Lidar-based mapping of leaf area index and its use for validating GLOBCARBON satellite LAI product in a temperate forest of the southern USA

Kaiguang Zhao *, Sorin Popescu

Spatial Sciences Lab, Department of Ecosystem Science and Management, Texas A&M University, College Station, TX-77840, USA

ARTICLE INFO

Article history:
Received 5 November 2008
Received in revised form 16 March 2009
Accepted 17 March 2009

Keywords:
Leaf area index
LAI
Airborne laser
Forest ecosystems
Lidar
GLOBCARBON
Carbon cycling
Hemispherical photograph
scale

ABSTRACT

Lidar provides enhanced abilities to remotely map leaf area index (LAI) with improved accuracies. We aim to further explore the capability of discrete-return lidar for estimating LAI over a pine-dominated forest in East Texas, with a secondary goal to compare the lidar-derived LAI map and the GLOBCARBON moderate-resolution satellite LAI product. Specific problems we addressed include (1) evaluating the effects of analysts and algorithms on in-situ LAI estimates from hemispherical photographs (hemiphoto), (2) examining the effectiveness of various lidar metrics, including laser penetration, canopy height and foliage density metrics, to predict LAI, (3) assessing the utility of integrating Quickbird multispectral imagery with lidar for improving the LAI estimate accuracy, and (4) developing a scheme to co-register the lidar and satellite LAI maps and evaluating the consistency between them. Results show that the use of different analysts or algorithms in analyzing hemiphotos caused an average uncertainty of 0.35 in in-situ LAI, and that several laser penetration metrics in logarithm models were more effective than other lidar metrics, with the best one explaining 84% of the variation in the in-situ LAI (RMSE = 0.29 LAI). The selection of plot size and height threshold in calculating laser penetration metrics greatly affected the effectiveness of these metrics. The combined use of NDVI and lidar metrics did not significantly improve estimation over the use of lidar alone. We also found that mis-registration could induce a large artificial discrepancy into the pixelwise comparison between the coarse-resolution satellite and fine-resolution lidar-derived LAI maps. By compensating for a systematic sub-pixel shift error, the correlation between two maps increased from 0.08 to 0.85 for pines (n = 24 pixels). However, the absolute differences between the two LAI maps still remained large due to the inaccuracy in accounting for clumping effects. Overall, our findings imply that lidar offers a superior tool for mapping LAI at local to regional scales as compared to optical remote sensing, accuracies of lidar-estimate LAI are affected not only by the choice of models but also by the absolute accuracy of in-situ reference LAI used for model calibration, and lidar-derived LAI maps can serve as reliable references for validating moderate-resolution satellite LAI products over large areas.

1. Introduction

As a key canopy structural characteristic, leaf area index (LAI) serves as an important input or state variable for a variety of process-based ecological and biogeochemical models, especially those pertinent to simulating energy and mass exchanges at the atmosphere-land interface or photosynthesis and respiration of ecosystems for carbon cycling (Turner et al., 2004). LAI is typically defined as the total one-sided area of green foliage per unit ground surface (Chen & Black, 1992). Both direct and indirect techniques, such as destructive sampling and optical methods, exist for measuring LAI in-situ. (Gower et al., 1999; Jonckheere et al., 2004). These in-situ techniques, however, are impractical for measuring LAI over large areas partly due to prohibitive costs (Cohen et al., 2003). Instead, researchers often resort to remote sensing for spatially-explicit mapping of LAI at landscape or regional levels. Reliable and accurate estimation of LAI, therefore, has become a primary task in exploiting the potential of remotely-sensed data for retrieval of biophysical variables, as demonstrated by both early studies in using optical imagery for estimating LAI and more recent efforts in mapping LAI with ranging measurements from lidar (Light Detection And Ranging) (Jensen et al., 2008; Morsdorf et al., 2006; Riano et al., 2004; Roberts et al., 2005).

Optical remote sensing of LAI relies on spectral sensitivities to changes in vegetative components. Strong relationships between LAI and some carefully selected vegetation indices are both observed from experimental data and revealed theoretically by physically-based canopy reflectance models (Chen & Cihlar, 1996; Myneni et al., 1997).
Among others, the utility of the normalized difference vegetation index (NDVI) for estimating LAI has been proven by extensive studies across various biomes using different remote sensing datasets such as Landsat TM/ETM+, MODIS, SPOT/VEGETATION and AVHRR (Eklundh et al., 2001; Garrigues et al., 2006; Tan et al., 2005; Yao et al., 2008). It is found that LAI-NDVI relationships not only depend on vegetation types but also vary seasonally and annually (Wang et al., 2005). Many studies report that NDVI saturates with high LAI, particularly for deciduous forests (Birky, 2001). At the global scale, several time series of moderate resolution LAI products have been generated from various satellite data, which, for example, include Terra/Aqua MODIS, GLOBCARBON (http://geofront.vgt.vito.be), and CYCLOPS (http://postel.mediasfrance.org). Accuracies of these products are of major concern to the scientific community. The improvement of LAI retrieval accuracy is limited by many interplaying factors such as the uncertainty of satellite reflectance measurements, the characterization of canopy architectures in retrieval algorithms, the natural variability of surface spectra and the mixture of species (Shabanov et al., 2005). To guide the informed use, practical efforts for validation or inter-comparison of these global satellite LAI products, especially GLOBCARBON and MODIS products, have been completed or are ongoing at a range of sites worldwide (Cohen et al., 2003; Garrigues et al., 2008; Wang et al., 2004). The major recognized difficulties in assessing the quality of these products include the limited number of representative validation sites and the scale discrepancies between in-situ and satellite measurements, although other factors, such as the mis-registration of satellite products with reference data, also will complicate the validation processes (Tan et al., 2006).

Recent advances in airborne laser scanners (commonly known as airborne lidar) bring a breakthrough in canopy remote sensing, with an enhanced capability of direct characterization of canopy vertical structures (Lim et al., 2003; Næsset, 2002; Popescu et al., 2004). Portions of lidar pulses penetrate into canopies and even strike on the ground, thus allowing for better characterization of all canopy layers, including understories (Nelson et al., 1988), and at the same time alleviating the saturation problem of optical remote sensing for forests of high LAI or biomass (Lefsky et al., 2002). The body of lidar literature on ecological and environmental studies is growing in such aspects as mapping terrain topography and estimating biophysical parameters such as biomass, canopy density, LAI and fuel parameters at various analysis units (e.g., individual tree, plot, stand, and woodland) (Brandtberg et al., 2003; Chen et al., 2006; Holmgren et al., 2003; Hopkinson & Chasmer, 2009; Mutlu et al., 2008; Næsset et al., 2005; Popescu & Wynne, 2004; Zhao et al., 2008a). A key factor concerning lidar-based estimation of forest attributes at scales above individual trees is the selection of appropriate lidar metrics as predictors and effective equations as model forms, preferably with certain physical meaning (Næsset, 2002; Zhao et al., 2009). Lidar metrics that have been previously investigated for LAI mainly include mean height, maximum height, percentile height, height of median energy, and canopy density metrics (e.g., ratio metrics), among others (Farid et al., 2008; Griffin et al., 2008; Hansen & Solberg, 2007; Jensen et al., 2008; Lovell et al., 2003; Morsdorf et al., 2006). In particular, ratio metrics such as laser penetration index and laser interception index prove useful for estimating LAI (Bariolotti et al., 2006). Kusakabe et al. (2005) also suggested the use of “mean free path” (penetration length into canopy) as a proxy for LAI.

Like optical remote sensing of LAI (Garrigues et al., 2006; Tian et al., 2002), estimating LAI from lidar data is also subject to scale issues, which include but are not limited to the selection of an “optimal” resolution at which to develop regression models as well as the scaling-up/down problems due to the scale-dependence of the regressed models (Patenaude et al., 2004; Zhao & Popescu, 2007). Though not indicated explicitly, almost all the lidar LAI models investigated previously are more or less scale-dependent, which is exemplified by the fact that the estimated LAI of a region by using a single metric extracted for the region does not equal the aggregated value of estimates over the sub-regions that partition the region (Zhao et al., 2009).

The primary goal of this study is to further explore the capability of discrete-return lidar for spatially-explicit mapping of LAI, with a secondary goal to examine the consistency between the lidar-derived LAI map and GLOBCARBON satellite LAI products over an eastern Texas forest. We addressed the following specific sub-problems: (1) to evaluate the effects of analysts and algorithms on uncertainties of in-situ LAI estimates derived from hemispherical photos (i.e., hemiphotos), (2) to investigate the effectiveness of a number of lidar metrics, including several newly proposed ones, for estimating LAI, (3) to assess the utility of integrating lidar with multispectral imagery (i.e., Quickbird-derived NDVI) for accuracy improvement of LAI estimates, and (4) to determine the extent to which mis-registration could affect the comparison between the lidar fine-scale and satellite moderate-resolution LAI maps as well as to develop a practical co-registration scheme for reducing the bias induced by registration errors in comparing the two LAI maps.

2. Materials

2.1. Study area

The study area is a 48-km² forested region in East Texas of the southern U.S. (30° 42′ N, 95° 23′ W) (Fig. 1). It mainly comprises pine plantations in various developmental stages, old growth pine stands in the Sam Houston National Forest with many of them having a natural pine stand structure, and upland and bottomland hardwoods. The major species include Loblolly pines (Pinus taeda L.) and deciduous trees such as water oak (Quercus nigra L.), red oak (Quercus falcata Michx.), sweetgum (Liquidambar styraciflua L.), and post oak (Quercus stellata Wangenh.). Much of the southern U.S. is covered by forest types similar to those of our study area, including similar species, productivity and patterns of land use and land cover. The area is characterized by a gentle topography and has an elevation varying from 62 m to 105 m with an average of 85 m.

2.2. Field measurements (hemiphoto)

Field work was undertaken from May to July in 2004. Hemispherical photographs (hemiphoto) were taken on 53 circular plots established across the study area, with plot centers geo-referenced by a differential GPS (Trimble Pathfinder). Of the 53 plots, 14 are mixed/ hardwood plots, and the remaining 39 are pine plots with 18 of them established in young pine plantations that have little variations of height and crown width and 21 in mature pine stands. These hemiphotos have a resolution of 3264 × 2448 pixels and were captured at 1.5 m above the ground using a horizontally-leveled CoolPix 8700 digital camera (Nikon) equipped with an FC-E9 fisheye lens converter (Nikon). Other ground inventory data, such as tree height, crown width and crown class, were also tallied (Popescu, 2007) but not directly used in this study.

