# Estimating Forest Canopy Height using Photon-counting Laser Altimetry

Mahmoud S. Awadallah[1], A. Lynn Abbott[1], Valerie A. Thomas[2], Randolph H. Wynne[2], and Ross F. Nelson[3]

[1] Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA, USA – {mawadala,abbott}@vt.edu

[2] Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, VA, USA – {thomas,wynne}@vt.edu

[3] Code 618, NASA Goddard Space Flight Center, Greenbelt, MD, USA – ross.f.nelson@nasa.gov

Paper Number: SL2013-034

## 1. Introduction

Reliable global estimates of forest biomass and its dynamics are critically important for understanding the global carbon cycle and its dynamics (e.g., Bellassen et al. 2011). Many studies have been published regarding the estimation of above-ground biomass using airborne or satellite analog lidar systems (e.g., Nelson et al. 1988; Naesset 2002; Drake et al. 2003; Popescu et al. 2003; Lim and Treitz 2004; Nelson et al. 2004; Bortolot and Wynne 2005; Sun et al. 2007; Thomas et al. 2006; van Aardt et al. 2008). In order to detect change at the global scale, repeated satellite observations are needed (Hall et al. 2011).

Laser altimetry is well suited to estimate vegetation height and structure (Herzfeld et al. 2013). However, at present, there are no operational lidar sensors in space that are designed to measure terrestrial surfaces. This situation will change with NASA's ICESat-2, which is scheduled for launch in October 2016. One goal of the ICESat-2 mission is to obtain elevation measurements that will enable independent determination of global vegetation height with a ground track spacing of less than 2 km over a two-year period. ICESat-2 will be equipped with the Advanced Topographic Laser Altimeter System (ATLAS), which is a multibeam system that will collect elevation data using a photon-counting technology. This approach yields clouds of discrete points, each resulting from the return of an individual green ($\lambda$ = 532 nm) photon.

ICESat-2 will be the successor to NASA's ICESat (Ice, Cloud and Land Elevation Satellite) mission, which acquired data near-globally during the period from 2003 to 2009. ICESat obtained canopy height estimates using the GLAS (Geoscience Laser Altimeter System) sensor, a laser altimeter for which elevation estimates are based on the analysis of the waveform returns (Schutz et al. 2005). Canopy height estimates were obtained with root-mean-square errors of 2 to 6 m. Unfortunately, much of the prior research into the estimation of forest biomass using lidar is not directly applicable to ICESat-2, because of the change in technology from the GLAS analog waveform system to the ATLAS photon counting system. Photon detectors introduce new challenges to the prediction of biomass, the largest being the often substantial amount of ambient noise in the atmosphere which appears in the photon cloud above and within the vegetation canopy and below the ground. Noise concerns can be mitigated for airborne platforms, but are expected to be particularly significant for the space-based ICESat-2. The anticipated noise levels will make the detection of the top of the canopy and the ground itself very challenging, particularly in complex forest ecosystems.

This paper describes research in which photon-counting lidar measurements from two sources have been assessed for their efficacy in estimation of forest canopy height. The two sensors are the Sigma Space Micro Pulse Lidar (MPL) system, which operates at 532 nm and is consistent with the planned ATLAS sensor on ICESat-2, and the Multiple-Altimeter Beam Experiment Lidar (MABEL), a multibeam sensor that was operated at high altitude on NASA's ER-2 platform. Both sources offer insights into the signal qualities that are anticipated for ICESat-2. Estimates of tree heights from these sources are compared with high-density airborne discrete lidar data collected for the same flight paths using NASA-Goddard's Lidar, Hyperspectral, and Thermal Imager (G-LiHT). The lidar sensor in G-LiHT uses small-footprint scanning analog technology, in contrast to the photon-counting approach used in the other sensors. Thus, the ability to detect forest canopy height and to calculate forest biomass along the G-LiHT flight lines is well understood. This paper will assess and compare the accuracy levels of all three sensors, with particular emphasis on ground detection under dense canopies.

The next section of this paper provides more details concerning the data sets that have been used in this study. Section 3 gives an overview of the novel canopy-ground separation method that we have developed, and Section 4 presents experimental results. Section 5 contains concluding remarks.

