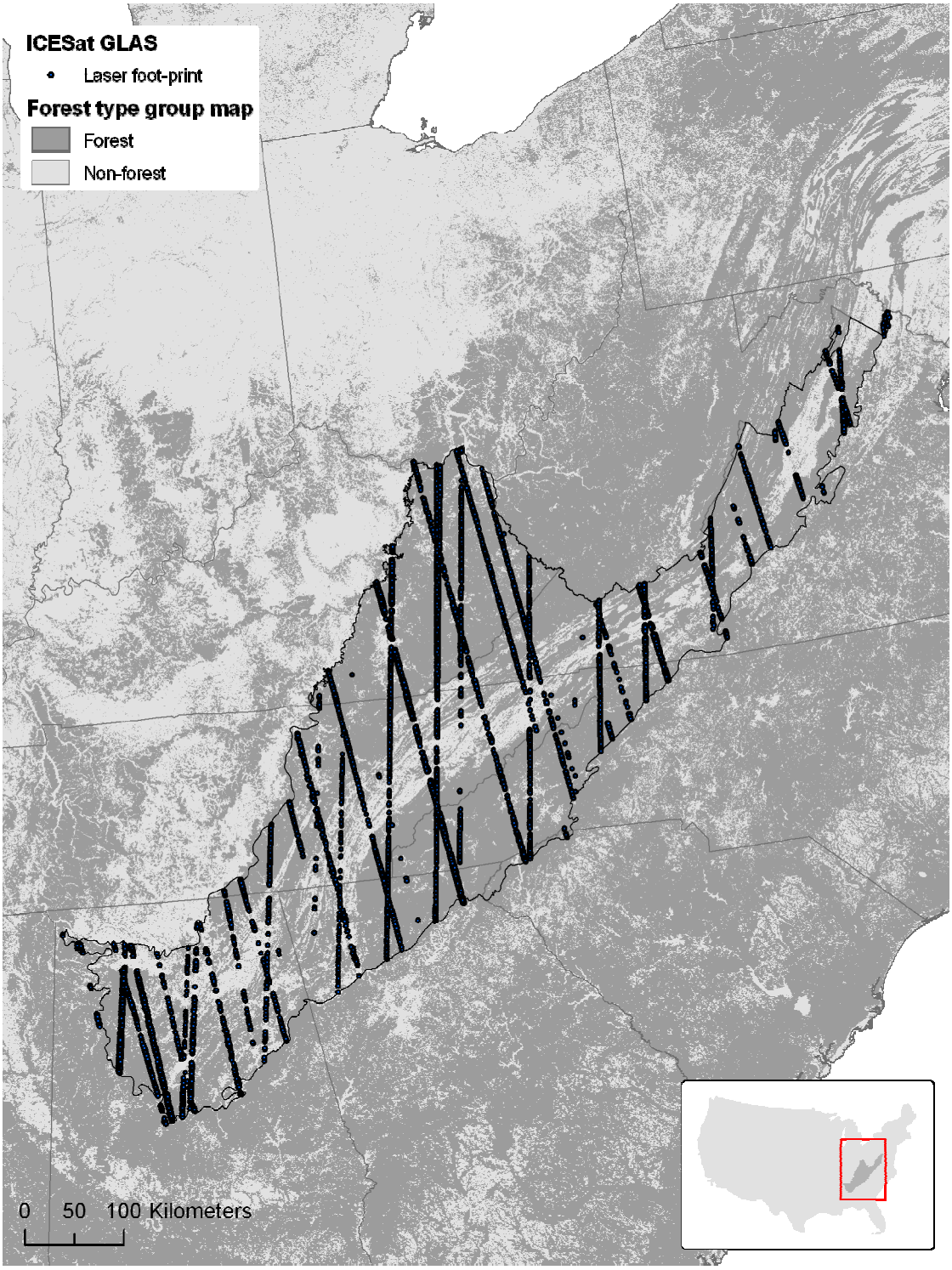


Figure A.8. Cloud-free GLAS observations over forested areas in the Appalachian region

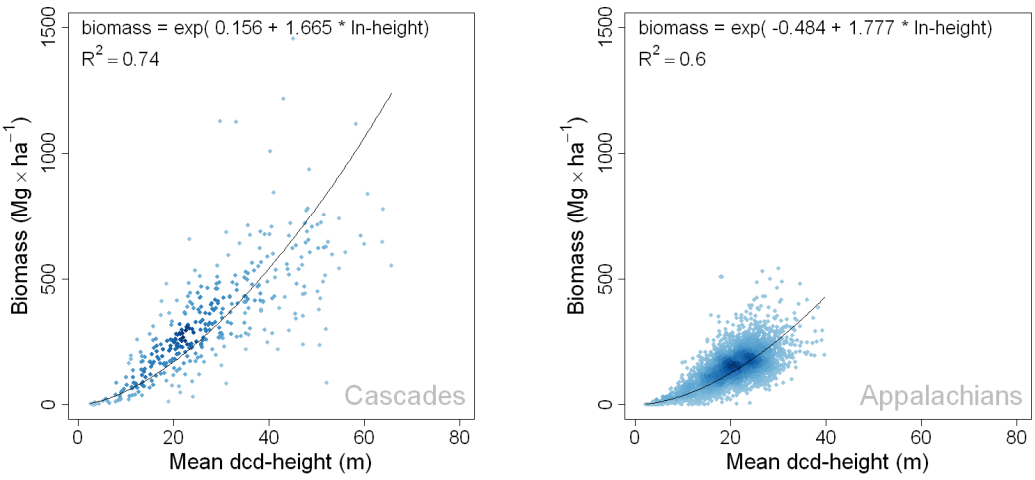


Figure A.9. Models of forest aboveground biomass for the Cascade (left) and Appalachian (right) study region based on mean-dcd height as a single predictor variable. Individual observations represent FIA annual field plots.