2.3. Lidar dataset

The airborne laser scanner data were acquired during the leaf-off season in March 2004 with a Leica-Geosystems ALS50 flying at an average altitude of 1000 m by M7 Visual Intelligence of Houston, Texas. The lidar system was operated to record two returns per pulse, i.e., first and last, with a reported horizontal and vertical accuracy of 20–30 cm and 15 cm, respectively. The system was configured to scan +/− 10° from nadir, resulting in a swath of about 350 m wide on the ground. The acquired dataset features a full coverage of the study area from either of two perpendicular directions, i.e., with 19 flight lines in the north–south direction and 28 in the east–west direction, resulting in an average of 2.6 laser hits per m² (Fig. 1). From the raw
2.4. Quickbird multispectral imagery

A Quickbird (Digital Globe, Inc.) scene acquired in the spring of 2004 is available over our study area (Fig. 1b). The image has a spatial resolution of 2.4 m with four spectral bands, i.e., blue (450–520 nm), green (520–600 nm), red (630–690 nm), and NIR (760–900 nm). Radiometric calibration and ortho-rectification have been applied to the image. An inspection of ten conspicuous feature points revealed that the image and lidar CHM were well geographically registered with an average error less than 2.4 m. The coordinate system of this image is also UTM 15N with the WGS84 datum.

The purposes of incorporating the Quickbird image into this study are to extract thematic information for distinguishing forest types and to assess the utility of NDVI for estimating LAI when integrated with lidar. For the first purpose, we applied the maximum likelihood classifier to the image mainly for differentiating pines, hardwoods, mixed forests, and grassland. The classification produced an overall accuracy of 86.5% via an on-screen evaluation of a random subset of 200 test pixels.

2.5. Moderate resolution satellite LAI product — GLOBCARBON

The initiative of GLOBCARBON is to deliver global parameters derived from multiple satellite image data for providing inputs to carbon cycle models (Plummer et al., 2006). A subset of monthly GLOBCARBON LAI product over our study area for July 2004 was obtained from http://geofront.vgt.be/geosuccess. The GLOBCARBON LAI algorithm is based on a 4-scale bidirectional reflectance model (Deng et al., 2006): it also applies land cover-specific clumping indices to correct for foliage clumping at plant and canopy scales. The LAI subset we obtained was derived through the combined use of SPOT/VEGETATION and ENVISAT/AATSR images of July 2004 and was delivered at a 1.0-km spatial resolution in the Plate Carree projection. The co-registration scheme for overlaying this moderate resolution LAI map with a fine-scale lidar-derived map is detailed in a later section.

3. Methods

3.1. Hemispherical analysis for in-situ LAI

The hemiphotos were processed to compute in-situ LAI by two analysts as well as with two algorithms. The motive for examining the two factors (i.e., analyst and algorithm) is to better understand the uncertainties in the in-situ LAI. Specifically, we first employed two experienced analysts to perform binary segmentation on the 53 hemiphotos into sky or obscured pixels using the commercial package HemiView (Delta-T Devices Ltd., UK, 1999), for evaluating the effect of subjectivity on hemispherical analysis that is caused by the manual and interactive selection of segmentation threshold values. We then applied two LAI algorithms to the segmented hemiphotos in order to assess the sensitivity of in-situ LAI estimates to inversion algorithms. As a result, a total of four sets of in-situ LAI estimates (two analysts by two algorithms) were produced.

The two inversion algorithms for computing in-situ LAI, though with slightly different assumptions, are both based on the “gap-fraction” formula of the Beer’s law in a canopy:

\[ \tau(\theta, \phi) = \exp\left[-G(\theta, \phi)L/\cos(\theta)\right] \quad (1) \]

where \( L \) is LAI or, more precisely, effective LAI as shall be described later, \( \tau(\cdot) \) denotes the gap fraction in the direction \((\theta, \phi)\) with \( \theta \) and \( \phi \) being the zenith and azimuth angles, respectively, and \( G \) is the Ross–Nilson \( G \)-function that refers to the fraction of a unit foliage area projected onto a plane normal to \((\theta, \phi)\) (Nilson, 1999). Eq. (1) assumes that foliage elements are distributed randomly and independently. From the segmented hemiphotos, we obtained estimates of \( \tau(\theta, \phi) \) over given sectors or annuli, which were used to compute \( L \) by inverting Eq. (1) with the following two algorithms (Fig. 2a).
3.1.1. Algorithm 1 for computing in-situ LAI

Algorithm 1 employs a specific leaf angle distribution, i.e., the ellipsoidal distribution, so that

\[
G(\theta; x) = \frac{\sqrt{x^2 \cos^2 \theta + \sin^2 \theta}}{x + 1.702(x + 1.112)^{-0.708}}
\]  

(2)

where the leaf distribution is supposed to be azimuthally symmetrical so that \( G \) only depends on the zenith angle \( \theta \), and \( x \) is a constant parameterizing \( G \) (Campbell, 1986). By combining Eqs. (1) and (2), LAI and \( x \) can be simultaneously estimated via iterative optimization procedures in terms of minimizing the squared error between the gap fractions \( \tau \) observed from hemiphotos and those predicted from the model (Fig. 2b), as implemented in HemiView (Delta-T Devices Ltd., UK, 1999). In this study, the observed gap fractions \( \tau \) for Algorithm 1 were obtained on 18 × 8 sky sectors with 5° and 45° for zenith and azimuth divisions of the upper hemisphere, respectively (Fig. 2a).

3.1.2. Algorithm 2 for computing in-situ LAI

In Algorithm 2, we also assumed that the leaf angle distribution is azimuthally-independent but did not restrict its functional form. Instead, Algorithm 2 is based on the Miller's theorem (Nilson, 1999):

\[
\int_0^{\pi/2} G(\theta) \sin \theta d\theta = 0.5
\]  

(3)

which, combined with Eq. (1), gives an LAI formula,

\[
L = -2 \int_0^{\pi/2} \ln |\tau(\theta)| \cos \theta \sin \theta d\theta = \int_0^{\pi/2} l_w(\theta) d\theta
\]

(4)

where, for notational convenience, the integrand has been denoted by

\[
l_w(\theta) = -2 \ln |\tau(\theta)| \cos \theta \sin \theta.
\]

However, two practical difficulties exist when applying Eq. (4) to the observed \( \tau(\theta) \). First, some observed gap fractions, especially those on annuli at large zenith angles, are zero-valued, thus causing numerical overflows in the evaluation of \( \ln |\tau(\theta)| \). In previous studies, a remedy for this is to simply add 1.0 or 0.5 pixel of openness to the problematic annuli or sectors (van Gardingen et al., 1999); however, such addition appears arbitrary. Second, an appropriate quadrature procedure is necessitated to evaluate the integral of Eq. (4) based on the gap fractions \( \tau(\theta) \) observed only at several zeniths \( \theta \)'s. To solve these difficulties, in Algorithm 2 we fitted a nonparametric curve using the smoothing spline to all the valid observations of \( l_w(\theta) \), excluding the problematic ones of zero \( \tau(\theta) \). The fitted spline allows computing \( l_w(\theta) \) at any \( \theta \), including the \( \theta_i \) associated with problematic annuli. Then, the resulting spline curve was numerically integrated to estimate LAI according to Eq. (4) (Fig. 2c).

3.1.3. Effects of analysts and algorithms on in-situ LAI estimates

The effects of analysts and algorithms on the in-situ LAI estimates were assessed with the repeated measures ANOVA analysis (SAS Help and Documentation System, 2007) as well as in terms of root mean square differences (RMSD). Since each factor has two levels (i.e., two analysts denoted by A1 and A2, and two algorithms denoted by A1 and A2), two RMSDs for each factor can be obtained from the four sets of in-situ LAI by taking the root mean square of the differences between the two levels of one factor at each of the two levels of another factor. Then, the average

![Fig. 2](image-url)

**Fig. 2.** Illustration of in-situ LAI inversion algorithms: (a) A sample hemiphoto was partitioned into 18 × 8 sky sectors with a division of 5° and 45° for zenith and azimuth, respectively. Gap fractions \( \tau \) can be calculated over each sector \( \theta_j \) or each annulus \( \theta_i \) by counting the relative occurrence of sky pixels. (b) Algorithm 1 seeks an optimal “ellipsoidal distribution” model in the form of Eq. (2) by minimizing the sum of square errors between the observed \( \tau(\theta_i, \theta_j) \) and the fitted values of the model; the resulting model for (a) gives a value of 0.980 and 0.894 for \( x \) and LAI, respectively. For comparison, the observed gap fractions \( \tau(\theta_i) \) over annuli are also plotted (circles). (c) Algorithm 2 fits a smoothing spline to the observed \( l_w(\theta) \) from hemiphotos and then integrates the fitted curve to obtain LAI, which is 1.17 for the hemiphoto of (a). The two LAI estimates are different by 0.276.
of two RMSDs (ARMSD) is used to represent the variability caused by that factor. For example, ARMSD for the factor of analysts was calculated by

\[
\text{ARMSD}_{\text{analyst}} = \frac{\|L_{-a1} - L_{-a2}\| + \|L_{-a1} - L_{-a2}\|}{2}
\]

where the subscripts of indicate the analyst and algorithm employed for computing LAI, and \(\|\|\) denotes the RMSD between any two sets of LAI.

Due to the lack of ground-truth LAI data of higher accuracy, we could not determine which of the four sets of in-situ LAI (two analysts by two algorithms) is the most accurate or which of the two in-situ LAI algorithm is more effective, although the assumption of Algorithm 2 is less restrictive. As a compromise, the average of the four sets was used in the subsequent regression analysis relating in-situ LAI to lidar metrics.

3.2. Analysis of lidar data characteristics for model development

A better understanding of the lidar data characteristics helps to derive more meaningful lidar metrics. The lidar data of this study, collected by a first/last return scanner, consist of three types of laser returns (throughout the paper, “return” is used interchangeably with “hit” or “echo”): single return/echo (SR), which corresponds to a pulse that produces only one echo; first return/echo (FR) and last return/echo (LR), which are the first and last echoes of a pulse that has multiple echoes. In the later presentation, we strictly discriminate SR from FR, although SR can be loosely deemed as either FR or LR. On the other hand, in terms of the targets intercepted, laser hits can be roughly categorized into either ground or canopy hits: Ground hits are often identified as those that are below a prescribed height threshold, as such, they include not only the true ground hits of zero height but also near-ground understory hits. Canopy hits can be further divided into those from crown surfaces and those from inside or below crowns (Fig. 3).