## 2. Technical Approach and Methodology

### 2.1 Overview

Data from photon-counting systems have a close affinity to small footprint discrete return data from airborne laser scanners, and these data have been used extensively for biophysical parameter estimation for over 15 years (e.g., Nilsson 1996; Næsset 1997; Means et al. 2000; Popescu et al. 2003; Popescu et al. 2004; Bortolot and Wynne 2005; van Aardt et al. 2006; Thomas et al. 2006). This affinity is a harbinger of improved precision of biomass estimation from spaceborne lidar, because a recent metaanalysis indicated that biomass estimated using discrete return systems has residual standard error (RSE) that is comparable or better than biomass estimated using waveform lidar, radar, optical data, or combinations of lidar and other remotely sensed data (Figure 1, Zolkos et al. 2013). By implication, then, reliable estimates of above-ground forest biomass should be possible using ICESat-2.

This paper describes an image-processing method has been developed to estimate the top-of-canopy and ground surfaces. To assess the accuracy of the proposed method and hence the potential utility of ICESat-2 data for the monitoring of forest height and biomass, the results obtained from simulated ICESat-2 data (from both Sigma Space MPL and MABEL) have been compared with high density airborne discrete lidar data collected for the same flight paths using NASA-Goddard's Lidar, Hyperspectral, and Thermal Imager (G-LiHT).

The lidar sensor in G-LiHT uses small footprint scanning analog lidar technology. Thus, the ability to detect forest canopy height and to calculate forest biomass along the G-LiHT flight lines is well understood, and maps of canopy height along the flight transects have already been generated by the G-LiHT science team (http://gliht.gsfc.nasa.gov/). We argue that if most of the noise can be identified and removed from the ICESat-2 data, then the photon-counting lidar data that will be available from ICESat-2 is conceptually similar to small footprint lidar. We have used the G-LiHT data to validate our height detection algorithm and to evaluate our ability to derive forest biomass from simulated ICESat-2 across a forest ecosystem gradient, as described below.

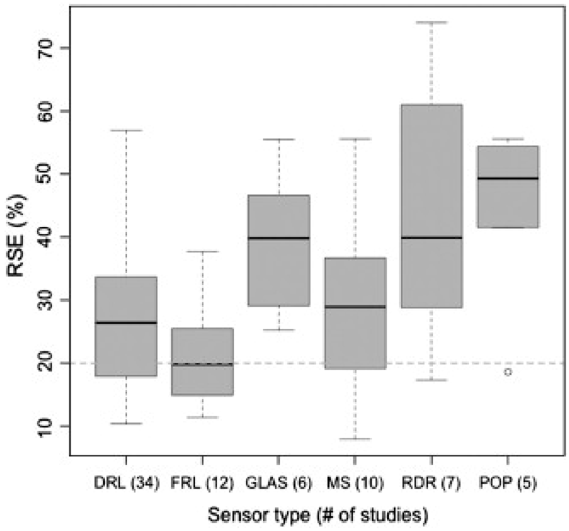



Figure 1: Residual standard error from estimates of above-ground biomass compiled from a metaanalysis by Zolkos et al. (2013). RSE(%) (RSE standardized by mean AGB from field measurements) categorized by sensor type, with dotted horizontal line at RSE = 20% of mean AGB. DRL = discrete return lidar, FRL = full return lidar, MS = multi-sensor, RDR = radar, POP = passive optical.

### 2.2 Study area

Two sources of simulated ICESat-2 data were used for this paper. These include low-altitude photon-counting acquisitions from the Sigma Space Micro Pulse Lidar (MPL) system, and high-altitude data from the Multiple Altimeter Beam Experimental Lidar (MABEL).

#### 2.2.1 Low-altitude Sigma Space Micro Pulse Lidar (MPL) data

The Sigma Space MPL sensor acquired low altitude (~600 m AGL) photon-counting data on October 8, 2009, over two mid-Atlantic, eastern US sites between sunset and 2200 hrs local time. As a result, there was very limited atmospheric noise in the original dataset. To simulate possible ICESat-2 data under varying sunlit atmospheric conditions, scientists at NASA's Goddard Space Flight Center generated multiple noise and beam strengths.

