Retrieval of Savanna Vegetation Canopy Height from ICESat-GLAS Spaceborne LiDAR With Terrain Correction

Ehsan Khalefa, Izak P. J. Smit, Alecia Nickless, Sally Archibald, Alexis Comber, and Heiko Balzter

Abstract—Light detection and ranging (LiDAR) remote sensing enables accurate estimation and monitoring of vegetation structural properties. Airborne and spaceborne LiDAR is known to provide reliable information on terrain elevation and forest canopy height over closed forests. However, it has rarely been used to characterize savannas, which have a complex structure of trees coexisting with grasses. This letter presents the first validation of spaceborne Ice Cloud and land Elevation Satellite Geoscience Laser Altimeter System (GLAS) full-waveform data to retrieve savanna vegetation canopy height that uses field data specifically collected within the GLAS footprints. Two methods were explored in the Kruger National Park, South Africa: one based on the Level 2 Global Land Surface Altimetry Data product and the other using Level 1A Global Altimetry Data (GLA01) with terrain correction. Both methods use Gaussian decomposition of the full waveform. Airborne LiDAR (AL) was also used to quantify terrain variability (slope) and canopy height within the GLAS footprints. The canopy height retrievals were validated with field observations in 23 GLAS footprints and show that the direct method works well over flat areas (Pearson correlation coefficient $r = 0.70$, $p < 0.01$, and $n = 8$ for GLA01) and moderate slopes ($r = 0.68$, $p < 0.05$, and $n = 9$ for GLA01). Over steep slopes in the footprint, however, the retrievals showed no significant correlation and required a statistical correction method to remove the effect of terrain variability on the waveform extent. This method improved the estimation accuracy of maximum vegetation height with correlations ($R^2 = 0.93$, $p < 0.05$, and $n = 6$ using the terrain index (g) generated from AL data and $R^2 = 0.91$, $p < 0.05$, and $n = 6$ using the GLAS returned waveform width parameter). The results suggest that GLAS can provide savanna canopy height estimations in complex tree/grass plant communities.

Index Terms—Canopy height, Geoscience Laser Altimeter System (GLAS), Ice Cloud and land Elevation Satellite (ICESat), Kruger National Park, light detection and ranging (LiDAR), savanna, terrain correction.

I. INTRODUCTION

LIGHT detection and ranging (LiDAR) remote sensing provides an accurate and efficient means of estimating and monitoring vegetation structural properties. It has the potential to directly measure the 3-D structure of vegetation canopies, including tree canopy height, density, and indirectly aboveground biomass, as well as terrain surface [1]–[3]. While LiDAR from the airborne and spaceborne platforms has been proved to be effective in forestry applications [4]–[8], few studies exist for savannas, so its usefulness in that environment has to be further investigated. However, some studies have used airborne LiDAR (AL) in savannas [9]–[11]. The full-waveform spaceborne satellite LiDAR data, from the Geoscience Laser Altimeter System (GLAS) aboard the Ice Cloud and land Elevation Satellite (ICESat), have not been thoroughly evaluated in savanna systems using field data specifically collected within the GLAS footprints yet. A study by Baccini *et al.* [12] showed potential results of mapping aboveground biomass over tropical Africa using multiyear Moderate Resolution Imaging Spectroradiometer satellite observation and a wide range of field measurements and compared the results with GLAS LiDAR height metrics. GLAS has been shown to provide accurate estimates of forest canopy height and aboveground biomass in several studies of closed forests [13]–[18].

This letter evaluates the capability of satellite LiDAR data from GLAS for retrieving savanna canopy height in a landscape with highly complex vegetation structure due to the tree/grass coexistence. AL data are utilized to investigate the terrain variability within the GLAS footprints and its effect on the canopy height retrieval accuracy. An accuracy assessment of savanna canopy height using plot measurements within the footprint locations collected during a field campaign is carried out. Terrain elevation is also evaluated.

II. METHODOLOGY

A. Study Site and Data

The study area is inside the Kruger National Park in South Africa. The park covers $19\,485$ km^2 and is one of the largest protected areas in Africa. Various habitats and ecological regions exist within the boundary of the Kruger Park, with at least 16 recognized ecozones, each characterized by specific vegetation, geology, soil, rainfall rate, and temperature [19].

