Large-scale retrieval of leaf area index and vertical foliage profile from the spaceborne waveform lidar (GLAS/ICESat)

Hao Tang, Ralph Dubayah, Matthew Brolly, Sangram Ganguly, Gong Zhang

Department of Geographical Sciences, University of Maryland, College Park, USA
School of Environment and Technology, University of Brighton, United Kingdom
Bay Area Environmental Research Institute (BNERI), West Sonoma, USA
NASA Ames Research Center, Moffett Field, USA

Abstract

Leaf area index (LAI) and canopy vertical profiles are important descriptors of ecosystem structure. The Geoscience Laser Altimeter System (GLAS) on board ICESat (Ice, Cloud, and Land Elevation Satellite) provided three-dimensional observations that can be used to derive these canopy structure parameters globally. While several canopy height products have been produced globally from GLAS, no comparable data sets for LAI and canopy profiles exist across large areas. In this study we develop a physically based method of retrieving LAI and vertical foliage profiles (VFPs) from GLAS observations over the entire state of California, USA. This method refines lidar derived LAI and VFP through a recursive analysis of GLAS waveforms. The model was then transferred to the Montane forests of the Sierra Nevada using GLAS sensor data. The comparison between GLAS-derived LAI data and airborne lidar data demonstrated the more general capability of our algorithm to provide total LAI and VFP across biomes. The logical extension of these efforts is a further application of our methods over much larger areas, and is the goal of our work presented here.

1. Introduction

Leaf area index (LAI) and vertical foliage profile (VFP, or foliage height profile) are important biophysical variables in terrestrial ecosystems. Recent studies have reviewed the importance and potential applications of LAI and VFP derived from large footprint waveform lidar (Tang et al., 2012, 2014) and have shown the efficacy of a physical model to derive these profiles from waveform lidar data when compared to destructively sampled profiles in a tropical rainforest (Tang et al., 2012). This model was then transferred to the Montane forests of the Sierra Nevada using GLAS sensor data. The comparison between GLAS-derived LAI data and airborne lidar data demonstrated the more general capability of our algorithm to provide total LAI and VFP across biomes (Tang et al., 2014). The logical extension of these efforts is a further application of our methods over much larger areas, and is the goal of our work presented here.

Large scale derivation of GLAS LAI and VFP products has the potential to serve as a source of validation data for passive optical data sets, as well as providing needed canopy information that may be used within ecosystem and other models. While observations from airborne lidar sensors have been used to derive both LAI and VFP, these data are limited spatially. Demonstration of the viability of using space-based retrievals of these from lidar over large areas opens the possibility of enhanced descriptions of the vertical organization of canopy elements that play large roles in the transfer of energy and mass between the surface and atmosphere in ecosystem models. For example, there currently exists no regional data set of the LAI profiles, let alone for areas as large as states and beyond. Providing such data would improve our understanding of LAI structure and dynamics, its role in terrestrial gross primary production (GPP) (Kotchenova et al., 2004), and global carbon cycling (Houghton, 2007). Furthermore, foliar profiles have long been postulated to have an impact on habitat quality, species...
richness and abundance (Goetz, Steinberg, Dubayah, & Blair, 2007). A
global data set of such profiles would be exceptionally useful for clarify-
ing the roles of LAI and VFP in these areas, as well as providing the
means to explore the impact of climatic, edaphic, and human impacts
on their magnitudes and variability.

The overall goal of this paper is to demonstrate large-scale LAI and
VFP retrievals using GLAS data for the entire state of California. First,
we describe the utilized inputs, including data from GLAS, MODIS and
Landsat. We next briefly review details of our algorithm for deriving
LAI and VFP from GLAS, initially presented in Tang et al. (2012) and
implement our method to create footprint level LAI and VFP estimates
from GLAS data over California. Our results include statistical analysis
of GLAS LAI across environmental gradients (e.g. land cover type and el-
evation) and comparative analysis of GLAS and Landsat LAI retrievals for
California.