Two locations of MPL data were used in the algorithm development, the first just south of Annapolis, Maryland, USA on the Smithsonian Environmental Research Center (SERC). In this area of gently rolling hills, the dense (>95% canopy closure), tall (25–35 m) canopy consists mainly of hardwood species, including oak, hickory, maple, and tulip poplar. Given the ±50–100 m topography and dense overstory, the SERC data represents one of the more challenging types of forest ecosystems in the eastern US. The second location centers on two flux tower sites, the Silas Little and Cedar Bridge towers, in the Pine Barrens of southern New Jersey, just northeast of Atlantic City, NJ, and south of Fort Dix, NJ. The Pine Barrens is a large (~2500 km$^2$), flat, sandy area comprising mainly pine/oak, oak/pine, and pitch pine/scrub oak. The forest canopy closure is roughly 75-80%, with evident gaps. From the standpoint of vegetation measurements made using laser altimetry, it is an ideal natural target.

#### 2.2.2 High-altitude MABEL data

The second source of simulated ICESat-2 data is from the Multiple Altimeter Beam Experimental Lidar (MABEL). MABEL is a dual-wavelength (532 nm and 1064 nm) high-altitude system that was specifically developed as demonstrator and validation tool for ICESat-2. Of the 14 G-LiHT missions that intersect with MABEL flight lines, 2 were evaluated in this study. Both locations are near the Atlantic coast in North America, south of Washington, DC; one location is near Elkton, Maryland, and the other is near Jacksonville, North Carolina. These eastern, mid-Atlantic coastal-plain transects are predominantly southern pines (loblolly pine, Virginia pine, shortleaf pine), with hardwoods in the drainages. Moving away from the coast, there is a mixture of upland hardwood/pine forests.

## 3. Detection of Ground and Top of Canopy

Over forest canopies, the ICESat-2 laser data will include returns from the top of the canopy and from the ground, so that various canopy height measurements can be derived from the photon distribution between these two levels. However, as can be seen from our preliminary work, ICESat-2 data over forest canopy will pose three major challenges to canopy-ground detection:

1. Noise will be present throughout the atmospheric column, within the canopy, and below the ground, and will vary according to a) local atmospheric conditions, b) target reflectivity at 532 nm (or 1064 nm) and c) illumination conditions due to day/night changes and, during daylight, due to topography-sun angle interactions.
2. Dense canopies will naturally occlude the ground, making it very difficult to distinguish ground from noise in some instances. This effect will be magnified for closed-canopy forest ecosystems (Figure 2).
3. Gaps in the canopy can be difficult to localize in the presence of signal noise. These gaps could impact the accuracy of height metrics, and therefore affect biomass estimates, under certain conditions (Figure 2).

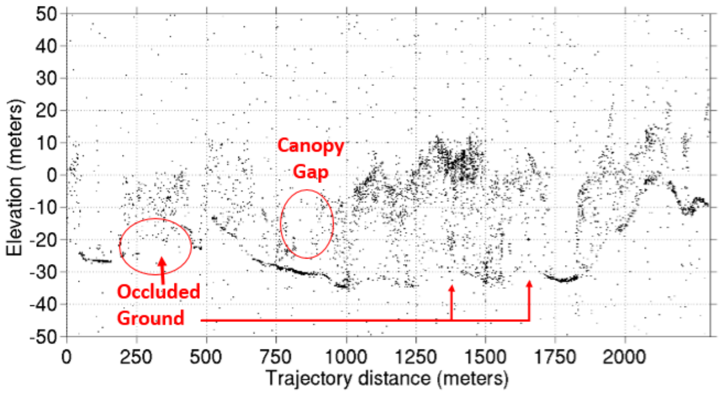



Figure 2: Unique challenges for forestry applications of ICESat-2 data. Noise can be seen throughout the entire vertical column. Canopy gaps and ground occlusion are evident.