All GLAS footprints were located in the Sabie/Crocodile Thorn Thickets ecozone, which is characterized by a very dense growth of thorny shrubs.
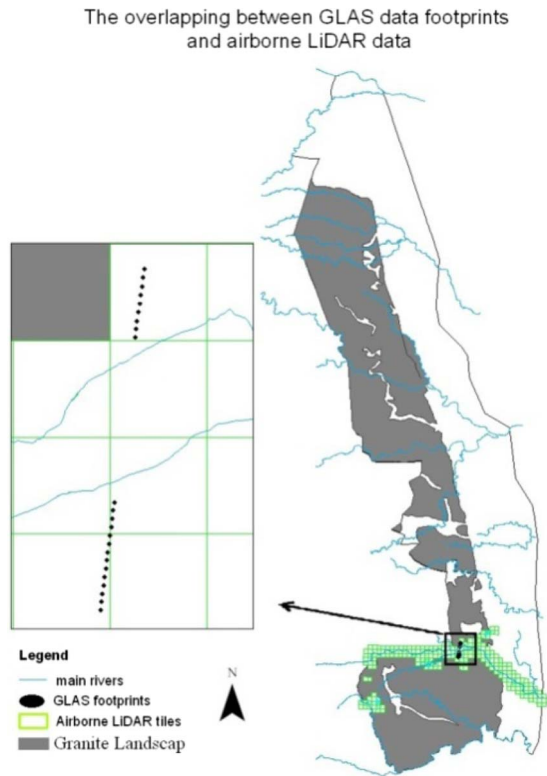
Fig. 1.   Location of the GLAS footprints which overlap with AL data tiles in the study area. The footprints are situated in (gray) the granitic bedrock region of Kruger National Park, in the Sabie/Crocodile Thorn Thickets ecozone.

GLAS Data: Two data products of GLAS release 29 were used in this study: Level 1A Global Altimetry Data (GLA01) and Level 2 Global Land Surface Altimetry Data (GLA14). ICESat-GLAS data were acquired in February–March 2009. GLA01 contains the transmitted and received echo waveforms. This product contains variables such as the filter threshold value for signal detection in digitizer counts, laser transmitter energy, received energy from all signals above threshold, sampled transmit pulse waveform, 4-ns background mean value, and standard deviation [20]. GLA14 contains the sensor position and pointing information as well as the calculated footprint position, size and shape, and land surface elevation. Transmit pulse and recorded waveforms are represented with characteristic shape parameters only. The recorded waveform is decomposed into a series of Gaussian peaks, as described in [20] and [21].

AL Data: The discrete return AL data set was acquired during the period of August 24–September 9, 2004, by the University of the Witwatersrand. Subsets of AL data were created using a radius of 35 m for each geolocated ICESat/GLAS footprint position. The analysis of AL indicated that the topography over the GLAS footprints for the selected footprints can be classified into flat terrain (0°–15° slope), moderate slope (15°–30°), and steep slope (> 30°).

Field Data: Fieldwork was conducted in August 2010. The entire selected GLAS footprints ($n = 23$) for field data collection are located in the granitic areas (Fig. 1). The criteria for selecting the field plot locations were based on the overlap between spaceborne LiDAR data and AL data, as well as

considerations of accessibility and travel times. In the field campaign, vegetation canopy height was measured using four sampling plots within the GLAS footprint. The GLAS footprints are elliptical with roughly 70-m diameter and varying eccentricity, separated by about 170 m along track. For analytical purposes, we approximated the footprints with circles of 35-m radius. One circular subplot in the center of the GLAS footprint was measured, followed by three further subplots whose center position was 20 m to the north, southeast, and southwest from the footprint center at angles of 0°, 120°, and 240°. The subplots were 10 m in radius. Tree heights of all woody vegetation, including trees and shrubs, taller than 1 m aboveground were systematically measured using a clinometer or a woodland stick.

B. Data Processing

GLAS Data Processing: Two methods that retrieve savanna canopy height from different waveform parameters were developed. The direct method utilizes the waveform parameters derived during the operational GLAS processing, mainly the GLA14 product. The statistical method analyzes the waveform shapes from the GLA01 product, using a digital elevation model (DEM) to adjust the waveform extent, which proved important in rugged terrain. All calculations were performed using the statistical analysis program package R 2.9.0.