2. Data and methods

2.1. GLAS

To this day GLAS has been one of the few operational spaceborne
lidar instruments intended for global observations of the Earth. The
GLAS system was developed for the ICESat mission and operated be-
tween 2003 and 2009. The primary objective of GLAS was to measure
polar ice-sheet dynamics with the mission scope then extended to mea-
ure the height of vegetation canopies (Harding and Carabajal, 2005).
The GLAS sensor emits energy with wavelength of 1064 nm at a frequen-
cy of 40 Hz, and records the returned waveform from a ~65 m footprint
(Abshire et al., 2005). Rather than providing wall-to-wall ob-
ervation, GLAS yields individual footprint data with a centroid separa-
tion distance of ~165 m (Neuenschwander, Urban, Gutierrez, & Schutz,
2008). Considering that the ICESat satellite primarily targeted polar re-
gions, data acquired over global terrestrial ecosystems are quite sparse
with large distances across track at equatorial and mid-latitudes. The
GLAS data have been used to produce several global canopy height
products (Lefsky, 2010; Los et al., 2012; Simard, Pinto, Fisher, & Bac-
cini, 2011).

The GLA01 and GLA14 data campaigns within the entire state of
California from 2003 to 2007 were used in this study (Fig. 1). GLA01
typically includes a 544-bin recorded waveform at a vertical resolution
of 1 ns (15 cm) for land surface products. GLA14 products are accurate
fits of the GLA01 waveform using up to 6 Gaussian peaks to represent
the data, in addition to providing surface elevation and footprint cen-
troid coordinates (Harding and Carabajal, 2005). We did not use data as-
related to the campaigns Laser 1A or 2A due to an acknowledged
signal truncation problem (Harding and Carabajal, 2005). Low energy
shots (waveform peak energy < 0.5 V) were filtered from the studied
data to ensure the best retrieval quality.

2.2. Ancillary input data

In our retrieval approach terrestrial vegetation (excluding grass-
lands and croplands) was classified into nine types based on the
MODIS land cover type product (MCD12Q1) at 500 m resolution. The
nine types were evergreen needleleaf forest, evergreen broadleaf forest,
deciduous needleleaf forest, deciduous broadleaf forest, mixed forest,
closed shrubland, open shrubland, woody savanna, and savanna. This
categorization followed the IGBP scheme of the MODIS land-cover
map (MCD12). The overall accuracy of MODIS Collection 5 Global
land-cover data classification was estimated to be about 75% globally
(Friedl et al., 2010).

Effective LAI derived from single angle looking remote sensing, such
as from GLAS, should be corrected to true LAI using vegetation foliage
clumping information. Such clumping effects could be ecologically
significant and can be quantified using a biophysical parameter known
as clumping index (Chen & Black, 1992). The lidar-derived effective
LAI can then be corrected into true LAI by dividing by the clumping
index. We built a clumping look-up table by assigning each land cover
class an average clumping index from multi-angular satellite POLDER
Chen, Menges, & Leblanc, 2005).

For the retrieval procedure presented in this study, MODIS LAI
(Myneni et al., 2002) was used as a filter to refine our GLAS LAI esti-
mates by assessing the validity of each GLAS LAI retrieval (more details
can be found in Section 2.4 and in Tang et al., 2014). The MODIS data
used for this purpose were obtained from the MCD15A2 Collection 5
data set acquired during summer (July and August) from 2004 to 2007
and excluded cloud contaminated or low quality pixels based on
associated QA data (Zhao, Heinsch, Nemani, & Running, 2005).
The maximum values of the temporal LAI series at each 1 km pixel were
then calculated to represent the potential maximum LAI thresholds.
Since MODIS LAI data may not represent the true maximum value
(at GLAS footprint scale) within a 1 km pixel, pixels with maximum
LAI values less than 1 were assigned the mean LAI value of its vegetation
type (according to the MCD12 land cover map). Pixels with a LAI value
greater than 6 were also re-assigned a value of 12 to correct the MODIS
saturation domain (Myneni et al., 2002). These processes will ensure a
proper upper boundary to filter GLAS LAI estimates.