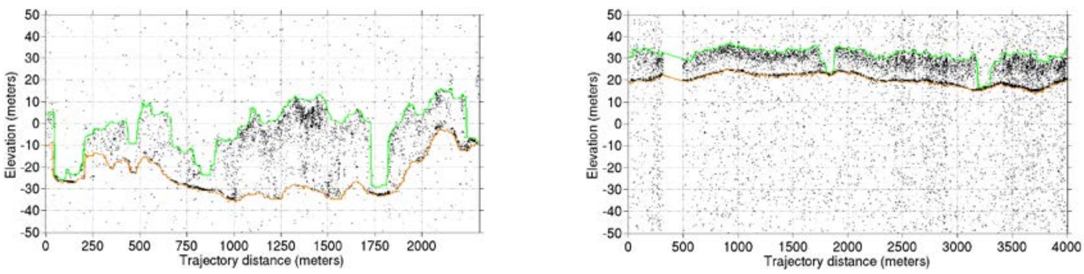
The main contribution of this research is the development of a novel, automated signal-analysis technique that can detect ground and top-of-canopy levels within ICESat-2 data. The main emphasis of the algorithm is on noise immunity and on localization of gaps in the canopy. The techniques addresses the canopy-ground detection problem using a combination of 1) noise filtering, 2) contour detection using deformable-model optimization, and 3) separation of ground and top of canopy. More discussion of the processing techniques are described in (Awadallah et al. 2013).

The work to date has been formulated as a problem of two-dimensional (2D) image analysis. Each data file consists of a set of ($x$, $y$, $z$) points from a single flight track. Approximately 10,000 to 25,000 values are present in each file, and the first step in our approach is to map these values onto a 2D grid. When displayed, the resulting image contains many points corresponding to the ground and canopy, along with noise points below the ground and above the canopy. For noise removal, a combination of median filtering, size filtering, and morphological processing have been applied. Much of the noise in the image consists of isolated points, often characterized as impulsive noise, and these techniques are well suited for removing this type of noise.

The next major processing step relies on *deformable models*, which refers to a class of optimization techniques that are widely used in image analysis today. The approach was first introduced by Kass et al. (1988). The fundamental idea is to perform an iterative search for the best fit of a 2D curve to a noisy image. The search is usually a “greedy” algorithm that updates the curve slightly at each interation step, in such a way that each update is locally optimum according to energy terms that characterize the curve’s shape and its immediate surroundings in the image (Awadallah et al. 2013). The approach does not guarantee that a globally optimum solution will be found, but represents a balance of accuracy and computation speed. These algorithms are sometimes known as “snakes” because of the curve’s appearance, over time, during the optimization procedure. We have extended the method of Chan and Vese (2001), who introduced a type of geometric active contour (GAC). In this approach, the curve is represented implicitly using a “level set” function, rather than the more traditional parametric form. A consequence is that multiple closed curves can be detected in the image. Our approach utilizes regional image statistics rather than more traditional intensity edges as a means of guiding the search. In our experiments, this approach has resulted in better detection of gaps in the canopy.

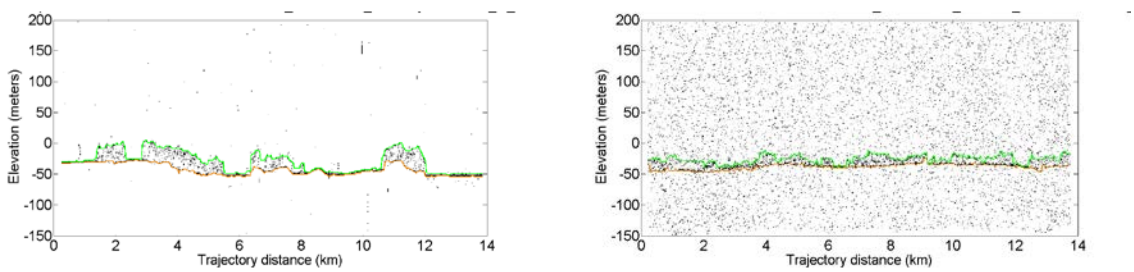
## 4. Results

Our algorithm was evaluated using an implementation in MATLAB. Figures 3 and 4 show example results obtained from the MPL and MABEL datasets respectively. The figures show that the proposed algorithm can estimate the top-of-canopy and ground surfaces despite the challenges of high noise level, canopy gaps, and occluded ground surface.




(a) (b)

Figure 3: Results using Sigma Space MPL lidar data. (a) Example from Smithsonian Environment Research Center, Maryland. (b) Example from Pine Barrens, New Jersey.




(a) (b)

Figure 4: Results using MABEL data. (a) Example from northeastern Maryland, near the city of Elkton. (b) Example from North Carolina, near Jacksonville.