Direct Method: In open and sparsely vegetated areas, the majority of the signal is likely to be returned from the terrain surface, which manifests itself in a ground peak with the largest amplitude in the waveform. In denser canopies, the laser energy penetration is reduced, resulting in a ground peak of lower amplitude [22]. The GLAS14 product provides waveform parameters estimated by a multiple Gaussian decomposition. The ICESat-GLAS Visualizer software iteratively fits six Gaussian curves to the waveform. This was used to compare Gaussian parameters. It is assumed that the Gaussian component with the largest amplitude corresponds to the ground return and the difference in elevation between the centroid of the ground peak and the beginning of the waveform signal is the maximum vegetation height [15]

$$H_{\max} = H_s - \mathrm{GP} \qquad (1)$$

where $H_{\max}$ is the maximum canopy height (in meters) estimated from LiDAR, H_s is the signal start (in meters), and GP is the ground peak (in meters).

Height estimates were also calculated from GLA01. As was explained earlier, raw waveforms have ambient system noise at the beginning and the end of the waveform signal. Therefore, the real signal of the waveform was identified by thresholding the raw waveforms. In this study, the threshold value for each raw waveform was individually determined to the mean background noise value estimated in the GLA01 product (GLAS product variable: d_4nsBgMean) plus four times the standard deviation of the background noise (GLAS product variable: d_4nsBgSDEV) [14]. To detect the ground peak location, the maximum amplitude was used as a reference. Hence, the peak with maximum amplitude was designated as the ground return.

Statistical Method: To consider the contribution of topographic relief on waveform shape, a terrain index (g) was calculated for each footprint location using a 7×7 subset DEM generated from AL data for each of the 23 GLAS footprints [17]

$$H_{\max} = b_0^* w - b_1^* g \qquad (2)$$

where w is the full-waveform extent and b_0 and b_1 are regression coefficients.

If an external DEM is not available, the terrain index g has to be substituted by another parameter. The width of the ground peak (i_Gsigma in GLA14), expressed as 2σ, is a potentially suitable indicator of the terrain effect on the waveform in sloped areas. On the assumption that the terrain index g can be replaced by the waveform width ($2^*\sigma$), vegetation canopy height can be estimated by

$$H_{\max} = b_0^* w - b_1^* (2^*\sigma). \qquad (3)$$

This method was adopted by Harding *et al.* [22] and offers a means of extracting GLAS canopy height estimates from the Gaussian decomposition parameters provided with the GLA14 product or extracted from GLA01 parameters. Crucially, it does not require an external DEM for terrain correction.

AL Processing: The AL footprints were classified into vegetation and ground hits, from which a digital surface model (DSM) and a digital terrain model (DTM) were generated using eCognition Developer 8 and Environmental Sciences Research Institute ArcMAP 9.3.1, respectively. These data were also used to calculate the terrain slope within each GLAS footprint ($n = 23$). A canopy height model (CHM) was calculated as the difference between DSM and DTM following the method by Hinsley *et al.* [23]. CHM values < 1 m were excluded from the data set in order to eliminate the effect of grasses and understory vegetation. Within each GLAS footprint, $H_{\max}$ was estimated by using the local maxima and minima of the generated CHM. The results from the airborne processing were used to validate $H_{\max}$ from the field plots and $H_{\max}$ obtained from the spaceborne LiDAR system.

III. RESULTS AND DISCUSSION

A. Direct Method Using Parameters of GLAS Products

A comparison of maximum canopy heights derived from GLA14 and GLA01 products gives a Pearson correlation $r = 0.78$ for the selected study plots. To validate the previous results, a comparison of the results of $H_{\max}$ from the field plots and $H_{\max}$ obtained from the spaceborne LiDAR system parameters (GLA01 and GLA14) is shown in Fig. 2.

The analysis of the AL data indicated that the topography over the selected GLAS footprints can be classified into terrain with low ($0°$–$15°$), moderate ($15°$–$30°$), and steep slopes ($> 30°$).