2.3. LAI derived from Landsat

We used a 30 m Landsat LAI map over California as a comparative
data set. This map was produced from Landsat surface reflectance data
based on canopy spectral invariant theory (Ganguly et al., 2012). The
orthorectified Landsat data were acquired from the Global Land Survey
(GLS) 2005 with core acquisition dates from 2005 to 2006 (Gutman
et al., 2008). Surface reflectance data were then generated through the
Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS)
after applying radiance calibration and atmospheric correction (Masek
et al., 2006). LAI at each 30 m pixel was retrieved based on a look-up
table method, in accordance with MODIS LAI algorithms (Ganguly
et al., 2008; Myneni et al., 2002; Yang et al., 2006). A comparison of
Landsat-derived LAI with MODIS LAI showed reasonable agreement
(Ganguly et al., 2012).
2.4. LAI retrieval algorithm

LAI and VFP can be derived from the vertical gap probability distribution as a function of height based on the MacArthur–Horn method (Chen, Rich, Gower, Norman, & Plummer, 1997; MacArthur & Horn, 1969; Nilson, 1971). This technique is essentially the optical point-quadrat sampling and computes vertical distribution of Leaf Area Density (LAD, unit: m²/m³), which is the density of foliage at discrete height layers within the canopy. The vertical arrangement of LAD constitutes VFP, and an integration of LAD at all canopy layers gives the total leaf area, or total LAI.

Note that we do not specifically distinguish between PAI (Plant Area Index) and LAI here for several reasons. First, differences between LAI and PAI are generally small. For example, as shown in our previous work over a tropical forest in Costa Rica (Tang et al., 2012), destructively sampled stem-area was vastly smaller than leaf-area, with LAI comprising 93% of total Plant Area Index and LAI derived from LVIS waveforms agreed well with these profiles. Additionally, our comparison with ground-based methods of LAI derivation in the Sierra Nevada, California, such as terrestrial scanning lidar and optical methods, again showed good agreement (Tang et al., 2014). Lastly, we expect the great majority of reflected laser energy from near-nadir looking NIR lidar comes from leaves, not branches, and Monte-Carlo simulation results support this contention (Hancock, Lewis, Foster, Disney, & Muller, 2012).

We estimated the vertical gap probability from lidar waveforms using the Geometric Optical and Radiative Transfer (GORT) model (Ni-Meister, Jupp, & Dubayah, 2001; Tang et al., 2012). This method was successfully implemented at multiple sites to derive landscape level LAI and VFP data using the airborne lidar system LVIS (Laser Vegetation Imaging Sensor) (Blair, Rabine, & Hofton, 1999; Tang et al., 2012). Comparisons using different ground measurements (including destructively sampled data, hemispherical photos, LAI-2000 and terrestrial lidar) demonstrated that waveform lidar can provide accurate LAI and VFP estimates even in the presence of high LAI levels such as those expected in tropical rain forests (Tang et al., 2012; Zhao et al., 2011). Efforts were also made using similar methods to derive LAI from GLAS data. When using these models satisfactory results were achieved when comparing LAI-2000 or LAI estimates with those obtained from airborne lidar (Garcia et al., 2012; Luo, Wang, Li, & Xi, 2013). However these studies were limited to small and flat study areas featuring no significant topographic variation, which was found to have a significant impact on the accuracy of GLAS measurement (Pang, Li, Lefsky, Sun, & Yu, 2006; Simard et al., 2011).

In this study, we incorporated a 3-step recursive method developed by Tang et al. (2014) to retrieve LAI and VFP from GLAS data (Fig. 2). This method was specifically designed to facilitate the separation of canopy and ground energy in GLAS waveforms. By incorporating prior knowledge of the LAI distribution (e.g. maximum potential LAI value of certain biomes), initial GLAS LAI estimates obtained from the GORT model can be refined. Pre-processed MODIS LAI was applied as the maximum potential LAI threshold. The method begins with the identification of the first Gaussian fit, assumed to be the ground return and used to calculate LAI and VFP. This estimate is then judged to be valid or invalid using the available prior knowledge. If considered invalid, the method iterates using the next feasible Gaussian fit until threshold conditions are met.