To assess the accuracy of the proposed algorithm, we utilized several G-LiHT data sets that intersect MABEL flight lines. The G-LiHT files were coregistered manually to the MABEL data by researchers at Colorado State University, so that the G-LiHT data could serve as a ground-truth reference. Our algorithm detected ground and top-of-canopy curves in the MABEL data automatically, and then a comparison was made between the corresponding profiles. Figure 5 shows example results for a MABEL transect taken from northeastern Kentucky, where lots of hills and small watersheds are present. This example demonstrates that the proposed algorithm can estimate ground and top of canopy reasonably well despite dramatic changes in elevation. Figure 6 shows results for noisy data obtained for fairly flat terrain in southern North Carolina, about 65 km west of Wilmington, NC.

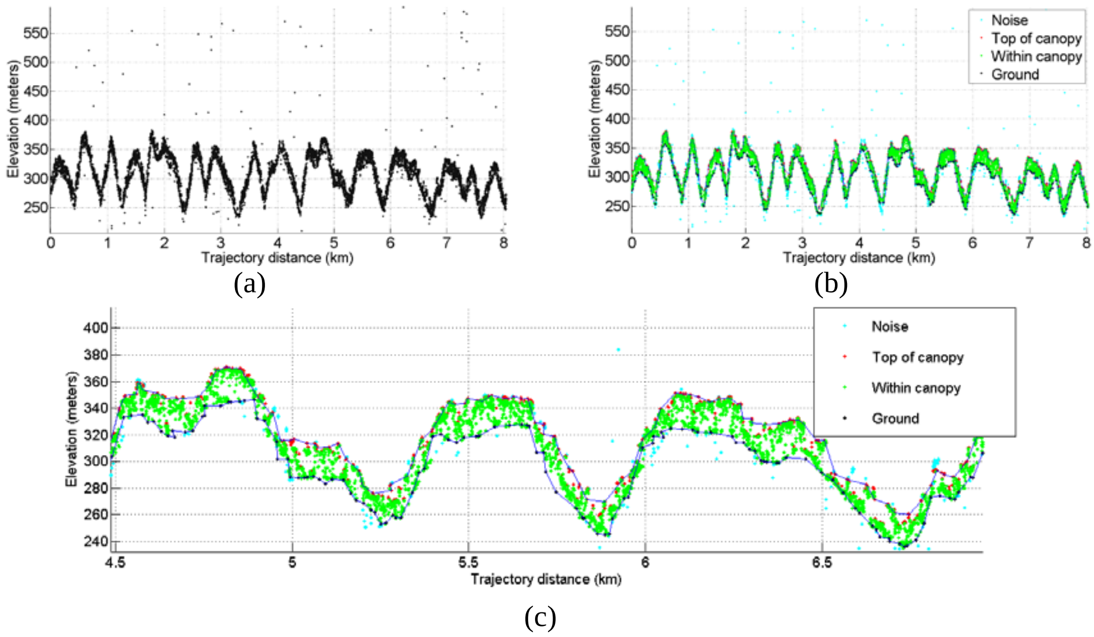



Figure 5: Profile view of a section of northeastern Kentucky. (a) Original MABEL data for an 8-km section. (b) The result of automatic estimation of ground and top of canopy. (c) Zoomed version of (b).