For conceptual convenience, we presume that crown-surface hits can only be SRs or FRs, and inside-/below-crown hits (i.e., inside-canopy hits) can only be LRs (Fig. 3). Because a pulse is unlikely to produce two separable echoes within a certain short distance near the ground due to the instrument dead-time (Morsdorf, 2006), ground hits can only be SRs or LRs when using a small separating height threshold. However, if the threshold is relatively high (e.g., greater than 4 m), ground hits can be any type of returns and therefore may not represent true ground returns, and in such a case, we still termed them as ground hits only for conceptual convenience so as to compute penetration metrics by analogy to the Beer’s law, as detailed later.

The above classification scheme of laser hits is also illustrated in Fig. 3. In addition, the following relationships should roughly hold as to the number of each type of laser hits over a region (Sasaki et al., 2008):

\[
\begin{align*}
N_{\text{fr}} & = N_{\text{lr}} \\
N_{\text{pulse}} & = N_{\text{sr}} + N_{\text{fr}} = N_{\text{ar}} + N_{\text{lr}} \\
N_{\text{total}} & = N_{\text{sr}} + N_{\text{fr}} + N_{\text{lr}} = 2N_{\text{fr}}
\end{align*}
\]

where \(N_{\text{pulse}}\) and \(N_{\text{total}}\), i.e., numbers of laser pulses and laser hits, are different in that a pulse with two returns is counted only once in \(N_{\text{pulse}}\) while it is counted twice in \(N_{\text{total}}\) because of its two returns; the other subscripts in Eq. (7) should be self-explanatory. Of particular note, these relationships only roughly hold due to boundary effects, i.e., the first return of a pulse does hit within the region but its last return is outside the region boundary or vice versa.

Ancillary information in lidar data, such as intensity and scan-angle, could be of practical value for some specific applications. For example, Hopkinson and Chasmer (2007) examined an intensity-based ratio metric, i.e., canopy power versus total power, and found it an effective predictor for estimating canopy gap fractions. According to a single-scattering model for the “hot-spot” viewing geometry (Sun & Ranson, 2000), the reflected laser intensity from a volume \(dV\) around crown surfaces is formulated as

\[\Delta \alpha \propto t_0 u_0(z) R_0 dV\]

where \(t_0\) is the incident laser intensity, \(u_0(z)\) is the leaf area density defined around the crown surface at a location \(z\), and \(R_0\) is the reflectance of leaves which are assumed to be Lambertian. It is shown in Eq. (8) that the intensity of a lidar return from the top crown surface is proportional to leaf area density \(u_0(z)\). Be aware, however, that \(t_0\) may vary slightly from pulse to pulse, depending on sensor-target distances, flight altitude and scan-angle; and that the lidar intensity measurements of most commercial airborne laser systems are not calibrated, which somehow limits the practical use of lidar intensity data for quantitative applications such as estimating leaf area densities (Boyd & Hill, 2007).

The scan angle of a laser pulse, as recorded by an airborne laser system and saved in the LAS format, often differs from the local look-angle that is defined as the angle between the pulse incidence and the local normal to the ground. This disparity may be due to the instability of flying platforms and the rugged topography. The local look-angle \(\theta_i\) of a pulse with two returns can be approximately calculated from the coordinates of its FR and LR by

\[
\cos \theta_i \approx \frac{|z_{\text{lr}} - z_{\text{fr}}|}{\sqrt{(x_{\text{lr}} - x_{\text{fr}})^2 + (y_{\text{lr}} - y_{\text{fr}})^2 + (z_{\text{lr}} - z_{\text{fr}})^2}}
\]

where the subscripts \(i, \text{fr}\), and \(\text{lr}\) refer to the pulse ID, first return and last return, respectively, and \(z\) is the height relative to the ground. For a single-returned pulse, \(\theta_i\) can be approximated as that of the nearest two-returned pulse. Eq. (9) is only an approximation for relatively flat terrain.

3.3. Lidar metrics as predictors for LAI

Based on the above analysis of lidar data characteristics, we examined a variety of lidar metrics to be related with LAI using regression models (Table 1). These metrics include both laser penetration indices and canopy height metrics, some of which have been previously examined (Griffin et al., 2008; Jensen et al., 2008; Morsdorf et al., 2006). Noteworthily, most of the metrics are chosen and constructed simply on the basis of heuristic clues rather than strict analytical/physical evidence. Therefore, although our choices of lidar metrics possess certain physical interpretation, their prediction abilities should be only justified with respect to experimental performances when regressed against in-situ LAI.
where the subscript “cs” refers to returns from canopy surfaces, and all the other subscripts should be self-explicit by referring to the above explanation. The reason for considering inside-canopy hits as in \( \text{in + grd/total} \) and \( \text{in + grd/pulse} \) is that inside-canopy hits can also reflect a certain degree of penetration (e.g., 1L and 5L of Fig. 3).

In addition, intensity-based ratios, as the variants to the above hit-number-based ratio metrics, can be calculated using the sum of intensity values of lidar hits rather than the numbers of hits. These intensity-based ratios were constructed in the same spirit as that of Hopkinson and Chasmer (2007). However, to constrain the value of ratios to fall within the interval [0,1], we considered only the ratios using the sum of intensity of all hits as the denominator, not those using the intensity sum of all first returns. The resulting intensity-based counterparts to the first three hit-number-based ratios of Eq. (10) are symbolized as \( \text{Igrd/total} \), \( \text{Iin + grd/total} \), and \( \text{Isr^grd/pulse} \), respectively. As an example, \( \text{Igrd/total} \) is given by,

\[
\text{I}_{\text{grd/total}} = \sum_{i=1}^{N_{\text{grd}}} \frac{l_i}{N_{\text{total}}}
\]

where \( l_i \) is the intensity for any returns, and \( l_i \) for ground returns only.

### 3.3.2. Laser penetration metrics adjusted by look-angle

Laser penetration through a canopy also depends on incident directions, often with a larger incident angle resulting in less penetration as implicitly indicated by Eq. (1). Due to the scanning mechanism of airborne laser scanners, incident directions of laser pulses oscillate slightly from pulse to pulse within/between the maximum scan-angles of the scanners (Fig. 3). Moreover, the displacements of different field plots relative to a flightline vary, thus contributing to the variation in incident angles among different plots. Owing to the cross-hatch flight patterns of our lidar data (Fig. 1a), most field plots were observed from multiple flight lines, thus increasing the unevenness of incident angles of laser pulses within or among field plots. To compensate for such unevenness, we designed several laser penetration metrics that are adjusted by the effect of look-angle \( \theta_i \), an example of which is given by,

\[
\text{R}_{\text{grd/total}} = \sum_{i=1}^{N_{\text{grd}}} \frac{1}{N_{\text{total}}} \frac{\cos \theta_i}{N_{\text{total}}}
\]

where \( \cos \theta_i \) is calculated with Eq. (9). We adjusted the look-angle effects only for the hit-number-based LPMs of Eq. (10), not for the intensity-based LPMs. The adjusted counterparts to the remaining seven penetration metrics of Eq. (10) can be obtained by analogy to Eq. (12) (Table 1). It is emphasized again that the formula form of the adjusted LPMs is not strictly physically-based but rather heuristically-derived.

### 3.3.3. Lidar height-related metrics (HRM) and foliage-density proxies

Conceptually, LAI is proportional to canopy volume, foliage density, or their product. Lidar-based surrogates for canopy volume are height metrics, e.g., mean canopy height (Chen et al., 2007). For foliage density, one lidar surrogate is the “mean free path” of lidar pulses (\( h_{\text{free}} \)), which is defined in Kusakabe et al. (2005) as the mean penetration length of lasers into the top canopy surface and is deemed to be inversely related to the foliage density; another potential surrogate to foliage density is the mean intensity of first canopy returns (\( \text{I}_c^{\text{rcs}} \)), which is identified by referring to Eq. (8) where \( u(2) \) is the foliage density near crown tops.

Based on the above heuristics for predicting LAI, we investigated a number of height-related metrics (HRM), including mean height of single and first returns (\( h_{\text{free}}, h_{\text{rst}}, h_{\text{total}} \)), the 10%–90% percentile heights of all returns with a 10% increment (\( h_{\text{rst}}, h_{\text{total}} \),...).
and the optimal threshold does not necessarily equal the camera setup height. The “mean free path” \( h_{\text{fre}} \) is calculated based on only canopy hits with the free path of a single-returned pulse being zero and that of a double-returned pulse being the distance between its first and last returns; the average intensity of canopy hits \( I_{chm} \) is calculated based on only canopy hits of first return.

### 3.4. Statistical analysis relevant to lidar-derived metrics

Several factors make it apparently difficult to develop regression models for predicting LAI using the lidar metrics described above:

1. To derive LPM, an appropriate height threshold needs to be determined for separating ground and canopy hits (Fig. 3). Earlier studies often chose the setup height of fish-eye camera as threshold (Morsdorf et al., 2006), but such a choice guarantees no optimality.

2. Hemiphotos were analyzed to derive in-situ LAI, but no exact plot size is readily defined for the extent to which a hemiphoto can view (Fig. 4). Thus, the appropriate plot size, at which lidar metrics should be extracted, is to be determined (Morsdorf et al., 2006; Riano et al., 2004).

3. It remains unclear as to which lidar metric or subset of metrics offers the most prediction power. More often than not, the metrics used in previous studies are study-specific.

4. The exact functional form of the relationship between LAI and each lidar metric is not theoretically available, although both linear and logarithmic models have been previously used and LPMs were preferred to be used in logarithmic models by analogy with the Beer’s law (Farid et al., 2008; Griffin et al., 2008; Jensen et al., 2008; Morsdorf et al., 2006).

5. Forest types, i.e., pines or hardwoods, may well influence the relationships between LAI and lidar metrics.

The compounding effects of the above five factors as well as the limited number of in-situ LAI observations make it rather impractical to enumerate all the combinations of lidar metrics, model forms, plot sizes, and forest types for model selection. As an expedient, we employed the following strategies to select predictors and models based mostly on \( R^2 \) or root mean square error (RMSE).

#### 3.4.1. Regression models using LPMs

When using the LPMs or adjusted LPMs as predictors, the logarithmic model (log-model) with no intercept term is adopted because of its resemblance to the Beer’s law. Moreover, to accommodate possible difference between pines and hardwoods, the stratified equation was used, i.e.,

\[
L = \beta_1 \cdot \ln(x) + \beta_2 \cdot t \cdot \ln(x) + \epsilon
\]

where \( x \) can be any of the 15 LPMs described in Table 1, \( t \) stands for a binary variable distinguishing between pines (\( t = 1 \)) and hardwood/mixed (\( t = 0 \)), \( \beta \)'s are coefficients to be determined, and \( \epsilon \) is the error term.