As a result of this stratification, eight GLAS footprints were in flat terrain, nine footprints in moderate terrain, and six footprints in steep terrain. In order to quantify the effect of terrain slope on both LiDAR systems, a comparison between results of $H_{\max}$ from field measurements and $H_{\max}$ obtained from both LiDAR systems stratified by terrain slope in the GLAS
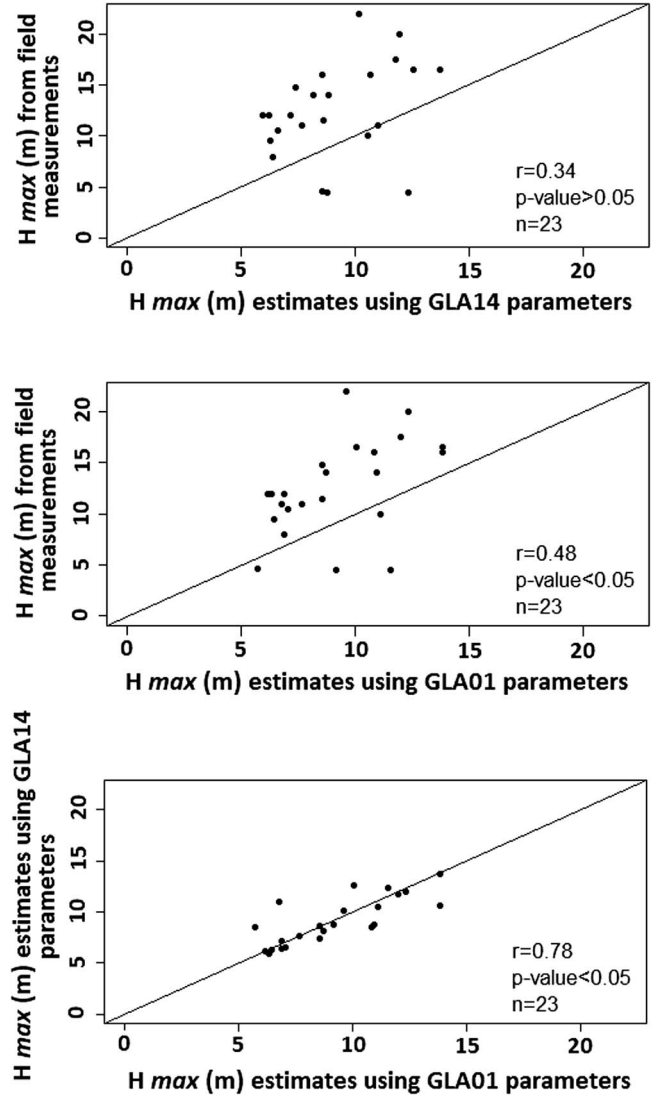

Fig. 2. Scatter plots of Pearson's correlation coefficient between $H_{\max}$ estimates using GLAS parameters of Gaussian decomposition and field measurements of $H_{\max}$. Upper plot shows GLA14, and middle plot shows GLA01. The lower plot shows the correlation coefficient between $H_{\max}$ using GLA14 and GLA01 parameters. In general, the results indicate a weak correlation when the comparison is made for all 23 footprints without considering terrain variability.

footprints was carried out. The results show that both LiDAR systems give a similar relationship to the field measurements of $H_{\max}$ (Fig. 3).

In flat terrain, the correlations between $H_{\max}$ from GLAS and AL with the field measurements were highest ($r = 0.70$ for GLAS and $r = 0.78$ for AL; both with $p < 0.001$).

Correlation coefficients were statistically significant in moderate terrain for both GLAS and AL when compared to field measurements of $H_{\max}$ ($r = 0.68$ and $p < 0.05$ for GLAS and $r = 0.70$ and $p < 0.01$ for AL). In steep terrain, the correlations with field-measured $H_{\max}$ were not significant anymore ($r = 0.17$ and $p > 0.05$ for GLAS and $r = 0.33$ and $p > 0.05$ for AL). Because of the small sample size, a bootstrap correlation analysis was carried out which fully confirmed these conclusions. The results show that the estimation of $H_{\max}$ from both air- and spaceborne LiDAR systems is affected by terrain slope.
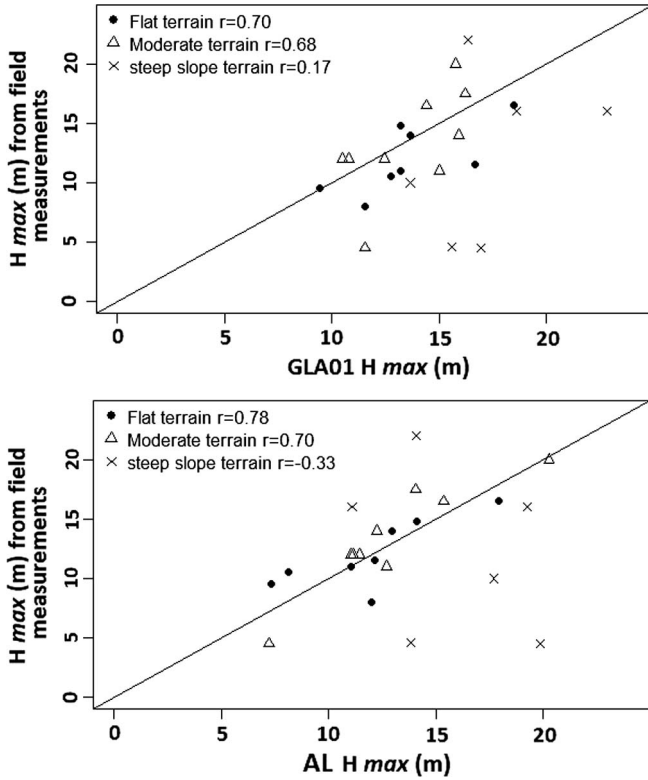
Fig. 3. Scatter plots between $H_{\max}$ within the GLAS footprints stratified by terrain slope. (Top) $H_{\max}$ from field measurements and from GLA01. (Bottom) $H_{\max}$ from field measurements and from AL. This figure shows that r decreases when the terrain slope increases.