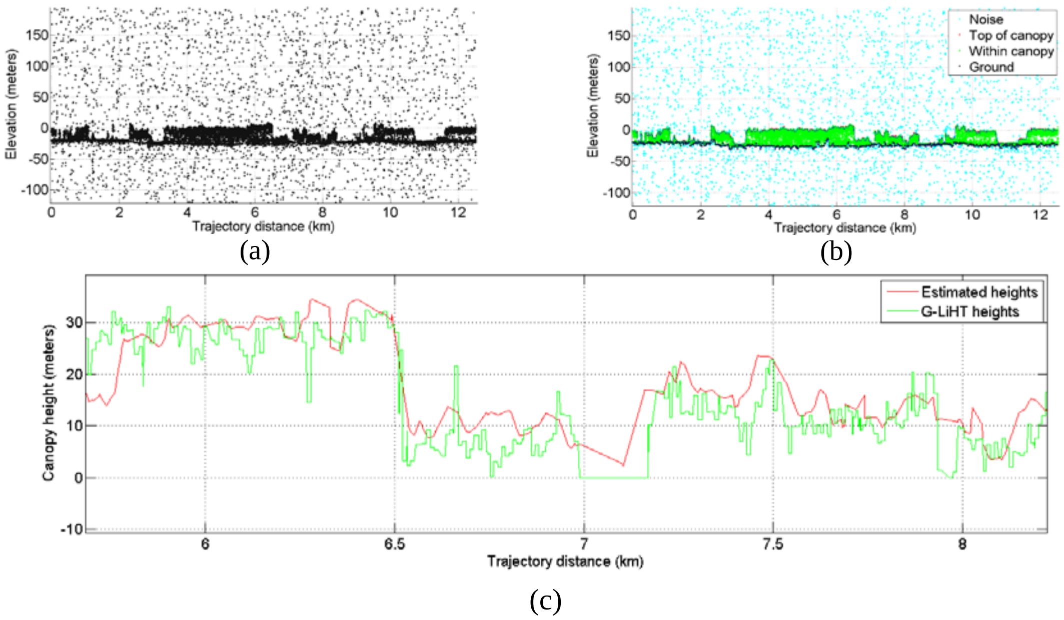



Figure 6: Profile view of a section of southern North Carolina, near Wilmington. (a) Original MABEL data for a 12-km section. (b) The result of automatic estimation of ground and top of canopy. (c) Comparison between the estimated canopy height for MABEL (red) and ground-truth values obtained from G-LiHT (green).

Figure 6(c) indicates a strong correlation between the two datasets, in spite of the high level of noise that is present in the MABEL data. Quantitative comparisons were performed for two sections of the data in Figure 6(c). For 5.6 to 6.5 km on the plot, a section of approximately 1 km, the median height and mean height obtained from G-LiHT were 27.99 and 27.51 m, respectively. The estimated values for the same section, for our automated algorithm with MABEL data, were 29.19 and 27.99 m, respectively. The computed error values for this section are therefore 1.2 m when comparing median canopy heights, and 0.48 m when comparing mean canopy heights.

We applied this same type of analysis to a section of approximately 100 m, from 7.7 to 7.8 km on the same plot. The median and mean canopy heights from G-LiHT were both 10.18 m. The automatically estimated heights from MABEL were 11.68 and 12.48 m, respectively. These values represent error values of 1.50 m when comparing median values, and 2.29 m when comparing mean canopy heights.

## 5. Conclusion

To our knowledge, no existing system is capable of automatically analyzing data that will be obtained from the ICESat-2 photon-counting sensor. The anticipated noise level is much higher than that for previous high-altitude lidar sensors. Because of the large amount of data that will be generated by the sensor, automated techniques are essential for extracting results in a way that is both fast and cost-effective. This paper has outlined a novel approach for addressing the canopy-ground detection problem, and the experimental results are very encouraging.

The estimation of vegetation biomass on a large scale is of critical importance in characterizing and understanding planetary-scale changes that are taking place in the Earth system. This paper describes an approach for automatic detection of vegetation height. It is expected that this work will facilitate follow-on studies that can bridge from vegetation height over multiple transects to large-scale estimates of biomass.

## Acknowledgements

The authors gratefully acknowledge the work of Dr. Mike Lefsky, of Colorado State University, for providing registered data sets in support of this work.

## References

M. Awadallah, S. Ghannam, A. L. Abbott, and A. Ghanem, 2013. Active Contour Models for Extracting Ground and Forest Canopy Curves from Discrete Laser Altimeter Data. *Proceedings: 13th International Conference on LiDAR Applications for Assessing Forest Ecosystems* (SilviLaser 2013), Beijing, China, Oct. 2013.

V. Bellassen, N. Delbart, G. Le Maire, S. Luyssaert, P. Ciais, and N. Viovy, 2011. Potential knowledge gain in large-scale simulations of forest carbon fluxes from remotely sensed biomass and height. *Forest Ecology and Management*, 261 (3), 515-530.

Z. J. Bortolot and R. H. Wynne, 2005. Estimating forest biomass using small footprint LiDAR data: An individual tree-based approach that incorporates training data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59, 342-360.