The value of a LPM varies with the plot size and height threshold that were selected to calculate the LPM (Fig. 4). To investigate the combined effects of predictors, plot sizes and height thresholds on model performance, the model in Eq. (13) was examined for each of the 15 LPM predictors that have been extracted on the 53 field plots using a range of plot radii from 4.0 m to 85.0 m incremented by 1.0 m as well as using a range of height thresholds from 0.1 m to 10.0 m incremented by 0.01 m. Hence, we fitted a total of 123,000 models in the form of Eq. (13) (15 LPMs by 100 height thresholds by 82 plot sizes). RMSEs and \( R^2 \) values of the fitted models were referred to for selecting the most effective subset of predictors and for assessing the sensitivity of model performances to the plot size and height threshold as well as for choosing the “optimal” plot size and height threshold to derive LPMs. Note that throughout the paper, the term “optimal” is used in a relative and loose sense to indicate the most appropriate choice of plot size or threshold with which the associated LPMs give the top \( R^2 \). As shall be seen later, the “optimal” height threshold, which is chosen to separate ground and canopy hits in calculating LPM, does not necessarily equal the camera setup height (1.5 m).

#### 3.4.2. Regression models using HRMs

When using HRMs as predictors, both univariate and multiple-variable regression models were investigated. For the univariate models, a total of 41 predictors were examined, which include 13 height metrics, two foliage density proxies, and 26 product metrics (i.e., 13 height metrics, two density proxies, and 26 product metrics). A total of 41 predictors were examined, which include 13 height metrics, two density proxies, and 26 product metrics (i.e., 13 height metrics, two density proxies, and 26 product metrics). A total of 41 predictors were examined, which include 13 height metrics, two density proxies, and 26 product metrics (i.e., 13 height metrics, two density proxies, and 26 product metrics).

\[
y = \beta_0 + \beta_1 x + \beta_2 x \cdot t + \epsilon
\]

where \( x \) can be any one of the 41 predictors, \( t \) again is the binary variable to differentiate pine plots from those of hardwood/mixed, \( \beta \)'s are the regression coefficients, and \( \epsilon \) is the normal error term.

Furthermore, to accommodate potential nonlinear relationships, the power transformations of LAI and predictors were also examined with the following model form,

\[
y = L^{\lambda_1} = \beta_0 + \beta_1 x^{\lambda_2} + \beta_2 x^{\lambda_3} t + \epsilon
\]

where \( \lambda_1 \) and \( \lambda_2 \) are the power transform exponents for LAI and the predictor, respectively, and all the other symbols have the same meaning as those of Eq. (14). The two exponents \( (\lambda_1, \lambda_2) \) were estimated by way of the bi-variate box–cox transform, which is a widely used statistical procedure to transform independent and dependent variables in linear regression (Box & Cox, 1964). The use of the box–cox transform renders the regression analysis more theoretically sound. In particular, a zero exponent in the box–cox transform reduces to the logarithm transform as in Eq. (13).

In the case of multiple regression, the pooled set of the above 41 predictors, without any power transformation, was used as candidates to be selected into a final reduced regression model with the stepwise regression procedure (Neter et al., 1996). Only the predictors that satisfy certain statistical criteria will enter the final model. The model was not stratified by forest types during the variable selection, due to a small sample size relative to the number of potential predictors.

In addition, we varied the plot size from 4 m to 85 m with a 1.0-m increment in radius to extract the HRMs and the two lidar foliage
density proxies for fitting the aforementioned univariate and multiple regression models. Therefore, for either Eq. (14) or Eq. (15), a total of 3362 models (41 predictors by 82 plot sizes) were evaluated. Results of the fitted models were compared for identifying effective HRM predictors as well as the appropriate plot size. Of particular note, we made no attempt to build a model that involves both LPMs and HRMs.

3.4.3. Model comparison

Except for the multiple-variable model fitted with stepwise regression, all the LPM-based or HRM-based single-variable models (Eqs. (13) or Eqs. (14) and (15)) have the same model dimension with only one lidar metric plus a binary categorical variable. Model comparison for these single-variable models is different from stepwise regression or variable selection in prior studies that attempt to identify an optimal subset of predictors from a candidate pool (Popescu et al., 2003). For our purpose, we used RMSE or equivalently RMSE as performance criteria to choose the optimal plot size and height threshold as well as the useful lidar metrics, as also described above. \( R^2 \) was calculated according to its most general definition, i.e., the unity minus the ratio of the error sum squares to the total sum squares (Neter et al., 1998).

Given two models of the same model dimension (e.g., two LPMs of the same type that, however, are extracted using two different height thresholds), common practice is to simply choose the one with higher \( R^2 \). However, techniques to quantify the significance of the \( R^2 \) improvement of one model over another are not adequately developed, possibly due to the intractability to model probability distributions of \( R^2 \), and thus are rarely seen in remote sensing literature. In this study, we adopted the Nash–Sutcliffe \( r^2 \) criterion (N-S \( r^2 \)), which was originally proposed to compare performance improvement of hydrological models (Nash & Sutcliffe, 1970), to assess the significance of increase in \( R^2 \). The equation for it is

\[
\begin{align*}
\Delta R^2 = & \frac{R^2_2 - R^2_1}{1 - R^2_1} = \frac{\text{PRESS}_1 - \text{PRESS}_2}{\text{PRESS}_1} \\
\end{align*}
\]

where \( R^2_1 \) and \( R^2_2 \) are \( R^2 \)s of Model 1 and Model 2. This criterion represents the proportion of initial variance unexplained by Model 1 but subsequently explained by Model 2. One advantage of \( r^2 \) over \( \Delta R^2 \) is that the same amount of \( R^2 \) difference \( \Delta R^2 \) is more weighted at a larger initial \( R^2 \). The critical \( r^2 \) value of significance is chosen subjectively or experientially. Following up other studies (e.g., Senbeta et al., 1999), we used 10% in this study, i.e., Model 2 is judged to be significantly or experientially better than Model 1 only if \( r^2 \) is greater than 0.10. In addition, we computed the predicted residual sums of squares (PRESS statistic) via the leave-one-out cross-validation (LOOCV) to assess model validity and generalization (Neter et al., 1996), as has also been used in Andersen et al., 2005. A close agreement in magnitude between RMSE and (PRESS)\(^{1/2}\) suggests that the fitted model tends to have less overfitting and more generalization.

3.5. Correlation of NDVI with in-situ LAI

NDVI on the field plots was derived from the Quickbird image. In deriving the plot-level NDVI, a scale discrepancy exists between the field plots and image pixels with a field plot covering a number of Quickbird pixels. Thus, two groups of NDVI at the plot level were calculated, depending on the order by which the aggregation of pixels and the calculation of NDVI are performed. The first group, NDVI1, was calculated by first computing NDVI for each Quickbird pixel and then aggregating the pixelwise NDVI to the plot level, i.e.,

\[
\text{NDVI} = \frac{\sum_{i=1}^{n} \text{NDVI}_i}{n} = \frac{\sum_{i=1}^{n} (\text{NIR}_i - \text{RED}_i)}{\sum_{i=1}^{n} (\text{NIR}_i + \text{RED}_i)}
\]

where NDVI is the NDVI value for a given image pixel and is computed from NIR and RED, i.e., radiances at the near-infrared and red bands, and \( n \) is the total number of pixels within the plot. In contrast, the second group of plot-level NDVI was obtained by first aggregating pixel-level spectral radiances up to the plot level and then using the resulting plot-level radiances to calculating NDVI, i.e.,

\[
\text{NDVI}^2 = \frac{\sum_{i=1}^{n} \text{NIR}_i - \sum_{i=1}^{n} \text{RED}_i}{\sum_{i=1}^{n} \text{NIR}_i + \sum_{i=1}^{n} \text{RED}_i}
\]

where the relevant symbols have the same meaning as in Eq. (17). Due largely to landscape heterogeneities at sub-plot levels, NDVI1 and NDVI2 are not equal, with NDVI2 being often larger. Such a discrepancy is considered as a scaling effect when estimating biophysical variables from NDVI (Zhao et al., 2009). The plot size used for computing the NDVI is the identified “optimal” radius that provides the best model performance for LPM in Section 3.4.1.

Correlation analysis was then performed to determine whether NDVI, when combined with the lidar metrics, provides additional predictive power for estimating LAI. We explored the strength of correlations between the plot-level NDVI and in-situ LAI as well as between the NDVI and the residual LAI of the lidar-based LAI models. Specifically, a significant correlation between the NDVI and the residuals of the predicted LAI indicates that a portion of the variance in LAI explained by lidar metrics can be explained by NDVI and thus, the integration of NDVI with lidar metrics improves on the LAI estimates as compared to the use of lidar metrics or NDVI alone. In the correlation analysis, we discarded the 15 hardwood/mixed plots and only considered the 39 pine plots mainly for two reasons. First, NDVI-LAI relations are species-dependent, and the number of hardwood/mixed plots is relatively small so as not to give robust inference. Second, the in-situ LAI was measured in the leaf-on season while the Quickbird image was acquired in the leaf-off season, and the use of only pine plots tends to minimize the discrepancy in seasonality because for Loblolly pines, hemiphotos appeared insensitive to changes in LAI within a season and the changes in LAI between leaf-on and leaf-off seasons are small and well correlated (Dewey et al., 2006).

3.6. Lidar-derived LAI map for comparing with GLOBCARBON LAI product

3.6.1. Generating a lidar LAI map

An appropriate plot radius was identified by examining the overall performances of all the aforementioned LAI models (\( R^2 \) and RMSE) fitted across a range of plot sizes (i.e., 4 m–85 m) and height thresholds. This identified radius was used to determine the spatial resolution (scale) at which to generate the lidar LAI map so that the pixel size of the map should approximately equal the chosen plot area. However, even at this identified plot size, thousands of LAI models were fitted, so only the one with the best performance (i.e., the highest \( R^2 \)) was selected to create the LAI map.