Canopy height predicted by the lidar waveform was calculated as the distance between the beginning of the waveform signal, the first return above the noise threshold (3.5 × standard deviation of identified noise), and the center of the chosen Gaussian fit. Our previous work showed that using this recursive method, GLAS could provide accurate LAI estimates under favorable topographic conditions, while a moderate level of accuracy could be achieved even over steep slopes (Tang et al., 2014). We extended this method by doubling the number of Gaussian fits to the GLAS waveform associated with the GLA14 data to 12 through a Gaussian decomposition method (Hofton, Minster, & Blair, 2000). The
default 6-Gaussian-curve fit, in certain instances, could fail to capture the true ground position in the presence of complex topography or highly vegetated areas (Duncanson, Niemann, & Wulder, 2010). Within the outlined methodology an additional model was included to improve GLAS LAI estimates through accounting for spatial variability in the soil/leaf reflectance ratio. This ratio value, a required input parameter in the GORT model, was primarily applied to normalize the reflected canopy return and ground components in a GLAS waveform. Lefsky (1999) initially suggested a constant value of 2.0 for such a ratio while Ni-Meister et al. (2010) developed a method to estimate its value directly using waveform lidar data. The latter method was based on the assumption that this value did not change at local scales. Armston et al. (2013) further improved Ni-Meister’s model by building a simple linear regression model between canopy energy and ground return with the slope of the regression taking on the value of the leaf/soil reflectance ratio. In this study we applied Armston’s method at quarter-degree scale (~20 km) to calculate this parameter and its variability, assuming that the ratio value did not vary significantly over a homogenous forest or over similar land cover types. In keeping with Lefsky (1999) a default value of 2.0 was assigned to quarter-degree cells with different land cover types or those with poor regression estimates ($r^2 < 0.2$ and $N < 10$).

Finally, we performed three other analyses. First, estimates of true LAI at the footprint scale were derived from the effective GLAS LAI values using the clumping-biome look-up table, noting that true LAI should be consistently greater than effective LAI (Chen et al., 2005). Secondly, we compared total LAI values derived from GLAS with Landsat LAI data. This was not a validation, per se, but rather a means to explore how the magnitude and variability of lidar-derived LAI compared with a commonly used passive optical product at fine scales. Lastly, much of California is mountainous. It was therefore of interest to explore if differences in GLAS vs. Landsat LAI were a function of topography as derived from the void-filled 90 m resolution SRTM DEM data (Reuter, Nelson, & Jarvis, 2007), as both products would be affected by topography, but in different ways.

3. Results

We derived a total of 16,529 GLAS LAI and VFP observations. Using the method of Armston et al. (2013), 42% of the GLAS footprints had a derived soil/leaf reflectance ratio different from the default of 2.0 (58%).

3.1. GLAS LAI

The distribution of GLAS LAI over California (Fig. 3) was highly skewed towards lower values of LAI with more than 50% of the footprints less than 2 (the median value was 1.76). This result was reasonable given that there were a large number of GLAS footprints occurring over land cover with expected low LAI values, such as savanna and shrubland.

The GLAS LAI values were grouped by land cover type according to our biome map (Fig. 4). Evergreen needleleaf forest and woody savanna comprised the highest frequency of GLAS observation numbers. Mixed forest showed the highest median LAI value (3.20) against all other biomes, followed by evergreen needleleaf forest (2.62). Woody savanna was observed to have a higher LAI value (1.23) than shrublands (1.05 for closed and 0.75 for open) and savanna (0.91) (and these differences were significant at $P < 0.001$). Evergreen broadleaf forest, deciduous needleleaf forest and deciduous broadleaf forest did not have sufficient observation numbers to perform a robust statistic analysis (results of all paired T-tests are given in Supplementary Table 1).
GLAS LAI was also stratified by elevation (Fig. 5). The LAI values did not differ significantly in lower elevation groups (from 250 m to 1000 m) (all P > 0.1). A linear regression analysis showed that increasing altitude led to a significant ($r^2 = 0.91, P < 0.001$) but slow decrease in LAI values ($\Delta$LAI $= 0.91$ per km elevation change) accompanied by a general reduction in frequency (results of paired T-test in Supplementary Table 2).

### 3.2. GLAS vertical foliage profiles (VFPs)

Vertical foliage profiles were created by averaging GLAS shots for each different land cover type (Fig. 6). Here VFP was expressed as the vertical distribution of Leaf Area Density (LAD). These averaged profiles exhibited a peak of foliage density at approximately 5 m height for forest and values closer to ground level (2 m or lower) for savanna and shrubland. Foliage density of savanna and shrubland diminished quickly above heights of 10 m and dropped to almost 0 when reaching about 20 m on average. In the mid-story range (5 m–10 m) differences between open shrubland, savanna and woody savanna were significant (P < 0.01). Evergreen needleleaf and mixed forests had a similar vertical foliage distribution as would be expected if the mixed forest was heavily comprised of evergreen needleleaf (paired T-test results in Supplementary Table 3).