Therefore, statistical methods are required to incorporate this important factor in estimating vegetation height in steep terrain.

B. Statistical Methods of Terrain Correction

Two methods of terrain correction were applied. The first method uses the waveform extent (w) extracted from GLA01 parameters and a terrain index (g), which is generated from AL data within the GLAS footprint. g is defined as the difference between the highest elevation and the lowest elevation within a 7×7 subset of the DEM from AL. It is an indicator of the slope effect on the waveform extent.

In order to improve the relationship between the waveform extent and terrain indices, a multiple regression with field measurements of $H_{\max}$ was used $(n = 6)$, following [17]. The resulting equation which uses the AL terrain index is

$$H_{\max} = 1.47w - 1.19g \qquad (4)$$

with $R^2 = 0.93$ and rmse $= 3.64$ m (both with $p < 0.05$).

The analysis showed the capability of estimating $H_{\max}$ from the GLAS waveform with consideration of terrain effects on the waveform structure. However, the largest variation was obtained for very steep slopes $(n = 6)$. The validation of $H_{\max}$ estimated from GLAS waveform parameters with field-derived $H_{\max}$ suggests that the terrain index is related to waveform structure. However, this approach to terrain correction requires terrain elevation data of appropriate resolution to aid the GLAS analysis.
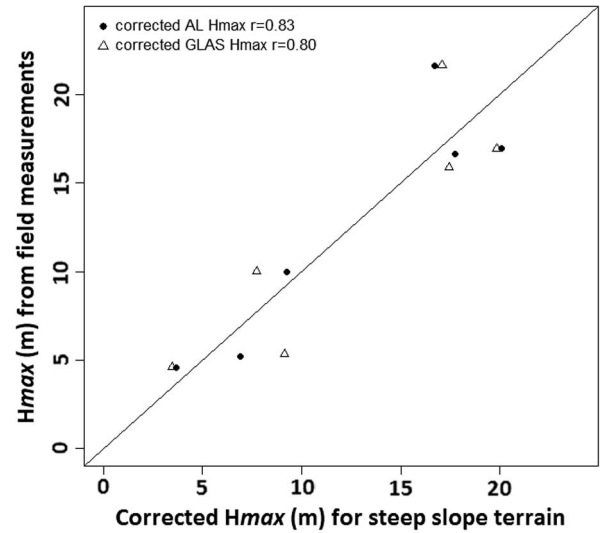

Fig. 4. Scatter plots between corrected $H_{\max}$ within the steep slope terrain GLAS footprints using statistical method of terrain correction and the $H_{\max}$ from field measurements. This figure shows that R increases when the terrain slope is corrected using the terrain index (g) extracted from AL data $r = 83$, and it also shows that the terrain correction could be done using the waveform width (σ) from GLAS parameters only $r = 83$ for the selected GLAS footprints.

In order to have a self-sufficient approach to slope correction, which relies only on GLAS waveform parameters for estimating $H_{\max}$, the second method of terrain correction uses a multiple regression with the GLAS waveform width σ replacing the AL terrain index g

$$H_{\max} = 1.65^*w - 1.44^*\sigma \qquad (5)$$

with $R^2 = 0.91$ and rmse $= 3.92$ m. Both coefficients are statistically significant $(p < 0.05)$.