3.6.2. Registration between lidar and GLOBCARBON LAI maps

Co-registration between the lidar and GLOBCARBON LAI maps constitutes a practical difficulty for making a pixelwise comparison. The difficulty arises from the fact that the two maps, both being of raster format, are not only geo-referenced in different coordinate systems but also generated at different resolutions with the GLOBCARBON LAI map having a 1 km\(^2\) pixel size that is much larger than that of the lidar map. Therefore, the registration of the GLOBCARBON map in reference to the lidar one may be problematic because this process may involve interpolating 1-km\(^2\) pixels to compute new values at a re-organized grid of 1 km\(^2\) pixels. In addition, even if the re-projecting and interpolation processes contribute no extra mis-registration, a systematic pixel-shift error may have already occurred in the GLOBCARBON map because the generation of satellite products often involves resampling.
the raw observations to a pre-defined grid that slightly offsets from the actual localities observed. This shifting error, though typically less than one satellite pixel in magnitude, can add an artificial discrepancy to the comparison result.

To estimate the systematic sub-pixel shifting mis-registration between the GLOBCARBON and lidar maps, we employed a subgrid-searching procedure to match the two LAI maps in terms of maximizing their correlation. This makes our study distinct from others that directly used commercial software for co-registration without explicitly considering mis-registration. Specifically, we assumed the shift vector \( \Delta^* = (\Delta x_{\text{shift}}, \Delta y_{\text{shift}}) \) to be within the range of \([-1, 1] \times [-1, 1]\) (i.e., a square) where the length unit is a GLOBCARBON pixel size. Then, we divided this square region of error into 200 by 200 subgrids, which means that \( \Delta x_{\text{shift}} \) or \( \Delta y_{\text{shift}} \) can take values of \(-1, -99/100, -98/100, ..., 99/100, 1\). Next, for each of these 40,000 potential shift errors, we co-registered the two maps accordingly, aggregated the lidar map up to the 1-km resolution, and computed the corresponding correlation between the satellite and aggregated lidar maps. An estimate of the shift error was finally determined as the vector that gave the highest correlation. Afterwards, we compared the two LAI maps by taking into account the estimated shift error.

3.6.3. Comparison of LAI between two maps

Even after being accurately co-registered, the GLOBCARBON and lidar LAI are not directly comparable because the former is true LAI while the latter is effective LAI. To be more precise, the predicted LAI from lidar refers to effective LAI because the in-situ LAI derived from hemiphotos for calibrating regression models is effective LAI, i.e., the notation \( L \) in all the aforementioned equations denotes effective LAI. Theoretically, effective LAI \( L_t \) can be converted to true LAI \( L_i \) with the following formula, as partially revealed in Fig. 2a,

\[
L_i = (1 - \alpha) \frac{\gamma_e}{\Omega_c} L = (1 - \alpha) \frac{1}{L_t}
\]

where \( \alpha \) is the wood-to-total area ratio, \( \gamma_e \) is the needle-to-shoot area ratio, which is 1.0 for broadleaved forests, and \( \Omega_c \) is the clumping index that accounts for non-random patterns at scales larger than shoots (Chen & Black, 1992). However, we lacked ancillary data to infer values of these correction factors for our study area. Chen et al. (2005) derived global land cover-specific \( \Omega_c \) and the values are 0.69 and 0.62 for closed broadleaf and evergreen forests, respectively. On the other hand, even if the effective LAI is not corrected, a strong linear relationship should exist between the true and effective LAI for a certain forest type. Furthermore, the slope of the zero-intercept line fitted to the data of true versus effective LAI should give an estimate of the overall correction factor. As such, we fit both non-zero and zero intercepted lines to the data points of GLOBCARBON versus lidar-derived LAI to test if there exists a strong linear relationship between them and to determine if the slope of the fitted line can reveal certain reasonable information about the correction factor.

4. Results

4.1. Uncertainties in in-situ LAI estimates

Effects of analysts and hemiphot analysis algorithms on the in-situ LAI estimates were examined. Fig. 5 presents the results for the comparison of in-situ LAI values obtained by two analysts as well as the comparison between the two algorithms. To avoid a plethora of figures, the scatterplots in Fig. 5 refer to the mean LAI values averaged over the two levels of a factor when making comparison between the two levels of another factor. It is observed that both the factors (i.e., analysts and algorithms) affect the estimation of in-situ LAI. In terms of the strength of correlation between the two levels of each factor, the effect of algorithms seems more distinct than that of analysts as revealed by the scattering patterns of two factors (Fig. 5). According to the repeated measures ANOVA test that used analysts and algorithms as within-subject factors, and forest types (i.e., pines versus hardwood/mixed) as between-subject factor, neither analysts nor algorithms have significant influences on estimating mean in-situ LAI values at a level of 0.05; however, the test did show that for both forest types, the factor of algorithms (p-value: 0.09) appears to affect the LAI estimation more than the factor of analysts (p-value: 0.36). In addition, the ARMSD for analysts and algorithms are 0.32 and 0.37 respectively, which also reveals that algorithms have more influences on the uncertainties in the in-situ LAI estimates, similarly to what the repeated measures ANOVA suggests.

4.2. Lidar-based LAI estimates using LPMs

4.2.1. Effects of plot size and thresholding height

The plot sizes used to extract LPMs affect model performances significantly. Results of the 123,000 fitted LPM models show that for all 15 predictors, the range of 12−35 m was generally appropriate choices of plot radius for extracting LPMs to be related with in-situ LAI. Over this range of plot size, only slight fluctuations in \( R^2 \) were observed, as shown in Fig. 6a where the changes in \( R^2 \) with plot size are plotted for the top four LPMs that were derived using a 3.6 m height threshold in separating ground and canopy hits. It is also
observed that when using too small a plot size (e.g., less than 8 m in radius), the fitted models degrade greatly with a relatively large decrease in $R^2$ (e.g., from 0.78 at a radius of 8.0 m to 0.61 at a radius of 6.0 m for the predictor $I_{\text{g}+\text{r} \text{d}/\text{total}}$), and when gradually increasing the plot size above about 25.0 m, the fitted models show a slightly declining performance. Based on these observations, we chose a plot size of 25.0 m as the “optimal” scale to build and apply the lidar-based LAI models.

The height threshold, below which lidar returns are identified as ground returns when deriving LPM, also greatly affects model performances (Fig. 6b). Results of the 123,000 fitted LPM-based models indicate that a height threshold within 2.0 m to 5.0 m is an appropriate choice for determining ground hits. Furthermore, in terms of $R^2$ values of all the models fitted across various plot sizes, a height of about 3.6 m was selected as the “optimal” threshold, which is 2.1 m higher than the camera set-up height (1.5 m).

4.2.2. Comparison of LPMs

Effectiveness of LPMs for estimating LAI varies significantly from one metric to another. Fig. 7 shows an example comparing the $R^2$ of the 15 LPMs extracted at the identified optimal setting (i.e., a plot radius of 25.0 m and a height threshold of 3.6 m). Across all plot sizes and height thresholds investigated, some look-angle-adjusted LPMs occasionally produced slightly higher $R^2$’s and lower RMSEs than their non-adjusted counterparts, but a paired t-test suggests that such improvement is not statistically significant ($p$-value: 0.73). Results indicate that intensity-based LPMs performed much worse than their hit number-based counterparts. For example, a mean N-S $R^2$ of 0.62 was observed for comparing $I_{\text{g}+\text{r} \text{d}/\text{total}}$ and $I_{\text{g}+\text{r} \text{d}/\text{total}}$ which implies that $R^2$ will be greatly increased by changing the predictor from $I_{\text{g}+\text{r} \text{d}/\text{total}}$ to $I_{\text{g}+\text{r} \text{d}/\text{total}}$. Results also indicate that in most cases, LPMs derived by using the total number of laser hits as denominator are superior to those using the total number of laser pulses as denominator; for instance, the average N-S $R^2$ is 0.52 when comparing $I_{\text{g}+\text{r} \text{d}/\text{pulse}}$ and $I_{\text{g}+\text{r} \text{d}/\text{total}}$, meaning that 52% of variance unexplained by $I_{\text{g}+\text{r} \text{d}/\text{pulse}}$ is explained away by $I_{\text{g}+\text{r} \text{d}/\text{total}}$.

Among the 15 LPMs, the top four are $I_{\text{f} \text{n}+\text{r} \text{d}/\text{total}}$, $I_{\text{f} \text{n}+\text{r} \text{d}/\text{total}}$, $I_{\text{g}+\text{r} \text{d}/\text{total}}$, and $I_{\text{g}+\text{r} \text{d}/\text{total}}$, the respective $R^2$’s of which, for example, are 0.841, 0.821, 0.805 and 0.795 at a plot size of 25 m and height threshold of 3.6 m (i.e., the identified optimal setting). These four LPMs perform consistently well across a range of plot sizes and height thresholds and they are far more effective than the remaining 11 predictors when used in the log-model of Eq. (13). The prediction abilities are very close between $I_{\text{f} \text{n}+\text{r} \text{d}/\text{total}}$ and $I_{\text{f} \text{n}+\text{r} \text{d}/\text{total}}$ or between $I_{\text{g}+\text{r} \text{d}/\text{total}}$ and $I_{\text{g}+\text{r} \text{d}/\text{total}}$ with the former metric in each pair slightly better than the latter one, which again shows that our look-angle-adjusted metrics did not necessarily improve performances over those that do not account for look-angle effects. Moreover, averaged over the identified appropriate ranges of plot size and height threshold, the mean N-S $R^2$ values are 0.16 for $I_{\text{g}+\text{r} \text{d}/\text{total}}$ versus $I_{\text{f} \text{n}+\text{r} \text{d}/\text{total}}$ and 0.14 for $I_{\text{g}+\text{r} \text{d}/\text{total}}$ versus $I_{\text{g}+\text{r} \text{d}/\text{total}}$. Both $R^2$ values are greater than the specified critical value of significance (i.e., 0.10), which suggests that the use of inside canopy plus ground hits as denominator is more effective than the use of only ground hits when deriving $I_{\text{g}+\text{r} \text{d}/\text{total}}$ or $I_{\text{g}+\text{r} \text{d}/\text{total}}$ to predict LAI.