Forest vertical strata were produced from the original GLAS VFP data (0.15 m vertical resolution) through classification into four height groups: 0–5 m for understory, 5–10 m for lower mid-story, 10–20 m...
for upper mid-story, and >20 m for upper-story (Table 1). Similar analysis to that shown in Section 3.1 was then performed on the forest strata to examine LAI variability as a function of land cover (Fig. 7). Differences between forest (needleleaf or mixed forest) and shrub (or savanna) were significant across all forest strata (all \( P < 0.001 \)). There were also significant differences at the layer at 10–20 m (\( P = 0.01 \)) and that above 20 m (\( P < 0.001 \)) between needleleaf forest and mixed forest. However, we could not identify any significant difference in understory (0 m–5 m) or lower mid-story (5 m–10 m) between those two forest types (\( P = 1.0 \) and \( P = 0.14 \) respectively).

Canopy layers displayed a consistent decreasing trend of total LAI with elevation (Fig. 8), although the overall decreasing rates were slightly different for each layer (\( \Delta \text{LAI} = 0.19, 0.23, 0.30 \) and 0.19 per 1 km elevation change respectively). Despite the general decreasing trend, median values of upper-story LAI appeared to reach maximum values at an altitude of ~1000 m. However, the differences among this layer and those of the adjacent elevation strata (groups of 250, 500, 750 and 1000 m) were not significant (\( P > 0.05 \) for all six comparisons. See Supplementary Table 4).

### 3.3. GLAS vs. Landsat

There was a fair agreement between GLAS LAI and Landsat LAI at the GLAS footprint scale (Fig. 9), with \( r^2 = 0.34 \), bias = 0.26, and a RMSD (Root Mean Square Difference) = 1.85. Compared with GLAS, Landsat LAI was larger at lower values of LAI (<2) but then saturated quickly at around LAI = 5. The histogram of differences (Fig. 10) between GLAS and Landsat is skewed because of large positive differences at high LAI (a long tail to the right), but there is also a large number of overestimates of Landsat relative to GLAS in the lower LAI regions.

While the overall bias is small, this was mostly a serendipitous canceling out of systematic overestimates and underestimates (saturation) (as opposed to random variations). An example of the fine scale variability of along-track GLAS LAI estimates and Landsat is shown in Fig. 11.

### 3.4. Slope analysis

There are well-documented impacts of slope on large footprint lidar waveforms (Pang et al., 2006; Simard et al., 2011). We therefore

---

**Table 1**

Statistics of GLAS LAI and LAI strata by biome type (ENF = evergreen needle forest, EBF = evergreen broadleaf forest, DNF = deciduous needleleaf forest, DBF = deciduous broadleaf forest).

<table>
<thead>
<tr>
<th>Biome type</th>
<th>( N )</th>
<th>Total LAI</th>
<th>LAI 0–5 m</th>
<th>LAI 5–10 m</th>
<th>LAI 10–20 m</th>
<th>LAI &gt; 20 m</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (±SD)</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>ENF</td>
<td>9899</td>
<td>3.28 ± 2.46</td>
<td>0.76</td>
<td>0.83</td>
<td>1.07</td>
<td>0.64</td>
</tr>
<tr>
<td>EBF</td>
<td>22</td>
<td>2.94 ± 1.58</td>
<td>0.97</td>
<td>0.91</td>
<td>0.76</td>
<td>0.31</td>
</tr>
<tr>
<td>DNF</td>
<td>40</td>
<td>1.13 ± 1.06</td>
<td>0.32</td>
<td>0.31</td>
<td>0.31</td>
<td>0.19</td>
</tr>
<tr>
<td>DBF</td>
<td>8</td>
<td>1.33 ± 1.41</td>
<td>0.48</td>
<td>0.41</td>
<td>0.33</td>
<td>0.11</td>
</tr>
<tr>
<td>Mixed forest</td>
<td>615</td>
<td>3.66 ± 2.38</td>
<td>0.80</td>
<td>0.90</td>
<td>1.19</td>
<td>0.78</td>
</tr>
<tr>
<td>Closed shrub</td>
<td>147</td>
<td>1.26 ± 1.03</td>
<td>0.47</td>
<td>0.39</td>
<td>0.31</td>
<td>0.09</td>
</tr>
<tr>
<td>Open shrub</td>
<td>649</td>
<td>0.92 ± 0.82</td>
<td>0.35</td>
<td>0.28</td>
<td>0.21</td>
<td>0.07</td>
</tr>
<tr>
<td>Woody savanna</td>
<td>3921</td>
<td>1.51 ± 1.21</td>
<td>0.49</td>
<td>0.46</td>
<td>0.42</td>
<td>0.14</td>
</tr>
<tr>
<td>Savanna</td>
<td>1228</td>
<td>1.14 ± 0.89</td>
<td>0.43</td>
<td>0.39</td>
<td>0.26</td>
<td>0.06</td>
</tr>
</tbody>
</table>