T. Chan and L. Vese, 2001. Active contours without edges. *IEEE Transactions on Image Processing,* 10 (2), 266–277.

J. B. Drake, 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, 147–159.

F. G. Hall, K. Bergen, J. B Blair, R. Dubayah, R. Houghton, G. Hurtt, J. Kellndorfer, M. Lefsky, K. J. Ranson, S. Saatchi, H. H. Shugart, and D. Wickland, 2011. Characterizing 3D vegetation structure from space: Mission requirements. *Remote Sensing of Environment*, 115 (11), 2753-2775.

U. C. Herzfeld, B. W. McDonald, B. F. Wallin, T. A. Neumann, T. Markus, A. Brenner, and C. Field, 2013. Algorithm for Detection of Ground and Canopy Cover in Micropulse Photon-Counting Lidar Altimeter Data in Preparation for the ICESat-2 Mission. To appear in *IEEE Transactions on Geoscience and Remote Sensing*.

M. Kass, A. Witkin, and D. Terzopoulos, 1988. Snakes: Active Contour Models. *International Journal of Computer Vision*, 1 (4), 321-331.

K. S. Lim and P. M. Treitz, 2004. Estimation of above ground forest biomass from airborne discrete return laser scanner data using canopy-based quantile estimators. *Scandinavian Journal of Forest Research*, 19, 558−570.

J. E. Means, S. A. Acker, B. J. Fitt, M. Renslow, L. Emerson, and C. J. Hendrix, 2000. Predicting Forest Stand Characteristics with Airborne Scanning Lidar. *Photogrammetric Engineering and Remote Sensing,* 66(11), 1367-1371.

E. Næsset, 1997. Estimating Timber Volume of Forest Stands Using Airborne Laser Scanner Data. *Remote Sensing of Environment*, 61, 246-253.

E. Naesset, 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, 80 (1), 88–99.

R. Nelson, W. Krabill, and J. Tonelli, 1988. Estimating forest biomass and volume using airborne laser data. *Remote Sensing of Environment*, 24 (2), 247–267.

R. Nelson, A. Short, and M. Valenti, 2004. Measuring biomass and carbon in Delaware using an airborne profiling LiDAR. *Scandinavian Journal of Forest Research*, 19 (6), 500–511.

M. Nilsson, 1996. Estimation of Tree Heights and Stand Volume Using an Airborne Lidar System. *Remote Sensing of Environment*, 56, 1-7.

S. C. Popescu, R. H. Wynne, and R. F. Nelson, 2003. Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. *Canadian Journal of Remote Sensing*, 29 (5), 564-577.

S. C. Popescu, R. H. Wynne, and J. A. Scrivani, 2004. Fusion of smallfootprint LiDAR and multispectral data to estimate plot-level volume and biomass in deciduous and pine forests in Virginia, USA. *Forest Science*, 50 (4), 551– 565.

B. E. Schutz, H. J. Zwally, C. A. Shuman, D. Hancock, and J. P. DiMarzio, 2005. Overview of the ICESat mission. *Geophys. Res. Lett.*, 32 (21), L21S01.

G. Sun, K. J. Ranson, J. Masek, A. Fu, and D. Wang, 2007. Predicting tree height and biomass from GLAS data. *Proceedings of the 10th International Symposium on Physical Measurements and Signatures in Remote Sensing*, Davos, Switzerland.

V. Thomas, P. Treitz, J. H. McCaughey, and I. Morrison, 2006. Mapping stand-level forest biophysical variables for a mixedwood boreal forest using lidar: an examination of scanning density. *Canadian Journal of Forest Research*, 36 (1), 34–47.

J. A. N. van Aardt, R. H. Wynne, and R. G. Oderwald, 2006. Forest volume and biomass estimation using small footprint lidar-distributional parameters on a per-segment basis. *Forest Science*, 52, 636-649.

J. A. N. van Aardt, R. H. Wynne, and J. A. Scrivani, 2008. LiDAR-based mapping of forest volume and biomass by taxonomic group using structurally homogenous segments. *Photogrammetric Engineering & Remote Sensing*, 74 (8), 1033–1044.

S. G. Zolkos, S. J. Goetz, and R. Dubayah, 2013. A meta-analysis of terrestrial above ground biomass estimation using lidar remote sensing. *Remote Sensing of Environment*, 128, 289-298.