4.3. Lidar-based LAI estimates using HRMs

4.3.1. Single-variable regression model

As compared to LPMs, HRMs prove less effective in predicting LAI. The obtained $R^2$ values ranged from 0.007 to 0.712 for the untransformed univariate linear models using each of the 41 HRMs as predictor that have been extracted across various plot sizes ranging from 4.0 m to 85 m in radius. Of the 41 HRM predictors, the 13 height metrics and two foliage density proxies appear to be weak predictors when used alone in Eq. (14); the 13 product metrics resulting from the multiplication of 13 height metrics with $I_{\text{g}+\text{r} \text{d}/\text{cs}}$ are also poor predictors; however, the 13 product metrics resulting from the multiplication of 13 height metrics with $I_{\text{g}+\text{r} \text{d}/\text{cs}}$ have significantly improved the results over the height metrics (e.g., a mean N-S $R^2$ of 0.67 for comparing $I_{\text{g}+\text{r} \text{d}/\text{cs}}$ and $I_{\text{g}+\text{r} \text{d}/\text{cs}}$ versus the range of plot sizes investigated). Moreover, the plot size influences model performance (Fig. 8), and a range of plot
sizes from 15 m to 70 m are identified as appropriate scales for extracting the lidar HRM metrics and applying the model. Among the 41 predictors for the univariate model, the product metric $h_{\text{chm}}^\lambda (1/h_{\text{free}})$ consistently yields the largest $R^2$ across the identified range of plot sizes, followed next by a similar product metric $h_{\text{sr}} + t/1/h_{\text{free}}$ due to the high correlation between $h_{\text{sr}} + t$ and $h_{\text{chm}}$. This result lends support to the original presumption that LAI is proportional to both canopy volume and foliage density and the use of their product can improve prediction.

On the other hand, the transformed single-variable models of Eq. (15) have a range of $R^2$ values from 0.071 to 0.751 with significant or marginal improvements over their untransformed counterparts of Eq. (14), as partially indicated by the comparison of $[h_{\text{chm}}(1/h_{\text{free}})]^\lambda$ and $h_{\text{chm}}(1/h_{\text{free}})$ in Fig. 8. For the plot radius range of 15–70 m, the estimated exponent in the box–cox transform for LAI (i.e., $\lambda_1$) is within close proximity to 1.0, and those for the 41 HRM predictors (i.e., $\lambda_2$) fall within the interval [0.268, 0.922]. Thus, we fixed the value of $\lambda_1$ in Eq. (15) to 1.0 for all the transformed single-variable models evaluated. The deviation of $\lambda_2$ from unity indicates the potential nonlinearity between LAI and the HRM metrics. Moreover, $\lambda_2$ of HRM predictors usually increases from a small value toward 1.0 as the plot size increases, which suggests that the model at larger plot sizes has more tendency to be a linear one. The plot size affects the performances of the transformed models in a similar way as that of the untransformed models, with an appropriate plot size range roughly from 15 m to 70 m. Such an effect of plot size is due to the scale dependence of both the extracted HRM metrics and the estimated $\lambda_2$.

### 4.3.2. Multiple-variable model from stepwise regression

Interestingly, the stepwise regression yielded the following multiple-variable model with two predictors $h_{\text{sr}} + t/1/h_{\text{free}}$ and $h_{\text{total}}/1/h_{\text{free}}$.

$$
\hat{L} = 0.41 + 2.80 \left(\frac{h_{\text{sr}} + t}{h_{\text{free}}} - 2.56 \frac{h_{\text{total}}}{h_{\text{free}}}\right) \\
= 0.41 + 2.80 \left(\frac{h_{\text{sr}} + t}{h_{\text{free}}} - 0.91 \frac{h_{\text{total}}}{h_{\text{free}}}\right)
$$

(20)

which has an $R^2$ of 0.80 (RMSE=0.32). To interpret the model meaningfully, we combined its two predictors into a single product $[(h_{\text{sr}} + t - 0.91h_{\text{total}})/(1/h_{\text{free}})]$, which is non-negative because $h_{\text{sr}} + t$ is always greater than or equal to $h_{\text{total}}$. As mentioned above, $1/h_{\text{free}}$ is treated as a proxy to foliage density (Kusakabe et al., 2005). Therefore, under our original presumption that LAI is proportional to the product of canopy volume and foliage density, $(h_{\text{sr}} + t - 0.91h_{\text{total}})$ can be considered as a proxy to canopy volume or, more precisely, effective crown volume containing foliage.

### 4.4. Cross-validation of the optimal LAI models

Based on the results in Section 4.2 and 4.3, a radius of 25 m was finally chosen as the most favorable plot size for relating the lidar metrics to in-situ LAI. Table 2 summarizes the best fitted LPM models at this plot size for different types of lidar metrics. As mentioned above, the LPM metrics, especially $r_{\text{in}} + grd/total$, are more effective than the HRM metrics; the product metric $h_{\text{chm}}(1/h_{\text{free}})$ outperforms the other HRM predictors; and the box–cox transform of the HRM predictors tends to improve the model fitting. Furthermore, it is found that the species variable $t$ is insignificant in the three stratified models of Table 2 (e.g., a $p$-value of 0.32 for the model with $r_{\text{in}} + grd/total$), which indicates that the stratification of pines versus hardwood trees did not significantly improve the accuracy in estimating effective LAI.

LOOCV was used to compute PRESS for the top model of each category (Table 2). Square roots of the obtained PRESS’s are comparable to the corresponding RMSEs for all these top models (Table 2); for example, the RMSE and squared root of PRESS for the log-model of $r_{\text{in}} + grd/total$ are 0.29 and 0.30, respectively. Such consistency in magnitude provides positive evidence that these models have good generalization abilities with no or little overfitting.

### 4.5. Correlation analysis relevant to NDVI

To be consistent with the development of lidar LPM models, NDVI1 and NDVI2 were calculated at the identified “optimal” plot size of 25 m in radius. NDVI1 was significantly different from NDVI2 ($p$-value $<0.000$ from a paired $t$-test). Both of them were found to be significantly correlated with in-situ LAI with a correlation coefficient of 0.763 and 0.788 for NDVI1 and NDVI2, respectively, but no significant correlation exists between NDVI and the LAI residuals of Table 2 (e.g., a $p$-value of 0.32 for the model with $h_{\text{total}}/1/h_{\text{free}}$), which indicates that the stratification of pines versus hardwood trees did not significantly improve the accuracy in estimating effective LAI.


table 2

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Model</th>
<th>$\hat{R}^2$</th>
<th>rmse</th>
<th>PRESS $(\hat{R}^2)^{1/2}$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPM</td>
<td>$L = -2.56\ln (r_{\text{in}} + grd/total) - 0.12\ln (r_{\text{in}} + grd/total)$</td>
<td>0.84</td>
<td>0.29</td>
<td>0.30</td>
<td>0.42</td>
</tr>
<tr>
<td>HRM</td>
<td>$L = 0.56 + 0.82 [h_{\text{chm}}(1/h_{\text{free}})] + 0.074r_{\text{in}}$</td>
<td>0.72</td>
<td>0.37</td>
<td>0.38</td>
<td>0.84</td>
</tr>
<tr>
<td>HRM</td>
<td>$L = 0.001 + 1.40[h_{\text{chm}}(1/h_{\text{free}})]^{0.67}$</td>
<td>0.76</td>
<td>0.34</td>
<td>0.36</td>
<td>0.54</td>
</tr>
<tr>
<td>HRM</td>
<td>$L = 0.41 + 2.80(h_{\text{sr}} + t)/h_{\text{free}}$</td>
<td>0.80</td>
<td>0.32</td>
<td>0.35</td>
<td>-</td>
</tr>
</tbody>
</table>

The test results show that stratification leads to no significant improvement.
metrics in linear models may not significantly improve the results as compared to the use of lidar metrics alone.

4.6. Comparison of lidar and GLOBCARBON LAI maps

4.6.1. A fine-resolution lidar-derived LAI map

The stratified log-model with $\frac{r_{in}+grd}{total}$ in Table 2 is the best fitted model (Fig. 9a) and thus was applied to the whole study area to generate a spatially-explicit LAI map (Fig. 9b). Although the stratification tends not to improve accuracy (Table 2), we still used this stratified model because the species information was readily available from the classified Quickbird image. The map has a spatial resolution of $45 \times 45$ m, as determined by equaling the area of a pixel to that of a circular plot of $25$ m in radius.

4.6.2. Co-registration and comparison of two LAI maps

The systematic sub-pixel shift error of the GLOBCARBON map relative to the lidar map was estimated to be $(\Delta x_{shift}, \Delta y_{shift}) = (-0.58, -0.62)$ km (GLOBCARBON pixel). The correction for this shift error greatly improved correlation between the two maps, e.g., from 0.08 to 0.85 for pines, as also revealed in Fig. 10 where correlation is reported without differentiating pines and hardwoods. Such a result

![Fig. 9. (a) The best fitted logarithm model (LAI vs. $\ln(\frac{r_{in}+grd}{total})$). (b) A spatially-explicit effective LAI map of 45 m resolution was generated by applying the model of (a) to the whole lidar dataset.](image-url)
implies that co-registration of the two maps by directly using commercial software can strongly bias comparison results (Fig. 10a).

After accounting for the shift error of (−0.58, −0.62) km, a total of 54 GLOBCARBON pixels were first selected that have at least 80% of their pixel areas falling into the boundary of the study area with 41 pixels of them falling completely within the study area. In a strict sense, none of the 54 pixels have pure land cover, due to the 1 km² grain size. By further referring to the classified Quickbird high-resolution image, we were able to label a pixel as pine or hardwood based on whether the pixel has at least 80% of pixel area covered by pines or hardwoods. As a result, of the 54 pixels, 24 were labeled as pines and 5 as hardwoods, and the remaining 25 pixels were discarded because they do not have coverage of pines or hardwoods above 80%.

The GLOBCARON LAI was well correlated with the aggregated lidar LAI with a correlation coefficient of 0.849 (pines, n = 24) and 0.854 (hardwoods, n = 5). Therefore, straight lines with non-zero intercepts can well fit the relationships between the GLOBCARBON and lidar LAI (Fig. 11). But, it is proven unsuccessful to fit the relationships using zero-intercept lines, resulting in negative $R^2$ values both for pines and hardwoods (Fig. 11). By referring to the slopes of the two fitted zero-intercepted lines (Fig. 11) and using the global average $\Omega$ values of 0.62 and 0.69 respectively for pines and hardwoods (Chen et al., 2005), we estimated the wood-to-total area ratios of pines and hardwoods to be 0.74 and −0.10, respectively, according to the relations in Eq. (19). Apparently, these two estimates are far from realistic as compared to those previously reported for similar land cover classes (Chen et al., 1997). We also computed the corrected LAI $L_c$ from the lidar-derived effective LAI $L_e$ using $L_c = L_e / \Omega$, assuming that $\Omega$ for pines is 0.69 (Chen et al., 2005). Then, the fitted nonzero-intercept equation reported in Fig. 11a for pines can be re-written as

$$ L_{\text{globearbon}} = 1.77 + 0.56 \cdot L_c \quad (21) $$

where the intercept differs from zero and the slope from the unity significantly (both p-values << 0.000). This suggests that direct comparison can lead to a large absolute difference in the two types of LAI. We did not report the counterpart equation for hardwoods due to the scarcity of hardwood pixels, i.e., only 5 pixels.