---

**Fig. 7.** GLAS LAI stratified by canopy position. Note that differences between needleleaf forest and mixed forest were only significant at the 10–20 m layer and >20 m layer (\( P = 0.003 \) and \( P < 0.001 \) respectively). Note differing y-axis scales.
analyzed the LAI differences between GLAS and Landsat as a function of slope to see what role, if any, slope might play in explaining them. We applied a linear regression model where $\Delta \text{LAI} = f (\text{slope})$ but found no relationship ($r^2 = 0.01$ and standard error (RSE) = 1.76):

$$\Delta \text{LAI} = 0.02 \times \text{Slope} - 0.2.$$ 

We also grouped LAI differences into eight different slope ranges from 0 to 40° with a 5° interval to examine the agreement of Landsat and GLAS LAI. The $r^2$ and RMSD were calculated within each slope range (Fig. 12). The $r^2$ generally decreased from 0.3 to about 0.2 as the slope increased up to 40°. The RMSD doubled from about 1 to a maximum value over 2 at slopes greater than 20°.

**Fig. 8.** Variability of GLAS LAI strata as a function of elevation for different canopy positions. The decreasing patterns were consistent in the 0–5 m, 5–10 m, 10–20 m and >20 m groups with slightly different rates. Note differing y-axis scales.

**Fig. 9.** Density scatter plot of Landsat LAI and GLAS LAI over California. The comparison reveals a fair agreement between the two data sets, but Landsat appears to saturate at about LAI = 5, and overestimates lower values relative to GLAS. Kernel density color bar refers to the distribution of LAI pairs with darker color indicating more clustered footprints.

**Fig. 10.** Histogram of LAI difference between GLAS and Landsat. The red dashed line gives the bias (0.26). However, this low bias does not reflect the systematic differences apparent in Fig. 9. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
4. Discussion

Lidar technology has been proven to provide accurate measurements of many ecosystem structural parameters. Regional mappings of canopy height, LAI and forest strata have been deployed across all major biomes using airborne lidar data (Asner et al., 2012; Morsdorf, Kotz, Meier, Itten, & Allgower, 2006; Tang et al., 2012; Wulder, Han, White, Sweda, & Tsuzuki, 2007; Zhao & Popescu, 2009). However, large-scale measurements of three-dimensional forest structure remain a major challenge, and this difficulty is mainly because of the limited availability of observations. One possible solution is to measure structure from space-based sensors. In this study we attempted to derive important biophysical structure in the form of LAI and VFP over a large region.

Our previous work has shown strong relationships between GLAS LAI and LAI derived from airborne waveform lidar (Tang et al., 2014) but was based on unique validation data (in the form of destructively sampled canopy profiles), or ground-based optical methods that are exceptionally limited spatially. These previous studies gave us confidence in the theoretical basis and efficacy of our models. For our work here, across an entire state, a comparable set of validation data for assessing our derivations does not exist. We therefore were limited to assessing differences in total LAI using passive optical data from Landsat. We found an overall fair agreement between derived GLAS LAI and 30 m Landsat LAI map. Despite their fundamental differences in sensor design and retrieval algorithms, this is quite an encouraging result. However, compared with GLAS data, Landsat was observed to both underestimate (saturate) LAI values for high LAI (Fig. 9) and overestimate LAI for lower values. There are variations between the methods of observation that could lead to differences in total LAI such as system design (passive vs. active) and pixel size (30 m pixel vs. ~65 m eclipse for GLAS). Note that there was also a temporal discrepancy between the two data sets (2005 only vs. 2003–2007).