These results were similar to those from (1) but, at this time, without relying on a terrain index from AL data within the GLAS footprints. This suggests that the potential for estimating $H_{\max}$ from GLAS waveform parameters can be achieved by using only the waveform width (σ), which becomes wider as the terrain slope in the footprint increases [5]. The analysis was repeated using the bootstrap analysis for the multiple regressions previously mentioned, and for the first method, this resulted in

$$H_{\max} = 1.53^*w - 1.36^*g \qquad (6)$$

with $R^2 = 0.95$ and rmse $= 3.6$ m (both with $p < 0.05$).

The second method gave

$$H_{\max} = 2.12^*w - 1.91^*\sigma \qquad (7)$$

with $R^2 = 0.95$ and rmse $= 2.42$ m. Both coefficients are statistically significant $(p < 0.05)$.

The results demonstrate the validity of estimating $H_{\max}$ from GLAS waveform parameters directly using statistics based on multiple regression analysis (Fig. 4).

A limitation of this study is the limited sample size, which is further reduced by the stratification according to slope. It was constrained by the logistics of the fieldwork in a difficult

environment, but such studies would be more statistically robust if more GLAS footprint locations were visited and sampled in the field. Nevertheless, the statistical inference presented here was confirmed by bootstrapping.

An additional possible limitation of this study is that, over the five years between the airborne and spaceborne LiDAR acquisitions, the savanna vegetation is likely to have changed to a certain degree. Fires, however, rarely destroy the woody vegetation in the Kruger Park, because many trees are adapted to fire. Megaherbivores such as elephants may have destroyed some trees. While tree growth tends to be slow, some growth will have happened. These factors introduce additional uncertainty into the data analysis, which implies that, for concurrent data collection, the quality of fit is likely to be better than that reported.

IV. CONCLUSION

The results of the validation of maximum vegetation canopy height derived from GLAS and AL have shown that the direct method can be used in moderate terrain while the statistical method of terrain correction should be applied in steep terrain to obtain accurate estimates of $H_{\max}$. Better results could be achieved for estimating $H_{\max}$ using only the GLAS waveform parameter "waveform extent" and "waveform width" over sloped areas.

REFERENCES

[1] M. Nilsson, "Estimation of tree heights and stand volume using an airborne LiDAR system," *Remote Sens. Environ.*, vol. 56, no. 1, pp. 1–7, Apr. 1996.

[2] R. Nelson, R. Oderwald, and T. G. Gregoire, "Separating the ground and airborne laser sampling phases to estimate tropical forest basal area, volume, and biomass," *Remote Sens. Environ.*, vol. 60, no. 3, pp. 311–326, Jun. 1997.

[3] M. A. Lefsky, W. B. Cohen, D. J. Harding, G. G. Parker, S. A. Acker, and S. T. Gower, "Lidar remote sensing of above-ground biomass in three biomes," *Global Ecol. Biogeogr.*, vol. 11, no. 5, pp. 393–399, 2002.

[4] J. A. B. Rosette, P. R. J. North, and J. C. Suarez, "Vegetation height estimates for a mixed temperate forest using satellite laser altimetry," *Int. J. Remote Sens.*, vol. 29, no. 5, pp. 1475–1493, 2008.

[5] G. Sun, K. J. Ranson, D. S. Kimes, J. B. Blair, and K. Kovacs, "Forest vertical structure from GLAS: An evaluation using LVIS and SRTM data," *Remote Sens. Environ.*, vol. 112, no. 1, pp. 107–117, Jan. 2008.

[6] S. C. Popescu, K. Zhao, A. Neuenschwander, and C. Lin, "Satellite LiDAR vs. small footprint airborne LiDAR: Comparing the accuracy of aboveground biomass estimates and forest structure metrics at footprint level," *Remote Sens. Environ.*, vol. 115, no. 11, pp. 2786–2797, Nov. 2011.

[7] L. I. Duncanson, K. O. Niemann, and M. A. Wulder, "Estimating forest canopy height and terrain relief from GLAS waveform metrics," *Remote Sens. Environ.*, vol. 114, no. 1, pp. 138–154, Jan. 2010.

[8] Q. Chen, "Retrieving vegetation height of forests and woodlands over mountainous areas in the Pacific coast region using satellite laser altimetry," *Remote Sens. Environ.*, vol. 114, no. 7, pp. 1610–1627, Jul. 2010.