Fig. 10. (a) A subset of the GLOBCARBON LAI map (1-km resolution) overlaid with the boundary of lidar-derived map (parallelogram): The solid line parallelogram refers to the co-registration by directly using ArcGIS (ERSI, INC.), and the dashed-line parallelogram refers to that from our grid searching procedure. (b) Contour plot showing the variation in the strength of correlation between the GLOBCARBON and lidar-derived LAI maps as a 2-D function of the relative shift of the two maps. The vector of shift error, i.e., the displacement between the solid-line and dashed-line parallelograms in (a), is estimated to be (−0.58, −0.62) that corresponds to the highest correlation. This result implies that the mis-registration error, even less than one pixel, could induce a relatively large artificial discrepancy in the comparison.

Fig. 11. Comparison of LAI between the GLOBCARBON and the aggregated lidar-derived maps at the 1-km resolution for (a) 24 pine and (b) 5 hardwood pixels, respectively. The two maps have been registered by accounting for the shift error identified in Fig. 10. Two types of lines were fitted to the scatterplots, one with non-zero intercepts (solid lines) and another with zero intercepts (dotted lines).
5. Discussion

5.1. Analysis of lidar LAI models

This study substantiates similar findings obtained in earlier studies by showing that lidar is a viable tool for mapping LAI (Morsdorf et al., 2006; Riano et al., 2004; Solberg et al., 2006). Moreover, we extended previous studies by investigating the combined effects of plot sizes, height thresholds, and types of lidar returns on the usefulness of lidar penetration metrics as predictors in regression models for estimating LAI. Among all the LPMs examined, the hit-number-based LPMs have stronger relationships with LAI than intensity-based LPMs. The LPMs adjusted by look-angles (e.g., $R_{grd/total}$) do not improve over the counterpart metrics that do not account for the look-angle effect (e.g., $R_{lastate}$). However, this should not preclude the potential utility of scan-angle information but instead only suggests that our formulation in Eq. (12) is not a sensible way to encode scan-angle into LPM for estimating LAI. Both plot size and height threshold influence the performances of LPMs to a significant degree. In particular, the appropriate plot size at which to extract lidar metrics needs to match the effective area over which the in-situ LAI was observed; the “optimal” threshold we identified for separating ground and canopy hits is about 2.0 m higher than the camera height. In this regard, previous studies that directly used camera height in computing laser penetration/ratio metrics might have not attained the full potential of their models (e.g., Morsdorf et al., 2006; Riano et al., 2004).

The logarithm LAI models with LPMs are inherently semi-empirical models, although by analogy to the Beer’s law they have certain physical interpretation. One should refrain from calling them physically-based models because the actual physical processes underlying the lidar scanning and the LPM derivation are fundamentally different from those of the Beer’s law. The penetration rates expressed by LPMs are not the real gap fraction observed in the field, partly due to the use of laser beams much wider than individual leaves in airborne laser scanning. In the numerators of our LPMs, ground echoes were treated as penetrating hits equally without distinguishing between single and last returns (e.g., 3S versus 4L in Fig. 3), and this may overestimate the true penetration rate because a laser pulse with the last echo being a ground hit (e.g., Beam 4 in Fig. 3) does not represent 100% penetration due to the blockage of energy at its first echo and between its first and last echoes. On the other hand, the use of the total number of hits rather than the number of pluses as denominator in LPMs can lessen such over-estimation or even possibly lead to an underestimate of the true penetration. Nevertheless, for practical applications, we suggest examining various types of laser echoes to derive LPMs and determining the most effective one via trial-and-error, as demonstrated by our result that $\hat{r}_{lastate}$ was chosen to be the best metric that considers both inside-canopy and ground hits as penetrating echoes in its numerator. The lidar data of this study contain only first/last returns. Intermediate lidar returns should provide extra information on vegetative components along the penetrating paths between first and last returns, and therefore, we expect that the availability of intermediate returns or full waveforms should augment the usefulness of lidar for estimating LAI.

We examined the effectiveness of lidar LPM metrics only in the log-model with no intercept. Although linear regression models were also used with some laser penetration indices (Kwak et al., 2007), this study did not investigate these linear models due in part to some of their artifacts: LPM often takes values between 0 and 1; therefore, the direct use of LPM in a linear model sets a limit on the dynamic range of LAI values that can be predicted, thus causing artificial saturation in LAI estimates. Also, the intercept term of these linear models, if negative, may yield unrealistic estimates in certain cases; the same problem may also occur to the log-models of this study when an intercept term is added to them. On the other hand, if we do introduce an intercept term to the log-models, the resultant models will have higher $R^2$’s and lower RMSEs than their zero-intercept counterparts, especially for those LPM predictors that performed poorly in zero-intercept log-models (e.g., at the plot radius of 15 m, the $R^2$ for the log-model of $h_{lastate}$ increases from 0.1 to 0.65 by adding the intercept term). Such $R^2$ improvement for log-models with intercepts also indicates that the mean biases in prediction of our zero-intercept models are nonzero (e.g., the LAI estimates from the best fitted model using $h_{lastate}$ has a mean bias of $-0.016$, although it is negligible compared to the RMSE of 0.29). In practice, it is relegated to modelers to determine whether to use an intercept term or not, as deemed the most appropriate by them.

Only the common regression analysis was used for model inference in this study. It is recommended that more robust regression techniques should be investigated for future studies. One potential improvement is to explicitly consider the error structure of lidar-based predictors that has been ignored in the ordinary least-squares regression (Jensen et al., 2008). The variance of a hit number-based LPM can be approximated by using the binomial distribution, e.g., $\mbox{var}(\hat{r}_{lastate}) = \hat{r}_{lastate} \cdot (1 - \hat{r}_{lastate}) / N_{total}$. The incorporation of such variance information cannot boost $R^2$ but can make inference more robust, e.g., with more realistic predictive error bars and less risky extrapolation. Meanwhile, existing studies seldom exploited the utility of advanced learning algorithms such as neural networks, support vector machine and Gaussian processes for lidar remote sensing of canopy attributes. These advanced algorithms often far outperform conventional methods in many supervised learning tasks, as demonstrated in Zhao et al. (2003, 2008b). Therefore, future studies should also test the ability of some popular advanced machine learning techniques to implicitly model the possible nonlinear relationships between lidar metrics and LAI or other forest structural attributes of interest.

Although HRMs such as lidar mean and percentile heights are reportedly useful to predict forest structural attributes such as biomass, timber volume, and basal area (Holmgren et al., 2003; Lefsky et al., 1999; Nelson et al., 2003) or sometimes leaf area (Jensen et al., 2008; Roberts et al., 2005), we found that these metrics were not effective predictors for LAI. However, the products of height metrics with foliage-density lidar surrogates tend to improve the prediction power for LAI over the use of HRM alone. The multiple regression model (Eq. (20)) resulting from the stepwise analysis also seems useful for estimating LAI, and the variables selected into this model partially conform to our original heuristics that LAI is proportional to foliage-density as argued in Kusakabe et al. (2005); as such, $(\hat{r}_{lastate} - 0.91h_{total})$ can be effectively deemed as a canopy volume surrogate.

5.2. Error sources and implications for future studies

Many a source of errors contributes to uncertainties in mapping LAI with lidar. First of all, the in-situ reference LAI used for model calibration is subject to various uncertainties (Dewey et al., 2006). Both analysts and algorithms, for example, influence the retrieval of in-situ LAI from hemispherical photos. Our results indicate that the uncertainties caused by these factors were in the same order of magnitude as or even larger than the RMSE of the fitted regression models (e.g., the averaged RMSD of 0.35 for in-situ LAI versus the RMSE of 0.29 for the LPM model in Table 2). The discrepancies in in-situ LAI caused by analysts are due to the subjectivity in segmenting hemisphots (Hanssen & Solberg, 2007), and those caused by methods result mainly from the different assumptions in the in-situ inversion algorithms. Algorithm 1 has a more restrictive assumption because it specifically in the exact leaf angle distribution, which becomes invalid for certain real forest conditions. A notable difficulty in developing in-situ LAI inversion algorithm concerns the modeling of foliage clustering at several structural levels. Continuous modeling efforts have been
devoted to solving the difficulty. Nilson (1999), for example, proposed formulae for inverting canopy structural variables such as LAI from gap data by explicitly considering the clustering of foliage into separate crowns. In practice, it seems difficult to assess the true accuracy of hemisphoto-derived in-situ LAI because reference LAI measurements of higher accuracy are expensive to collect and often remain unavailable. On the other hand, emerging ground-based laser scanners hold promise to be an unprecedented non-destructive technique for accurately measuring ground-reference LAI (Hopkinson et al., 2004). Ground and airborne lasers complement each other in terms of accuracy and coverage, and the combined use of both techniques has the potential of revolutionizing traditional forest inventory and currently remains an active research area.

Furthermore, the varying choices of lidar models represent another major source of uncertainty. Model selection involves choosing both predictors and model forms. Prior research has examined a wide range of lidar-derived metrics for estimating forest stand characteristics with various regression techniques (Næsset et al., 2005); however, for estimating a given forest structural attribute, researchers often developed study-specific plot-/stand-level models that, more often than not, differ from each other in terms of the selected “optimal” laser metrics and model forms. Nelson (2008) recently emphasized that the use of various models with the same lidar data could introduce differences in the predicted biomass that are of practical significance. Results of this study also show that the best models selected with different forms or metrics and fitted to the same training data could result in a range of $R^2$ from 0.27 to 0.84. Common practice in earlier studies for estimating canopy variables with lidar is to rely on variable or model selection techniques to identify optimal predictors from a set of potential predictors (Næsset et al., 2005; Patenaude et al., 2004). These techniques tend to seek a parsimonious model while maintaining a reasonable fitting accuracy based on such statistical criteria as the Akaike and Bayesian information criteria (Chen et al., 2006; Jensen et al., 2008), but models fitted so may lack proper physical interpretations. In this study, we made no attempt to examine the combined use of LPMs and HRMs. Practical applications in the future may consider simultaneously using both types of metrics to glean extra prediction power, and the validation of such a fitted model with an independent dataset is highly recommended in order to minimize overfitting due to the addition of extra predictors.