Which is “correct”, GLAS or Landsat? While we cannot answer this question definitively, saturation of LAI by passive optical sensors is well known (Shabanov et al., 2005; Yang et al., 2006). We further do not know of a plausible physical explanation that would lead us to suspect that GLAS observations for high LAI areas are overestimates. A plausible hypothesis could be that passive optical sensor data do not detect energy from lower canopy portions or understory in dense forests. Progress in this area will depend on the availability of improved field data, coupled with terrestrial scanning and airborne lidar observations and radiative transfer simulation modeling.

Comparisons with Landsat aside, there are several potential sources of error in our GLAS LAI algorithm. In particular there are two major input parameters involved in the GLAS-LAI algorithm: a clumping index based on MODIS land cover classification and the leaf/ground reflectance ratio. The use of such ancillary data is the essence of data fusion, but the impacts of errors in the parameters on LAI variations need
to be assessed. In our algorithm, MODIS land cover classification provides the clumping information required for converting GLAS effective LAI to total LAI values. For any given effective LAI value, error in true LAI varies linearly with clumping index. For example, a 10% error in clumping index would generally lead to a 10% error of true LAI value but does not change the effective LAI value. Errors in clumping index, in turn, are caused by errors in land cover classification, the spatial variability of clumping index within land cover types, and the accuracy of the fundamental derivations of the index itself. Here we do not account the spatial variability of clumping index within biomes because the variation of clumping index for any given biome is small (Chen et al., 2005).

Leaf/ground ratio is the other important input parameter of our model. Methods have been developed to derive this value relying on lidar data themselves (Armstrong et al., 2013; Ni-Meister et al., 2010). However, this ratio cannot be obtained for all GLAS footprints because the mathematical solutions do not converge. In these cases we were left to assign a default value of 2.0. Sensitivity of LAI with respect to leaf/ground ratio value has been analyzed in our previous work (Tang et al., 2012). We found that the effect of the ratio on LAI is not large. For example, expected variations of the ratio would introduce an error of about ±0.5 for LAI ≈ 4, and this impact decreases as LAI gets smaller. Uncertainty introduced from the ratio should constitute roughly 10% or less of total LAI given an average LAI value of ~2 in California.

We discovered little relationship between LAI differences and slope with our regression analysis, while there was a trend of decreased $r^2$ (and increased RMSD) with steeper slope (Fig. 12). The relationship was not significant largely because of the distribution of GLAS-slope data: more than 80% of the data had a slope less than 20°. But when we grouped GLAS data by slope ranges, the slope effect on LAI difference is more evident. Steep slopes affect the lidar return by spreading the Gaussian ground return and blending canopy elements into the ground return. Thus, estimates of LAI from GLAS near the ground (below 5 m) may sometimes be in error. Tang et al. (2014) showed that for the Sierra Nevada, differences between GLAS LAI and airborne waveform lidar LAI were minimized for slopes less than about 15°. This suggests that our iterative method of finding LAI overcomes impacts of slopes less than this magnitude, or effects of slope are overwhelmed by other sources of error or variation.

Our analysis also revealed that GLAS data captures spatial LAI variability across different environmental gradients. For example, we found a significant relationship between GLAS LAI distribution and elevation groups (Fig. 5). An increase of 1 km elevation would typically lead to a decrease of ~0.9 LAI unit. This result agrees with published findings from previous studies of LAI and altitude gradients (Luo et al., 2004; Moser, Hertel, & Leuschner, 2007; Pfeifer, Goncino, Disney, Pellikka, & Marchant, 2012). We also found a high variability of GLAS LAI values within and across different land cover types (Fig. 4).

One of the most exciting aspects of our work is the ability to derive LAI profiles at the landscape scale to develop a better understanding of their differences across land cover types. For example, we found that foliage density peaks increase in both height location and maximum value, starting from shrubland, savanna, woody savanna and going to forest (Fig. 6). This finding is also supported by the distribution of LAI within forest vertical strata (0–5 m, 5–10 m, 10–20 m, and >20 m, as shown in Fig. 7). In particular, we concluded that total LAI difference between needleleaf and mixed forest was due to their difference in upper-story rather than understory (Figs. 4 and 6). Moreover, we discovered that LAI strata contribute differently towards the total LAI change across elevation gradients (Fig. 8). Our results suggest that such data should improve our ability to evaluate the impacts of climatic and edaphic factors, as well as disturbance on spatial and vertical organization of canopies by revealing the existence of generalizable relationships between canopy structural information and these factors.