[9] Q. Chen, D. Baldocchi, P. Gong, and M. Kelly, "Isolating individual trees in a savanna woodland using small footprint lidar data," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 8, pp. 923–932, Aug. 2006.

[10] S. R. Levick and K. H. Rogers, "Structural biodiversity monitoring in savanna ecosystems: Integrating LiDAR and high resolution imagery through object-based image analysis," in *Proc. Object Based Image Anal.*, 2008, pp. 477–491.

[11] J. Wu, J. A. N. Van Aardt, G. P. Asner, R. Mathieu, T. Kennedy-Bowdoin, D. Knapp, K. Wessels, B. F. N. Erasmus, and I. Smit, "Connecting the dots between laser waveforms and herbaceous biomass for assessment of land degradation using small-footprint waveform lidar data," in *Proc. IGARSS*, 2009, pp. II-334–II-337.

[12] A. Baccini, N. Lapote, S. J. Goets, M. Sun, and H. Dong, "A first map of tropical Africa's above-ground biomass derived from satellite imagery," *Environ. Res. Lett.*, vol. 3, no. 4, p. 045011, Oct.–Dec. 2008.

[13] D. J. Harding and C. C. Carabajal, "ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure," *Geophys. Res. Lett.*, vol. 32, no. 21, pp. 1–4, Nov. 2005.

[14] M. A. Lefsky, M. Keller, Y. Pang, P. B. De Camargo, and M. O. Hunter, "Revised method for forest canopy height estimation from Geoscience Laser Altimeter System waveforms," *J. Appl. Remote Sens.*, vol. 1, no. 1, p. 013537, Sep. 2007.

[15] J. A. Rosette, P. R. J. North, J. C. Suárez, and J. D. Armston, "A comparison of biophysical parameter retrieval for forestry using airborne and satellite LiDAR," *Int. J. Remote Sens.*, vol. 30, no. 19, pp. 5229–5237, Oct. 2009.

[16] J. Boudreau, R.F Nelson, H. A. Margolis, A. Beaudoin, L. Guindon, and D. S. Kimes, "Regional aboveground forest biomass using airborne and spaceborne LiDAR in Québec," *Remote Sens. Environ.*, vol. 112, no. 10, pp. 3876–3890, Oct. 2008.

[17] M. A. Lefsky, D. J. Harding, M. Keller, W. B. Cohen, C. C. Carabajal, F. Del Bom Espirito-Santo, M. O. Hunter, and R. de Oliveira, Jr., "Estimates of forest canopy height and aboveground biomass using ICESat," *Geophys. Res. Lett.*, vol. 32, no. 22, pp. 1–4, Nov. 2005.

[18] R. Nelson, J. Boudreau, T. G Gregoire, H. Margolis, E. Nsset, T. Gobakken, and G. Ståhl, "Estimating Quebec provincial forest resources using ICESat/GLAS," *Can. J. Forest Res.*, vol. 39, no. 4, pp. 862–881, Apr. 2009.

[19] L. Gillson and K. I. Duffin, "Thresholds of potential concern as benchmarks in the management of African savannahs," *Philos. Trans. Roy. Soc. B, : Biol. Sci.*, vol. 362, no. 1478, pp. 309–319, Feb. 2007.

[20] A. C. Brenner, H. J. Zwally, C. R. Bentley, B. M. Csatho, D. J. Harding, M. A. Hofton, J. B. Minster, L. A. Roberts, J. L. Saba, R. H. Thomas, and Y. Yi, "Geoscience Laser Altimeter System (GLAS)—Derivation of range and range distributions from laser pulse waveform analysis for surface elevations, roughness, slope, and vegetation heights," Algorithm Theoretical Basis Document - Version 4.1 2003.

[21] H. J. Zwally, 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, "ICESat's laser measurements of polar ice, atmosphere, ocean, and land," *J. Geodyn.*, vol. 34, no. 3/4, pp. 405–445, Oct./Nov. 2002.

[22] D. J. Harding, M. A. Lefsky, G. G. Parker, and J. B. Blair, "Laser altimeter canopy height profiles methods and validation for closed-canopy, broadleaf forests," *Remote Sens. Environ.*, vol. 76, no. 3, pp. 283–297, Jun. 2001.

[23] S. A. Hinsley, R. A. Hill, P. E. Bellamy, and H. Balzter, "The application of LiDAR in woodland bird ecology: Climate, canopy structure and habitat quality," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 12, pp. 1399–1406, Dec. 2006.