Scale or spatial resolution (e.g., plot size used to extract LPM or HRM) also plays a critical role in developing lidar-based LAI models (Zhao et al., 2009). The determination of an appropriate scale constitutes an extra source of uncertainty. The reason for choosing an appropriate plot size in this study is mainly because no direct information is available on how far a fish-eye camera can “see” when taking hemisphotos. Trial-and-error could be a basic method to determine an “optimal” plot size (Riano et al., 2004). Previous studies also provide experiential evidence on how to identify a reasonable plot size (Morsdorf et al., 2006). For our study area, we found that the appropriate ranges of plot radii for LPM and HRM are about 12–35 m and 15–70 m, respectively, which suggests that LPMs are more sensitive to variations in plot sizes than HRMs in estimation of LAI. On occasion, even a radius up to 80 m can be used without too much significant performance degradation for some LPM and HRM predictors. Such an upper limit of radius is larger than that reported in previous studies (e.g., Riano et al., 2004) probably because our study area, comprising a large portion of pine plantations, is relatively homogeneous. The “optimal” plot size that consistently yielded the highest $R^2$ for most of the lidar metrics was around 25 m in radius, which differs from the actual field plot size (11 m in radius) that we used for measuring other stand attributes (Popescu & Zhao, 2008). Of particular note is that according to the Beer’s law, a camera will “see” through a shorter distance under thicker canopies of high LAI, and vice versa. For example, a hemisphoto taken under a relatively open and sparse canopy may observe horizontally for a few hundreds of meters. Therefore, a fixed plot size for deriving lidar metrics as used in this study may be insufficient for model calibration, and future studies may investigate the use of a varying plot size to extract lidar metrics when using hemisphoto-derived LAI as reference data to develop LAI models. One possible way is to rely on the observed in-situ LAI values to determine plot sizes because these LAI values provide hints on how far a camera can see. In addition, the DGPS employed for georeferencing field plots has a horizontal accuracy of about 5 m. The misregistration between the dependent (in-situ LAI) and independent (lidar metrics) variables also contributes to uncertainties in the fitted models, and the use of a relatively large plot size, e.g., 25 m, might alleviate such uncertainties to a certain degree.

The temporal discrepancy in the acquisition dates of the lidar and field data, i.e., one in the leaf-off season and another in the leaf-on season, unarguably causes an extra uncertainty in lidar estimates of LAI (Jensen et al., 2008; Næsset, 2005). However, the magnitude of this error was not investigated in this study and thus remains unclear. At first glance, such a time lag seemingly poses a limitation to this study, but the obtained results imply that it becomes less of an issue for two major reasons. First, because no natural or human-induced disturbance occurred over the study area during the period of this time lag, the changes in LAI for pines during this short period are deemed small or proportional to the initial LAI (Dewey et al., 2006); such types of changes only affect the exact coefficient values of models but not model accuracies. Second, the fitted lidar models, as have been validated by LOOCV, were applied in a consistent way with model calibration in that both calibration and application of the models convert the leaf-off lidar data into leaf-on LAI; such consistency guarantees that the model outputs are indeed what is desired (i.e., leaf-on LAI), although the model inputs are leaf-off data. In this regard, results of this study take on a special implication: Regression-based estimation of LAI or possibly other biophysical variables does not necessarily require ground-reference and remotely-sensed data being temporally coincident as long as the model fitted so is validated to have good performances and is fitted and applied consistently. In such a sense, depending on the purposes or accuracy requirements of applications, the same lidar dataset acquired at one time may be used to estimate LAI at various times if models are calibrated with the proper reference data; however, the resultant models may differ from one time to another with respect to model forms and predictors.

5.3. Integration of NDVI and lidar metrics

NDVI derived from the Quickbird image was found to be well correlated with in-situ LAI. The correlation strength suggests that NDVI could be helpful in predicting LAI but was less effective than lidar metrics (Griffin et al., 2008). However, computation of NDVI is scale-dependent due to both spatial heterogeneity and nonlinearity of relevant equations. NDVI2 showed a slightly better correlation with LAI than NDVI1. Thus, for practical applications that involve deriving coarse-scale NDVI from fine-scale multispectral images, it may be preferred to first aggregate the fine-scale images up to the coarse resolution and then compute NDVI from the aggregated spectral values. Meanwhile, future studies may investigate scaling effects of NDVI on estimation of biophysical parameters from high-resolution images using such techniques as that of Jiang et al. (2006). More important, we found that the inclusion of NDVI as an additional predictor in a linear form did not significantly improve the results over the lidar-based models. This accords with previous findings that the use of lidar metrics alone in regression models is generally sufficient to estimate LAI with no need to incorporate NDVI (Griffin et al., 2008; Jensen et al., 2008). Of particular note, this sufficiency does not exclude the usefulness of multispectral images to derive thematic information for developing stratum-specific lidar models, nor does it exclude the possibility that the combined use of lidar metrics and NDVI in more complicated model forms such as those established by
advanced machine learning algorithms may lead to improved performances as compared to the use of lidar metrics alone.

5.4. Comparison of GLOBCARBON and lidar LAI

GLOBCARBON LAI was observed to be highly correlated to the lidar-derived LAI over the study area. The obtained correlation coefficients rank high among the existing studies of validating moderate resolution satellite LAI products (Garrigues et al., 2008; Wang et al., 2004, 2005), and to our best knowledge, our result is more positive than those of early work on validating GLOBCARBON LAI (GLOBCARBON Demonstration Products and Qualification Report version 3.1). Undoubtedly, a major reason for this relatively high correlation is the compensation for a systematic sub-pixel shift of GLOBCARBON relative to the lidar LAI map. We suspect that this error is due to resampling satellite pixels to a pre-defined grid during the production of the LAI map from the raw satellite spectral measurements. The estimated magnitude of this shift error is around 0.60 km (60% of a GLOBCARBON pixel) in both x and y directions for our study area. Other mis-registration errors, either random or deterministic, could also influence the comparison of LAI (Tan et al., 2006) but have not been examined yet in this study. In this regard, we recommend that instead of directly using commercial software to co-register maps of significantly different spatial resolutions, future validation efforts should involve devising practical strategies to address mis-registration problems such as a systematic shift error, which, if not properly accounted for, can lead to an artificial and significant discrepancy in the comparison.

GLOBCARBON LAI is true LAI while the lidar-derived one is effective LAI. Even after applying the same stratum-specific clumping indices as used for producing GLOBCARBON LAI to our lidar-derived effective LAI, the absolute differences between the corrected lidar LAI and GLOBCARBON LAI were still large as revealed in Eq. (21). We do not know for certain how to explain away such inconsistency, but plausible explanations could be inferred from the equation fitted between the GLOBCARBON and corrected lidar LAI (Eq. (21)). Satellite sensors measure the integrated radiometric signals from within the full canopy layer but the hemiphotos only captured foliage above a certain height; therefore, leaves of certain understory layers that have been sensed by satellites are not included in the hemiphotos and lidar-derived LAI, causing a positive bias in GLOBCARBON LAI. A downward deviation of the fitted coefficient of Eq. (21) from the unity (i.e., 0.56) could be caused by the inclusion of non-foliage elements in the in-situ and lidar-derived LAI; this is because we did not differentiate between woody components and leaves in the hemi photo analysis whereas inversion algorithms used to produce satellite LAI often consider only foliage components in the models. Another potential reason for the discrepancy is the scale effect that the GLOBCARBON LAI was retrieved directly from coarse resolution images whereas the aggregated lidar LAI was obtained by averaging the fine-resolution LAI map. In general, it becomes more of a problem to directly compare moderate and fine resolution LAI maps unless they refer to the same physical quantity, for example, by correctly accounting for the clumping effect and scale incompatibilities.

Along with results of earlier studies, our findings urge the need not only for potential improvement on retrieval algorithms of moderate resolution satellite LAI products but also for more appropriate validation procedures (Cohen et al., 2003; Yang et al., 2006). In particular, efforts are entailed to develop more reliable LAI reference data. A commonly used approach to creating high-resolution LAI reference map is to rely on multispectral images to scale the in-situ LAI over local regions by either referring to statistical models or physically-based algorithms (Tan et al., 2005). One potential drawback of such an approach is the use of multispectral images which may by itself cause problems over medium to high biomass forests. To this end, the approach developed in this study to estimate LAI with lidar provides a competitive or even superior alternative for generating local fine-scale LAI maps that can be used as more reliable sources for validating satellite LAI products such as those of MODIS and GLOBCARBON.

6. Conclusions

Lidar holds great potential for estimating LAI with higher accuracies than multispectral images, particularly in forests of medium to high biomass. This study demonstrates the usefulness of several lidar metrics for predicting LAI in a pine-dominant eastern Texas forest. The investigated lidar predictors include a variety of laser penetration indices and canopy height-related metrics. Selection of model forms and lidar predictors is critical for ensuring good performances. In analogy to the Beer’s law, LPMs with logarithm models proved more effective than HRMs. Effectiveness of LPMs depends on not only plot sizes and types of laser hits used to compute LPM but also height thresholds selected to differentiate canopy and ground hits. The use of canopy height metrics alone shows only partial success for estimating LAI, but the product metrics resulting from the canopy height metrics times the foliage-density metrics also proved to be effective lidar surrogates for LAI.

Lidar-derived LAI maps can be more reliable than those derived from multispectral images to serve as reference data for validating coarse-resolution LAI products from other sources such as satellite images. Mis-registration between the lidar-derived and GLOBCARBON LAI maps could severely bias the pixel-by-pixel comparison. The use of a grid-searching procedure to accommodate a subpixel shifting mis-registration error considerably improved the correlation between the two maps, e.g., from 0.08 to 0.85 for 24 pine pixels; but their absolute differences still remain large, possibly due to the uncertainties in accounting for clumping effects.

At the current stage, it is still impractical to develop LAI estimation models from airborne lidar without recourse to ground-reference LAI measurements. Therefore, to enhance the utility of lidar for mapping LAI, besides testing more lidar predictors and models, future studies should also place an emphasis on collecting in-situ LAI data of higher accuracy, e.g., with the aid of the fast emerging ground laser scanning, to calibrate airborne lidar-based models. In addition, future studies may investigate the utility of lidar for creating other canopy structural characteristics as reference layers to validate relevant satellite products.

Acknowledgements

This research was supported by a grant from NASA New Investigator Program (NNX08AR12G). We thank Curt Stripling, all forestry personnel of the Texas Forest Service, Muge Mutlu and Alicia Griffin for their help with field data collection/compilation; Jared Stuky and Mark Karauach for their patient proofreading of the first draft of the paper; Drs. Ross Nelson, Yujie Wang, Xuesong Zhang, Marian Eriksson and X. Ben Wu as well as Mr. Borstlap Geert for helpful discussion. We also thank the four anonymous reviewers for their insightful comments.

References