The availability of LAI profiles across large scales should be valuable for both the initialization and validation of ecological models by providing more robust and realistic representations of canopy structure. Previous studies suggest an underestimate of more than 50% of GPP when neglecting vertical foliage stratification (Aber, 1979; Kotchenova et al., 2004; Sprintsin, Chen, Desai, & Gough, 2012). A new canopy radiative transfer model (ISBA–A–gs) then recommends a minimum of 10 canopy layers to better estimate FAPAR and GPP (Carrer et al., 2013). A fully derived and validated continental or global scale VFP product from GLAS, using a similar process to those developed in this study, would therefore be of great interest.

There is one note of caution considering the initialization of such models with lidar-derived profiles. If the models are run at a coarse modeling scale, e.g. say 50 km × 50 km as is typical of many carbon models, the use of an average profile for a land cover type would be misleading. As shown in Fig. 6, the average profiles show peaks very close to the ground partially due to the artifact of sampling. Hurtt et al. (2010) have shown that impact of spatial scale, either in the data through averaging, or in the model through coarse grid size cells, can have dramatic impacts on estimates of carbon flux, among others, between the canopy and the atmosphere when compared to the scale individual trees. Care must be taken to adequately represent the true vertical and spatial variability of the LAI profiles, either through shifts to finer model resolutions or sub-grid scale parameterizations of LAI. For example, individual-based ecosystem models, such as the Ecosystem Demography model (Hurtt et al., 2010), have now been implemented at 1 ha spatial resolution. Furthermore, at coarser modeling scales, ED can potentially ingest the actual probability distribution (PDF) of LAI profiles for a grid scale, as observed from lidar. Thus, the great power of a remote sensing approach to LAI profile derivation is that it captures this PDF directly. If applied globally to GLAS data, our algorithm would provide the first data set of LAI profiles across biomes for use in modeling efforts.

Similarly, such profiles also have the potential to help quantify habitat heterogeneity, a fundamental component of biodiversity studies. Relationship between habitat heterogeneity and vegetation structure from airborne lidar have been explored at local scales (Ferger, Schleuning, Hemp, Howell, & Böhning-Gaese, 2014; Goetz et al., 2007; Swatantran et al., 2012). Analyses using continental scale data sets, such as the Breeding Bird Survey (BBS) (Sauer et al., 2008), may be facilitated by the inclusion of GLAS LAI profiles in models of species richness and diversity (Culbert et al., 2013; Goetz, Sun, Zolkos, Hansen, & Dubayah, 2014).

5. Conclusion

In this study we have demonstrated the feasibility of deriving LAI and VFP from spaceborne waveform lidar over large areas. The state-level GLAS LAI product not only showed a fair correspondence with 30 m Landsat LAI maps produced over California, but also highlighted potential issues with the latter: namely, saturation at high LAI values relative to GLAS. This suggests that further analyses should be conducted to validate both data sets using other data, and from these, refinement of algorithms should then be applied for both active and passive optical remote sensing.

While our previous validation studies have been limited to tropical and coniferous forests, our work here spanned a range of biomes, land cover types and environmental gradients. Because our modeling is based on physical principles, we are optimistic that future validation studies in different biomes will provide further support of the accuracy of our methods. Independent of these activities, our results here provide a unique and valuable data set on two key environmental variables that has been previously been unavailable over large areas, and thus provides the basis for a continental and global LAI data set from GLAS. It is our belief that such data sets, once available, will be of a great importance in deepening our understanding of the role spatial and vertical canopy structure plays in ecosystem processes, and the factors which impact that structure through time.
Acknowledgments

This work was funded by NASA under grant NNX12AK07G (Dubayah) and an Earth And Space Science Graduate Fellowship NNX12AN43H (Dubayah/Tang). We thank Maosheng Zhao and Qing Ying for their help in processing MODIS data. We also thank the NSIDC (National Snow & Ice Data Center) User Services for the help on data acquisition and NASA Earth Exchange (NEX) for computing resources.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.rse.2014.08.007.

References


